Production Externalities and Two-Way Distortion in Principal-Multi-Agent Problems

Ben Lockwood*, University of Warwick and CEPR
Department of Economics,
University of Warwick
Coventry CV4 7AL
b.lockwood@warwick.ac.uk

First Version: August 1997
This version: April 1999

Abstract

This paper studies an otherwise standard principal-agent problem with hidden information, but whether there are positive production externalities between agents: the output of any agent depends positively on the effort expended by other agents. It is shown that the optimal contract for the principal exhibits two-way distortion: the effort of any agent is oversupplied (relative to the first-best) when his marginal cost of effort is low, and undersupplied his marginal cost of effort is high. This pattern of distortion cannot otherwise arise in optimal single- or multi-agent incentive contracts, unless there are countervailing incentives. However, unlike the countervailing incentives case, the pattern of distortion is robust to the precise form of the externality.

*I would like to thank an anonymous referee for very valuable comments.

1. Introduction

There is now a considerable literature on the principal-agent problem with multiple agents, both with hidden action and hidden information. Multiple-agent problems only di¤er from single-agent problems if there is some interaction between the agents. The two main forms of interaction that are of interest are …rst, production externalities between agents (the output of a particular agent depends on the e¤ort of other agents), and second, statistical correlation in the environments of the agents.

Most of the literature so far has focussed on the second kind of interaction. For example, in the hidden information case, a literature, starting with Demski and Sappington[4], has focussed on the implications of (positive) correlation of the cost of e¤ort of two agents. It turns out that, even in the case where agents are identical ex ante, before their private information is revealed, the optimal contract for the principal treats the agents asymmetrically, recruiting one of the agents as a "policeman", who can report on the type of the other agent (Demski and Sappington[4], Glover[6], Ma, Moore and Turnbull[13]). Again, in principal-agent problems with hidden actions, attention has focused on the case where the production functions that map e¤orts of the agents into outputs are subject to correlated random disturbances. In this setting, comparative compensation contracts, such as contests, may be optimal (Mookherjee[16], Nalebu¤ and Stiglitz[17]).

By contrast, the implications of production externalities for the contract design have little studied[1] in the hidden action case, and not at all (to my knowledge) in the hidden information case. This paper presents an analysis of a principal-multi-agent model with hidden information where there are (positive) production externalities between agents. The main …nding is that, in the optimal contract for the principal, the distortions that arise relative to the …rst-best are quite novel: they cannot arise in principal-multi-agent models without production externalities, even with correlated costs, unless the reservation utilities of the agents vary with their cost-type in such a way that agents face countervailing incen-

[1] Exceptions are Che and Yoo[2], Itoh[9], Kandal and Lazear [10], and Mookherjee[16], all of whom consider hidden action models. However, none of these papers is very close to this one, for reasons explained in detail in the conclusions.

tives in revealing cost information to the principal[2]. Here, we abstract from countervailing incentives (by assuming that all agents have a reservation utility of zero) and show that nevertheless, there is two-way distortion in output: an agent will choose an ine¢ciently low value of output for some values of his private information, and an ine¢ciently high value for other values.

The basic principal-agent model studied here is one where a number of agents choose e¤ort to produce outputs: the output of agent i depends not only on his own e¤ort, but also positively on the average e¤ort made by all other agents (the externalities are those studied by Cooper and John[5] under the heading of "input games"). The basic model is extended in Section 5 to allow for a richer structure of production externalities.

The marginal cost of e¤ort to agent i is parameterized by a variable μ_i, which is i^0s private information, and the μ_i are independently distributed[3]. There are no countervailing incentives; reservation utility is zero for all agents. A contract o¤ered by the principal to each agent is a choice of output and a monetary transfer, conditional on the vector of reported μs from all agents.

We study incentive-compatible contracts for the principal in this setting, i.e. contracts where it is either a dominant or Nash equilibrium strategy for each agent to tell the truth. We assume that the number of agents is "large"; under this condition, we show that the principal is no worse o¤ o¤ering a dominant-strategy incentive-compatible contract than a Nash incentive-compatible one (see Proposition 1), and so we can without loss of generality consider just the former class of contracts[4]. We show that, under quite weak conditions[5],

[2] See Lewis and Sappington[12], Maggi and Rodriguez-Clare[15]. We discuss this literature in more detail in Section 7.

[3] In fact, we assume a measure space of agents, where the distribution of marginal costs across agents is common knowledge. However, this can be interpreted as the limiting case of a model with a …nite number of agents, where the marginal costs of agents are independently and identically distributed.

[4] This is in contrast to the literature on two-agent models with private (correlated) information, where the dominant-strategy incentive constraints are more restictive than the Bayes-Nash incentive constraints.

[5] See Theorem 1 below for a full statement of su¢cient conditions; these comprise standard conditions on cost and revenue functions, plus some weak conditions on the spillovers between agents, and …nally the requirement that the cost function must be separable in the agent's e¤ort and cost parameter, and convex in the latter.

for the principal's optimal contract in this class, the output of agent i is oversupplied for low values of μ_i, and undersupplied for high values of μ_i.

The intuition is simple. First, as in the standard principal-agent model, for any value of μ_i, the informational rent captured by any agent i is increasing in the effort put in by that agent. It follows from this that informational rent captured by any agent i is decreasing in average output of agents $j \neq i$, as an increase in the average output of agents $j \neq i$ decreases the amount of effort agent i needs to put in to produce a given output. So, there is an interaction between the production externality and informational rent. This interaction means that the principal has an additional incentive (over and above the production externality) to raise the output of any agent i. This incentive co-exists with the standard incentive -absent the externality - for the principal to restrict agent i's output in order to reduce agent i's own informational rent, and so two-way distortion is the outcome.

So, one way of expressing this intuition is to observe that in the setting of this paper, the principal, rather than the agents, faces countervailing incentives; that is, he faces incentives both to lower and raise the output of any particular agent relative to the first-best. This intuition also relates to the general point, made e.g. by Sappington[18], that a principal may introduce distortions in other instruments to better limit agents' rents. In this case, the extra instrument that the principal has, when facing any particular agent, is the (average) output of other agents. Seen in this way, the main contribution of this paper is to establish the precise pattern of the distortion in the other instrument.

A second notable feature[6] of the optimal contract is that the transfer from principal to agent has a yardstick property i.e. the transfer to some agent i is (at some point) decreasing in the output of the other agent(s). This is the case even though the types of the agents are uncorrelated, so the principa1 cannot exploit the correlation between agents'

[6] It is of course, well-known that in principal-multi-agent schemes with statistical correlation of costs across agents, comparative compensation of the agents is often optimal. As shown by Nalebuff and Stiglitz[17], and Mookherjee[16], contests (where agents are compensated only on the basis of the ordinal ranking of their outputs) are sometimes optimal, and contests certainly have the yardstick property. Moreover, necessary and sufficient conditions for contracts to be independent are very strong (Mookherjee[16]).

types to extract additional informational rents, as in Cremer and McLean[3], Demski and Sappington[4].

The arrangement of the rest of the paper is as follows. The model is presented in Section 2, and dominant-strategy incentive-compatible contracts are characterized in Section 3. The main results on two-way distortion are presented in Section 4. Section 5 extends these results to a richer class of production externalities. Section 6 discusses the yardstick property of the optimal contract, and Section 7 discusses the related literature, especially the work on countervailing incentives an principal-multi-agent problems, and concludes.

2. The Model

The model is an otherwise standard principal-multi-agent model with production externalities between agents. It is analytically convenient (for reasons explained in the next section) to work with a "large" number of agents. Let the space of agents be (I, S, μ), where $I = [0, 1]$; S is the Borel σ-algebra on I, and μ is the Lebesque measure. Every agent provides effort level $e_i \in \Re_+$ at cost

$$c_i = c(e_i, \theta_i)$$

where $\theta_i \in [\underline{\theta}, \overline{\theta}] = \Theta$ parameterizes i's cost of effort. We assume the following properties of $c(.,.)$;

A1. $c_e, c_\theta, c_{\theta e}, > 0, \;\; c_{ee}, c_{\theta ee}, c_{\theta\theta e} \geq 0.$

These inequalities include the standard assumptions of positive and increasing marginal cost of effort, and the single-crossing condition $c_{\theta e}$. A special case that satisfies A1 is $c(e, \theta) = \theta e$. The parameter θ_i is private information of agent i.

The agent also receives a transfer $t_i \in \Re$ of a numeraire good from the principal, so his utility from the pair (e_i, t_i) is

$$u_i = t_i - c(e_i, \theta_i) \tag{1}$$

Every agent has a reservation utility of zero, so there are no countervailing incentives for agents.

Spillovers are specified as follows. We suppose that the agents are all engaged in production processes, where the marginal product of any agent's effort is affected by the average effort of the others, $e = \int_I e_i d\mu$. In other words, the set of agents I is a team, and the team production technology is such that the spillover for agent i from the effort of agent $j \neq i$ is the same as from the effort of any other agent $k \neq i$. This is a natural simplifying assumption often made when studying games with production externalities and large numbers of players (e.g. Cooper and John's [5] "input games"). It is relaxed in Section 5.

Following Cooper and John[5], we suppose the externality takes the following form:

$$q_i = e_i g(e) \tag{2}$$

where q_i is output of i. We assume that g is twice continuously differentiable, and that it satisfies:

A2. $g(e), g'(e) > 0, e \in \Re_+$.

These assumptions are reasonable: $g' > 0$ says that the spillover is positive, and $g > 0$ requires in particular that any agent i can produce even if all others do not i.e. $g(0) > 0$. We choose units so that $g(0) = 1$. An example satisfying A2 is $g = 1 + e^{\alpha}$, $0 < \alpha \leq 1$.

The production function (2) implies that the cost to agent i in effort units of producing q_i also depends on aggregate output q. First, integrating over all the agents, (2) implies

$$q = eg(e) \tag{3}$$

where $q = \int_I q_i d\mu$. Then, as $g'(e) > 0$, $e \in \Re_+$, the relationship (3) can be inverted on $\Re_+$ to give $e = \gamma(q)$. But then from (2), we can write

$$e_i = q_i s(q); \ s(q) = \frac{1}{g(\gamma(q))} \tag{4}$$

So, $q_i s(q)$ is the amount of effort required for agent i to produce output q_i. Note that $s'(q) = -s^2 g' \gamma' < 0$ as both $g', \gamma' > 0$. That is, the higher aggregate output, the lower the effort required for i to produce some fixed output q_i. Note also that as we have assumed $g(0) = 1$, $s(0) = 1$ also.

The output of agent i generates revenue for the principal of $r(q_i)$, where $r(\cdot)$ is strictly increasing and strictly concave. The idea here is that agent's outputs are differentiated and sold in separate markets[7]. The principal keeps the aggregate revenue net of payments, and so gets profit

$$\pi = \int_I [r(q_i) - t_i] d\mu \tag{5}$$

Following Demski and Sappington[4], we define a contract[8] as a compensation-output pair (t_i, q_i) for each agent $i \in I$ as a function of all the cost announcements $\hat{\theta} = \left(\hat{\theta}_i\right)_{i\in I}$. As all agents are ex ante identical, we can focus on anonymous contracts where (t_i, q_i) depends only on $\hat{\theta}_i$ and the distribution of announced characteristics. These contracts are defined formally in the next section.

The order of events is now as follows. First, the principal chooses an anonymous contract. Then, every agent $i \in I$ simultaneously announces a type $\hat{\theta}_i \in \Theta$. Finally, production takes place and transfers are made.

3. Incentive-Compatible Contracts

We begin by defining anonymous contracts. We assume that $\theta_i : I \to \Theta$ is a measurable function. Consequently, we can define the measure ν on Θ by $\nu(A) = \mu(\{i \in I | \theta_i \in A\})$, for all A in the Borel σ-algebra on Θ; so, ν is the distribution of (true) costs on Θ. Also, let $\Psi \in P(\Theta)$ be the distribution of announced costs on Θ; that is, $\Psi(A) = \mu\left(\left\{ i \in I \middle| \hat{\theta}_i \in A \right\}\right)$ for all Borel sets $A \subset \Theta$. Obviously, if all agents tell the truth, then $\Psi = \nu$. Note that[9]

---

[7] An alternative assumption would be that the agents' outputs are identical, and so sold in the same market, in which case revenue would be $r = r(\int_I q_i d\mu)$. In this case, the analysis is exactly the same, except we must strengthen A1 slightly by imposing $c_{ee} > 0$ to ensure an unique solution to problem P below.

[8] It would of course be possible in principle to have contract where the principal chooses a compensation-input pair (t_i, e_i). However, following much of the principal-agent literature, we suppose that effort is non-contractible (Hart[7]).

[9] For this to be the case, we require the sets $\{i \in I | \theta_i \in A\}$, $\left\{ i \in I \middle| \hat{\theta}_i \in A \right\}$ to be measurable with respect to μ for all Borel sets $A \subset \Theta$. The first sets are all measurable as the map $f(i) = \theta_i$ is assumed measurable with respect to μ. The second sets are all measurable if, in turn, the "announcement function" h (i.e. $h(i) \equiv \hat{\theta}_i$) mapping I into Θ is assumed measurable with respect to μ. In the announcement game, agents are restricted to play anonymous strategies i.e. agent i announces $\hat{\theta}_i = \sigma(\theta_i)$, where $\sigma : \Theta \to \Theta$ is measurable with respect to ν.(see below). So, the "announcement function" is the composition $h = \sigma \circ f$, where $f(i) = \theta_i$. As, both σ, f are measurable by assumption, so is h (Hildenbrand [8], p42).

$\nu, \hat{\nu} \in P(\Theta)$, where $P(\Theta)$ is the set of (Borel) probability measures on Θ. We now have:

Definition 1. An anonymous contract is a pair of functions $t : \Theta \times P(\Theta) \to \Re$, $q : \Theta \times P(\Theta) \to \Re_+$ where agent i is offered $(t_i, q_i) = (t(\hat{\theta}_i, \hat{\nu}), q(\hat{\theta}_i, \hat{\nu}))$ iff he announces a type $\hat{\theta}_i$ and distribution of announced costs is $\hat{\nu}$.

So, with an anonymous contract, (t_i, q_i) depends only on i's announced cost and the distribution of announced costs $\hat{\nu}$. Consequently, the payoff to any agent $i \in I$ with $\theta_i = \theta$ who makes a cost announcement $\hat{\theta}$ depends only on $\theta, \hat{\theta}, \hat{\nu}$;

$$u(\theta, \hat{\theta}, \hat{\nu}) \equiv t(\hat{\theta}, \hat{\nu}) - c(q(\hat{\theta}, \nu)s(\hat{q}), \theta), \quad \hat{q} = \int_{z \in \Theta} q(z, \hat{\nu})d\hat{\nu} \tag{6}$$

Note that $\hat{q}$ is the average output across all agents, given a distribution $\hat{\nu}$ of announced characteristics.

So, given a fixed anonymous contract, the agents play an "announcement" game. As any agent's utility depends only on his own action θ and the aggregate distribution of actions, this is an anonymous game (Mas-Colell[14]). It is therefore natural to restrict i's announcement to depend only on his cost characteristic. So, following Mas-Colell[14], we assume that a strategy profile in the announcement game is a measurable function $\sigma : \Theta \to \Theta$, where i's strategy is $\sigma(\theta_i) = \hat{\theta}$ iff $\theta_i = \theta$.

We can now define dominant-strategy and Nash incentive-compatible contracts.

Definition 2. An anonymous contract is dominant-strategy incentive-compatible iff

$$u(\theta, \theta, \hat{\nu}) \geq u(\theta, \hat{\theta}, \hat{\nu}), \text{ all } \theta, \hat{\theta} \in \Theta, \text{ all } \hat{\theta} \neq \theta, \text{ all } \hat{\nu} \in P(\Theta) \tag{7}$$

That is, truth-telling ($\sigma(\theta) = \theta$) is a dominant-strategy for any agent in the announcement game. Nash incentive-compatible contracts are defined similarly;.

Definition 3. An anonymous contract is Nash incentive-compatible iff

$$u(\theta, \theta, \nu) \geq u(\theta, \hat{\theta}, \nu), \text{ all } \theta, \hat{\theta} \in \Theta, \text{ all } \hat{\theta} \neq \theta \tag{8}$$

That is, truth-telling is a Nash equilibrium in the announcement game; it is best for any agent to tell the truth, given the distribution of announced costs is the true one.

Our first result gives conditions under which a contract is dominant-strategy or Nash incentive-compatible. This, and all subsequent results, are proved in the Appendix.

Proposition 1. A contract (t_D, q_D) is dominant-strategy incentive- compatible iff it satisfies

$$t_D(\hat{\theta}, \hat{\nu}) = c(q_D(\hat{\theta}, \hat{\nu})s(\hat{q}), \hat{\theta}) + \int_{\underline{\theta}}^{\hat{\theta}} c_\theta(q_D(z, \hat{\nu})s(\hat{q}), z)dz + A_D, \; A_D \in \mathbb{R} \tag{9}$$

for all $(\hat{\theta}, \hat{\nu}) \in \Theta \times \mathcal{P}(\Theta)$, where $\hat{q}$ is defined as above, and

$$\frac{\partial q_D(\hat{\theta}, \hat{\nu})}{\partial \hat{\theta}} \leq 0 \text{ almost everywhere on } \Theta \tag{10}$$

Moreover, there is a dominant-strategy incentive-compatible contract that yields the principal the same payoff as her highest payoff from the Nash incentive-compatible contract.

Proposition 1 indicates that the principal can restrict attention to dominant-strategy incentive-compatible contracts. So, we drop the "D" subscript on q_D, t_D without loss of generality. The dependence of this pair on the measure of announced characteristics, $\hat{\nu}$, is suppressed below for brevity except where appropriate, so we may write an anonymous contract simply as $(t(\theta), q(\theta))_{\theta \in \Theta}$.

It is also a result of independent interest for the following reason. It is well-known that when the number of agents is finite (e.g. with two agents) in problems of this types, the principal can generally do better with Nash incentive-compatible contracts than with dominant-strategy contracts, under the assumption that the truth-telling equilibrium prevails, as the constraints placed on contract design are less demanding (Demski and Sappington [4]).

The key assumption that generates this equivalence for the principal is that the number of agents is "large", not any of the other assumptions of the model. For then, from the point of view of any particular agent $i \in I$, the behavior of other players in the "announcement" game is non-stochastic in the aggregate i.e. every player faces a fixed distribution of announcements $\hat{\nu}$. To see this, suppose that we have a general production technology where the spillover s depends on the entire distribution of output, χ, not just the average q. Then, inspection of the proof of Proposition 1 reveals that the result goes though as before, where $s(\chi)$ replaces $s(q)$.

4. Contract Design and Two-Way Distortion

The problem faced by the principal is to choose a (dominant-strategy) incentive-compatible contract to maximise his profit, defined in (5), from among the class of such contracts. The problem can be formulated as follows. First, let $F : \Theta \to [0; 1]$ be the distribution function of costs defined as $F(x) = \nu([\underline{\theta}; x])$, and suppose that $F(.)$ is absolutely continuous, with density $f(.) > 0$. Also, let $w(\theta) \equiv u(\theta; \theta; \nu)$. Then the principal's payoff is;

$$\begin{aligned}
\pi &= \int_{\underline{\theta}}^{\bar{\theta}} [r(q(\theta)) - t(\theta)] f(\theta)\, d\theta & (11)\\
&= \int_{\underline{\theta}}^{\bar{\theta}} \left[r(q(\theta)) - c(q(\theta)s(q); \theta) + \int_{\theta}^{\bar{\theta}} c_{\theta}(q(z)s(q); z) dz \right] f(\theta)\, d\theta - w(\bar{\theta}) \\
&= \int_{\underline{\theta}}^{\bar{\theta}} [r(q(\theta)) - \psi(q(\theta)s(q); \theta)] f(\theta)\, d\theta - w(\bar{\theta})
\end{aligned}$$

In the second line we have used (9), and the fact that $A_D = w(\bar{\theta})$, as shown in the Appendix, where $w(\theta) \equiv u(\theta; \theta; \nu)$. In the third line, we have integrated by parts, and finally

$$\psi(e; \theta) = c(e; \theta) + \frac{1}{h(\theta)} c_{\theta}(e; \theta) \qquad (12)$$

where $h(\theta) = f(\theta)/F(\theta)$ is the hazard rate[10] for the distribution of θ. So, $\psi(e; \theta)$ has an obvious interpretation as the perceived cost, from the principal's point of view, of extracting output $q(\theta)$ from a type-θ when aggregate output is q. The second term in (12) is the informational rent accruing to the agent and is positive by A1, so the perceived cost always strictly exceeds the true cost ($\psi(e; \theta) \geq c(e; \theta)$), and does so strictly unless $\theta = \underline{\theta}$.

The principal therefore solves[11] the following problem:

$$\begin{aligned}
&\max_{q(.) \geq 0} \int_{\underline{\theta}}^{\bar{\theta}} [r(q(\theta)) - \psi(q(\theta)s(q); \theta)] f(\theta)\, d\theta \text{ s.t.} \\
q'(\theta) &\leq 0 \\
q &= \int_{\underline{\theta}}^{\bar{\theta}} q(\theta) f(\theta) d\theta
\end{aligned}$$

[10] $h(x)$ can be interpreted as the approximate conditional probability (for small Δ) that cost parameter θ does not fall below $x - \Delta$ given that it has already fallen from $\bar{\theta}$ to x (Laffont and Tirole[11], p66).

[11]. Note that the choice of $q; t$ must also ensure that the agent participation constraints $w(\theta) \geq 0$ are satisfied. First, as $w' < 0$ by standard arguments, the only potentially binding participation constraint is $w(\bar{\theta}) \geq 0$. As θ is decreasing in $w(\bar{\theta})$; it is immediately obvious that the principal sets $w(\bar{\theta}) = 0$.

Call this problem P. Even in the absence of the monotonicity constraint $q'(\theta) \leq 0$, this is not a concave problem, due to the presence of externalities in the perceived cost function. Denote by $(q^*(\theta))_{\theta \in \Theta}$ a solution to P. We will say that a solution to P is *interior* if $0 < q^*(\theta) < \infty$, all $\theta \in \Theta$.

We can characterize the solution to P under the following assumption which ensures an interior solution;

A3. $r'(0) > \psi_e(0;\theta)$; $\lim_{q \to \infty} r'(q) = 0$; $\lim_{e \to \infty} c_e(e;\theta) = \infty$:

Assumption A3 imposes quite standard Inada-type conditions on revenue and cost functions.

Proposition 2. *If A1-A3 hold, and the monotone hazard rate condition $h'(\theta) \leq \theta;\ \theta \in \Theta$ holds, then there exists an interior solution to problem P, and at this solution, $q(\theta)$ solves*

$$r'(q(\theta)) = \psi_e(q(\theta)s(q);\theta)s(q) + E\left[\psi_e(q(\theta)s(q);\theta)q(\theta)\right]s'(q) \qquad (13)$$

where the expectation is taken with respect to θ:

Note that (13) equates the marginal revenue generated by an increase in $q(\theta)$ to the perceived marginal cost to the principal - taking into account informational rent and the production externality - of an increase in $q(\theta)$. The first term on the right-hand side of (13) is the internal marginal cost of raising $q(\theta)$ incrementally, and the second term (which is negative, as $s' < 0$) is the external marginal benefit of raising $q(\theta)$ in terms of reduced costs for all agents.

We can now turn to analyze the distortions induced by the presence of both informational rent and externalities at the solution to problem P, and which are implicit in the first-order condition (13). The benchmark is the full-information case, where the principal can observe the cost parameter of each agent. In this case, the principal sets marginal benefit of an increment in $q(\theta)$ equal to true marginal cost, ignoring informational rent i.e. we replace the perceived marginal cost function in (13) by the true one to get

$$r'(q(\theta)) = c_e(q(\theta)s(q);\theta)s(q) + E\left[c_e(q(\theta)s(q);\theta)q(\theta)\right]s'(q) \qquad (14)$$

Again, the first term on the right-hand side of (14) is the internal marginal cost of raising $q(\theta)$ incrementally, and the second is the external marginal benefit of raising $q(\theta)$.

Compare (13) to (14) first for the familiar case without externalities. In this case, we can take $s(q) \equiv 1;\ s'(q) \equiv 0$. Then, the full-information and incentive-compatible first-order conditions are

$$r'(q(\theta)) = c_e(q(\theta); \theta) \tag{15}$$

$$r'(q(\theta)) = c_e(q(\theta); \theta) + \frac{1}{h(\theta)} c_{\theta e}(q(\theta); \theta) \tag{16}$$

respectively. So, inspection of (15),(16), plus the fact that $c_{\theta e} > 0$ from A1, indicates that without externalities, when θ is private information, marginal cost is "too high", due to the presence of informational rents, and effort is undersupplied for all values of θ except the lowest: This is a standard result (Laffont and Tirole [11]).

In the general case, by reference to (14), we have the following definition.

Definition 4. Output is oversupplied by an agent of type θ if

$$r'(q(\theta)) < c_e(q(\theta)s(q); \theta)s(q) + E\left[c_e(q(\theta)s(q); \theta)q(\theta)\right] s'(q)$$

and undersupplied by an agent of type θ if

$$r'(q(\theta)) > c_e(q(\theta)s(q); \theta)s(q) + E\left[c_e(q(\theta)s(q); \theta)q(\theta)\right] s'(q)$$

So, with oversupply, marginal revenue of an increment in output is below the marginal cost of an increment in output (taking into account spillover effects), and conversely, with undersupply, it is above. We now have the main result of the paper.

Theorem 1. Assume A1-A3 hold, the monotone hazard rate condition holds, and that $c(\theta; e) = \gamma(\theta)c(e)$; with $\gamma'' \geq 0$: Then, there is two-way distortion in the solution to problem P. That is, there exists $\underline{\theta} < \theta' < \bar{\theta}$ such that for $\theta \in [\underline{\theta}; \theta')$, effort is oversupplied, and for $\theta \in (\theta'; \bar{\theta}]$, effort is undersupplied.

This result can be interpreted as follows. At the solution to P, the principal always equates marginal revenue to perceived marginal cost $\psi_e s(q) + E[\psi_e q(\theta)]s'(q)$: So, Theorem 1 says that when θ is high, perceived marginal cost is greater than true marginal cost, and when θ is low, perceived marginal cost is less than true marginal cost. This is to be compared to the standard case without externalities, where perceived marginal cost is greater than true marginal cost for all θ. So, the new insight here is that with production externalities, when θ is low, perceived marginal cost is less than true marginal cost, even though perceived total cost is always greater then true total cost; it is this that generates the two-way distortion.

The intuition for this new result is as follows. From (13), (14), the difference between the perceived and true marginal cost of output is

$$\frac{c_{\theta e}(q(\theta)s(q);\theta)}{h(\theta)}s(q) + E\left[\frac{c_{\theta e}(q(\theta)s(q);\theta)}{h(\theta)}q(\theta)\right]s'(q) \qquad (17)$$

The first term in (17) is due to informational rent, and is always positive. The second term is negative as $s' < 0$. It captures the effect that an increase in $q(\theta)$ has on the information rent accruing to other agents via the spillover. Specifically, a small increase Δ in $q(\theta)$ leads to a reduction $\Delta s'(q)$ in the effort required by all agents, and this in turn leads to a reduction of

$$E\left[\frac{c_{\theta e}(q(\theta)s(q);\theta)}{h(\theta)}q(\theta)\right]\Delta s'(q)$$

in the informational rent captured by these agents. Whether the perceived marginal cost of output is above or below the true marginal cost depends on the relative magnitude of these two terms. When $\theta \approx \underline{\theta}$, $1/h(\theta) \approx 0$, and so the second term in (17) dominates the first term, implying that the perceived marginal cost of output is below the true marginal cost, and leading in turn to oversupply by our definition.

We now comment on the sufficient conditions for two-way distortion. First, assumptions A1-A3 are not at all restrictive. Assumption A1 is quite standard in the principal-agent literature. Assumption A2 imposes weak and reasonable conditions on the spillover function g, and A3 imposes quite standard Inada-type conditions on r and c. Finally, the condition that c be separable in e, θ and convex in θ is quite weak.

5. Multiple Teams

Probably the main restriction of the model of this paper is that only the aggregate effort of agents affects the marginal cost of effort of any particular agent. One simple way of relaxing this assumption somewhat is to suppose that there are two groups of agents, or teams, $I_l;\ l = a, b$, with $I_a \cup I_b = I;\ I_a \cap I_b = \emptyset$. Then, it is natural to suppose that the aggregate effort of team a; e_a, has some impact on the productivity of a member of team b, but less than the effect it has on the productivity of a member of team a. This can be captured formally by writing

$$q_i^a = e_i^a g(e^a + \sigma e^b);\ q_i^b = e_i^b g(e^b + \sigma e^a))$$

where superscripts denote team membership, and $0 < \sigma < 1$ measures the between-team spillover, which is less than the within-team spillover as $\sigma < 1$. Also, we assume that $g(.)$ satisfies A2. Using the identity $e^l = \int_{i \in I_l} e_i^l d\mu$; we have

$$q^a = e^a g(e^a + \sigma e^b) \tag{18}$$

$$q^b = e^b g(e^b + \sigma e^a) \tag{19}$$

Now, it is easily checked that the Jacobian of the system (18),(19) is non-singular on $\mathbb{R}_+^2$ (see e.g. (21 below), so we can invert (18),(19) to get

$$e^a = \phi^a(q^a, q^b)$$

$$e^b = \phi^b(q^a, q^b)$$

By the symmetry of technology, $\phi^a(x, y) \equiv \phi^b(y, x)$. Let $\phi_j^a(q^a, q^b)$ denote the derivative of ϕ^a with respect to its jth argument; $j = a, b$; and the same for ϕ^b. For future reference, note that

$$\phi_i^i = \frac{1}{D}[g_j + e^j g_j'] > 0;\ \phi_j^i = \frac{-\sigma e^i g_i'}{D} < 0 \tag{20}$$

where $g_i = g(e^i + \sigma e^j)$; and

$$D = (g_a + e^a g_a')(g_b + e^b g_b') - \sigma^2 e^a g_a' e^b g_b' > 0 \tag{21}$$

So, from (20), an increase in output by team a requires an increase in effort by team a, but an increase in output by team b allows members of team a to reduce their effort, while producing a constant output, due to the inter-team spillover.

Now, for a member i of team a, we can define

$$e_i^a = \frac{q_i^a}{g(\phi^a(q^a, q^b) + \sigma\phi^b(q^a, q^b))} = q_i^a s^a(q^a, q^b) \tag{22}$$

and $s^b(q^a; q^b)$ can be defined similarly. By the symmetry of technology, $s^a(x; y) \equiv s^b(y; x)$. Let $s^a_j(q^a; q^b)$ denote the derivative of s^a with respect to its jth argument; $j = a; b$. So, $s^a_j(q^a; q^b)$ measures the change in the effort required by $i \in I_a$ to produce one unit of output, when q^j increases. When $q^a \simeq q^b$, this change is negative[12] for $j = a; b$.

The preferences of both principal and agents are as before; any agent $i \in I_l$ has a cost of effort function $c(\theta_i; e_i)$ which satisfies A1 and A3, and the agents maximise their transfer from the principal net of the cost of effort. As before, the output of any agent $i \in I_l$ generates revenue $r(q_i)$ for the principal, where r satisfies A3, and the principal wishes to maximize the sum across teams of revenue minus transfers.

In the multi-team case, an anonymous contract for team l is defined as above i.e. as a pair of functions $t_l : \Theta \times P(\Theta) \times P(\Theta) \to \Re;\ q_l : \Theta \times P(\Theta) \times P(\Theta) \to \Re_+$ where agent $i \in I_l$ is offered $(t_i; q_i) = (t_l(\hat{\theta}_i; \hat{\gamma}); q_l(\hat{\theta}_i; \hat{\gamma}))$ iff he announces a type $\hat{\theta}_i$ and the distribution of announced costs for both teams is $\hat{\gamma} = (\hat{\gamma}_a; \hat{\gamma}_b)$.

Then it is easy to check that Proposition 1 goes though, modified in the obvious way i.e. the principal can do no better with a Nash incentive-compatible contract than with a dominant-strategy one, and the transfer to a member of team $l = a; b$ who reports $\hat{\theta}$ is

$$t_l(\hat{\theta}; \hat{\gamma}) = c(q_l(\hat{\theta}; \hat{\gamma})s^l(\hat{q}_a; \hat{q}_b); \hat{\theta}) + \int_{\hat{\theta}}^{\bar{\theta}} c_\theta(q_l(\hat{\theta}; \hat{\gamma})s^l(\hat{q}_a; \hat{q}_a); z)dz$$

where $\hat{q}_l = \int_{I_l} q_i d\mu$. Now assume that the two teams are identical in size ($\mu(I_a) = \mu(I_b)$), and in the distribution of costs across group members ($\gamma_a = \gamma_b$). So, the distribution function of costs in either team is F, with density f. As before, suppress the dependence of $q_l(.)$ on γ. The principal therefore solves the following problem P':

$$\max_{q_a(.);q_b(.) \geq 0} \sum_{l=a;b} \int_{\underline{\theta}}^{\bar{\theta}} \left[r(q_l(\theta)) - \psi(q_l(\theta)s^l(q_a; q_b); \theta) \right] f(\theta)\, d\theta \text{ s.t.}$$

$$q_l'(\theta) \leq 0;\ l = a; b$$

$$q_l = \int_{\underline{\theta}}^{\bar{\theta}} q_l(\theta) f(\theta) d\theta;\ l = a; b$$

Due to the symmetry of the problem, we focus on the class of symmetric solutions to P' where $q_a(.) = q_b(.) = q(.)$. Under assumption A3, there will be an interior symmetric solution to this problem i.e. Proposition 2 extends, and the first-order condition characterizing

[12] See (A.23) in the Appendix.

$q(\cdot)$ is

$$r'(q(\theta)) = \psi_e(q(\theta)s^a(q;q);\theta)s^a(q;q) + E\left[q(\theta)\psi_e(q(\theta)s^a(q;q);\theta)\right](s^a_a(q;q) + s^a_b(q;q)) \quad (23)$$

Also, note that Definitions 3 and 4 of undersupply and oversupply carry over directly to this case. We then have the following extension of the main result to the multi-team case;

Theorem 2. Assume that the assumptions of Theorem 1 on r, c, and F hold. Then, there is two-way distortion in the solution to problem P'. That is, there exists $\underline{\theta} < \theta' < \bar{\theta}$ such that if $\theta \in [\underline{\theta}, \theta')$, effort is oversupplied, and for $\theta \in (\theta', \bar{\theta}]$, effort is undersupplied.

It seems likely that a version of this result could be proved for the case of n teams, although the statement and proof would be cumbersome. So, our two-way distortion result does not depend crucially on the precise form of the externality between agents.

6. Yardstick Transfers

So far, we have restricted attention to contracts where agents directly report their types (direct mechanisms, in the parlance of the implementation literature). In practice, principals generally use contracts where the transfer from principal to agent(s) depends on output, rather than a reported type (indirect mechanisms). However, the class of incentive-compatible contracts described in Proposition 1 can easily be written in this form.

Let $\hat{q}^{-1}$ be the inverse of $q_i = \hat{q}(\theta_i)$; this inverse always exists as the monotonicity condition is satisfied by assumption of a monotone hazard rate (Proposition 2). Now consider the transfer schedule

$$t(q_i;q) \equiv t(\hat{q}^{-1}(q_i);q) = c(q_i s(q); \hat{q}^{-1}(q_i)) + \int_{\hat{q}^{-1}(q_i)}^{\bar{\theta}} c_\theta(q_i s(q); z)dz \quad (24)$$

Note also that the transfer schedule (24) satisfies the yardstick property, as defined in the introduction; namely, that the transfer to some agent i is decreasing in the output of other agent(s). To see this, differentiate to get

$$\frac{\partial t(q_i;q)}{\partial q} = s'(q)q_i[c_e(q_i s(q); \hat{q}^{-1}(q_i)) + \int_{\hat{q}^{-1}(q_i)}^{\bar{\theta}} c_{\theta e}(q_i s(q); z)dz] < 0 \quad (25)$$

7. Conclusions and Related Literature

This paper has shown that in an otherwise standard principal-agent problem with hidden information, the presence of positive production externalities between agents leads, under quite general conditions, to two-way distortion, with the output of any agent i being oversupplied when his marginal cost of e¤ort is low, and undersupplied when his marginal cost of e¤ort is high. As remarked in the Introduction, two-way distortion cannot arise in principal-multi-agent models with hidden information of the type studied in the literature[13], and a fortiori, it cannot arise in the standard single-agent case.

The literature related to the analysis of this paper is small. There is to my knowledge, no work that studies production externalities in principal-multi-agent models with hidden information. There are a small number of papers which allow for production externalities in principal-multi-agent models with hidden actions [Che and Yoo[2], Itoh[9], Mookherjee[16], Kandal and Lazear[10]]. However, Che and Yoo[2] and Mookherjee[16], are concerned entirely with the study of the cost-minimization problem for the principal (characterizing the minimum cost of inducing a given pair of actions by the two agents), and do not discuss the issue of whether actions that are then chosen by the principal are above or below their …rst-best levels.

Itoh [9] studies choice of e¤ort level as well as the cost-minimization problem given e¤ort levels, but his focus is rather di¤erent. Speci…cally, each of two agents can choose not only an e¤ort level that enhances the success probability of his own project, but also the level of another e¤ort variable ("helping" e¤ort) that enhances the success probability of the other agent's project. The main objective of his paper is to establish conditions under which the principal will choose a positive level of "helping" e¤ort in the incentive-compatible contract. By contrast, in the model of this paper, e¤ort is one-dimensional, but has a joint product; it enhances the output not only of the agent who exerts it, but other agents.

Finally, Kandal and Lazear[10] allow for general production spillovers, but they do not characterize the principal's optimal incentive-compatible contract. Rather they study a

[13] As shown by Demski and Sappington[4], and Ma, Moore, and Turnbull [13], output is always undersupplied, whether the announcement game equilibrium is in dominant or Bayes-Nash strategies.

particular “equal shares” contract where each of N agents gets 1/N of the revenue (or output), and study Nash equilibria in e¤ort levels in this setting. Their focus is on the role of “peer pressure” i.e. social norms or informal monitoring and punishment within the group of agents in enhancing Nash equilibrium e¤ort levels.

However, as mentioned in the introduction, it is well-known that two-way distortion can arise in the single-agent case when the standard set-up is modi…ed so that the agent faces countervailing incentives (Lewis and Sappington[12], Maggi and Rodriguez-Clare[15]). This case arises when the reservation utility of the agent, as well as his cost of acting for the principal, depends on his private information. It was …rst observed by Lewis and Sappington[12] that two-way distortion could arise in this case. For example, consider the problem of regulation of a monopolist with unknown cost (Baron and Myerson,[1]) where the regulator chooses the output of the …rm ,and a transfer payment to the …rm i.e. a non-linear price. Such a model can be interpreted as a special case (i.e. without spillovers) of the one considered in this paper. Suppose, plausibly, that the …rm’s reservation pro…t (the pro…t it could make by exiting the regulated market and producing elsewhere) depends negatively upon its marginal cost parameter[14], μ. In this case, the …rm has an incentive to understate its marginal cost (to increase its reported reservation pro…t, in order to induce the regulator to set a higher price), as well as to overstate its marginal cost (again, to induce the regulator to set a higher price).

A complete analysis of a principal-agent problem with countervailing incentives is presented in Maggi and Rodriguez-Clare[15], where it is shown that the pattern of the two-way distortion (and whether or not there is pooling) depends crucially on whether the reservation utility of the agent is convex or concave in his private information. If it is a concave or mildly convex function, then the agent’s output is ine¢ciently low when his cost of production is low, and ine¢ciently high when his cost of production is high. If the reservation utility is strongly convex, then the opposite is the case[15].

[14] As in the model of this paper without spillovers, we assume that the cost of the …rm is $c(q;\mu)$ where q is output and the analogue of A1 above is satis…ed.

[15] Also, in the concave/mildly convex case, production is e¢cient at the highest and lowest costs, and also at an interior cost. In the strongly convex case, production is e¢cient only at an interior cost value.

One intuition for their result, in the context of the Baron-Myerson model, is as follows. Suppose that μ can only take on two values, “high” or “low”. If the …rm has, on balance, an incentive to overstate its cost, the optimal action for the principal is to allow the low-cost …rm to produce at a point where marginal revenue is greater than marginal cost. This is because high-cost …rms will wish to imitate low cost …rms at the full-information optimum, and these constraints can be slackened by reducing output and price of low-cost …rms, thus making their price-output pair less attractive to high-cost …rms. An identical logic applies in the other case: if the …rm has, on balance, an incentive to understate his cost, the optimal action for the principal is to allow the high-cost …rm to produce at a point where marginal revenue is less than marginal cost.

Now return to the case where μ is a continuous variable. The derivative of reservation pro…t with respect to μ measures the strength of the incentive that the …rm has to understate its marginal cost slightly in order to increase its reported reservation pro…t. If the derivative of reservation pro…t with respect to μ is decreasing in μ (i.e. the reservation pro…t is concave in μ), then the marginal incentive to understate μ in order to increase its reservation pro…t is stronger when μ is high, and so high (low) μ types have on balance an incentive to understate (overstate). But then by the argument in the previous paragraph, the …rm’s output is ine¢ciently low when its cost of production is low, and ine¢ciently high when its cost of production is high. The argument is similar in the case where the reservation pro…t is convex in μ:

The above discussion makes it clear that the intuition for two-way distortion in the countervailing incentives case is somewhat involved. By contrast, in our setting, there is a clear and simple intuition for the two-way distortion, as explained in Section 4 above. Moreover, the pattern of two-way distortion identi…ed here is robust: it does not depend on the precise nature of the production externality, as long as it is positive and satis…es the very weak assumptions in A2 above. Finally, Maggi and Rodriguez-Clare[15]assume a much more special class of cost functions than those considered in this paper[16].

[16] In the notation of this paper, they assume $c(q;\mu) = \mu q$:

References

D.Baron and R.Myerson,“Regulating a monopolist with unknown costs”, Econometrica, 50 (1982), 408-415

Y.Che and S.Yoo, “Optimal incentives for teams”, SSRI Working Paper 9808, University of Wisconsin

J.Cremer and R.P.McLean,“Optimal selling strategies under uncertainty for a discriminating monopolist when demands are interdependent”, Econometrica 53, (1985), 345-361

J.S.Demski and D.Sappington,“Optimal incentive schemes with multiple agents”, Journal of Economic Theory 33, (1984), 152-171

R.Cooper and A.John, “Coordinating coordination failures in Keynesian models”; Quarterly Journal of Economics 103, (1988), 441-463

Glover, “A simpler mechanism that stops agents from cheating”, Journal of Economic Theory 62, (1994), 221-229

O.D.Hart, Incomplete contracts and the theory of the …rm, (1995), MIT Press

W.Hildenbrand, Core and equilibria of a large economy, (1974), Princeton University Press

H.Itoh, “Incentives to help in multi-agent situations”, Econometrica, 59, (1991), 611-637

E.Kandal and E.P.Lazear, “Peer pressure and partnerships”, Journal of Political Economy 98, (1990), 225-264

J.J.La¤ont and J.Tirole, A theory of incentives and procurement in regulation, (1993), MIT Press

T.Lewis and D.Sappington,“Countervailing incentives in agency problems”, Journal of Economic Theory 49, (1989), 294-313

C.Ma, J.Moore, and S.Turnbull, “Stopping agents from ‘cheating”’, Journal of Economic Theory 46, (1988), 355-372

Mas-Colell,A, ”On a theorem of Schmeidler”, Journal of Mathematical Economics (1984) 13, 207-223

G.Maggi and A.Rodriguez-Clare, “On countervailing incentives”, Journal of Economic Theory 66, (1995), 238-263

D.Mookherjee,“Optimal incentive schemes with many agents”, Review of Economic Studies

51 (1984), 433-446

B.Nalebu¤ and J.Stiglitz, "Prizes and incentives: towards a general theory of compensation and competition", Bell Journal of Economics, 1983, 21-43

D.Sappington, "Optimal regulation of a multi-product monopoly with unknown technological capabilities", Bell Journal (1983), 453-462

Appendix

A. Appendix

Proof of Proposition 1. (i) First, as every agent is of measure zero, each agent takes Ψ as fixed when deciding on her announcement $\hat{\theta}$. So, from Definition 2, for truth-telling to be a dominant strategy, $\hat{\theta} = \theta$ must maximize (6), holding Ψ fixed. Define

$$w(\theta;\hat{\Psi}) \equiv u(\theta;\theta;\Psi) \tag{A.1}$$

to be the utility from truth-telling, conditional on Ψ. Standard arguments (see e.g. Laffont and Tirole[11]) imply that the necessary and sufficient conditions for this are as follows;

$$\frac{\partial v(\theta;\Psi)}{\partial \theta} = -c_{\theta}(q(\hat{\theta};\hat{\Psi})s(\hat{q});\theta) \text{ almost everywhere on } \Theta \tag{A.2}$$

$$\frac{\partial q(\theta;\Psi)}{\partial \theta} \leq 0 \text{ almost everywhere on } \Theta \tag{A.3}$$

where (A.2), (A.3) are the envelope and monotonicity conditions respectively. Integrating (A.2), we can write

$$w(\theta;\Psi) = w(\bar{\theta};\Psi) + \int_{\theta}^{\bar{\theta}} c_{\theta}(q(\hat{\theta};\hat{\Psi})s(\hat{q});z)dz \tag{A.4}$$

Also by definition from (A.1), $w(\theta;\Psi) \equiv t(\theta;\Psi) - c(q(\hat{\theta};\hat{\Psi})s(\hat{q});\theta)$, implying

$$t(\hat{\theta};\Psi) \equiv c(q(\hat{\theta};\hat{\Psi})s(\hat{q});\theta) + w(\hat{\theta};\Psi) \tag{A.5}$$

Combining (A.4) and (A.5) gives (9), with the constant A_D equal to $v(\bar{\theta};\Psi)$.

(ii) From Definition 3, for truth-telling to be a Nash strategy, all we need is that $\hat{\theta} = \theta$ must maximize (6), holding Ψ fixed at γ: But then a similar argument implies that the contract will be Nash incentive-compatible as long as

$$t_N(\hat{\theta};\gamma) = c(\hat{\theta}q_N(\hat{\theta};\gamma)s(q);\hat{\theta}) + \int_{\underline{\theta}}^{\bar{\theta}} c(q_N(\hat{\theta};\gamma)s(q);\hat{\theta})dz + A_N;\ A_N \in \Re \tag{A.6}$$

$$\frac{\partial q_N(\hat{\theta};\gamma)}{\partial \hat{\theta}} \leq 0 \text{ almost everywhere on } \Theta \tag{A.7}$$

where $\hat{q} = \int_{\theta \in \Theta} q(\theta;\gamma)d\gamma$ is aggregate output given truth-telling. We know by definition that given a Nash-incentive compatible contract $(q_N;t_N)$, there-is a truth-telling equilibrium in the announcement game. At this equilibrium, $\hat\gamma = \gamma$. So, if we choose $A_D = A_N$, $q_D(\theta) = q_N(\theta)$ all $\theta \in \Theta$, then $t_D(\theta;\gamma) = t_N(\theta;\gamma)$, all $\theta \in \Theta$, so there is a dominant-strategy incentive-compatible contract that yields the principal the same payoff as a Nash incentive-compatible contract when, in the induced announcement game, agents tell the truth.

(iii) Now consider the Nash incentive-compatible contract $(\hat q_N(\hat\theta;\gamma), \hat t_N(\hat\theta;\gamma))$ that maximizes the principal's expected payoff, under the assumption that agents tell the truth[17]. Given this contract, there may be other Nash equilibria in the announcement game, where a positive measure of agents do not tell the truth. It is clear that in any other such equilibrium, the principal can be no better off than in the truth-telling equilibrium. Thus, the maximum payoff that the principal can get from any Nash incentive-compatible contract is no higher than the payoff that the principal can achieve from a dominant-strategy incentive-compatible contract. □

Proof of Proposition 2. (i) We proceed to solve problem P by initially ignoring the monotonicity constraint $q'(\theta) \le 0$. In general, the effect on profit of a small increase in with respect to $q(\theta)$, taking into account the dependence of q on $q(\theta)$, is;

$$\frac{\partial \pi}{\partial q(\theta)} = r'(q(\theta))f(\theta) - \psi_e(q(\theta)s(q);\theta)s(q)f(\theta) - E\left[q(\theta)\psi_e(q(\theta)s(q);\theta)\right]s'(q)f(\theta) \quad \text{(A.8)}$$

First, suppose that there is a non-interior solution where $q(\theta) = 0$, for some $\theta \in \Theta$. Evaluating $\frac{\partial \pi}{\partial q(\theta)}$ at this point, we get

$$\begin{aligned}\frac{1}{f(\theta)}\frac{\partial \pi}{\partial q(\theta)} &= r'(0) - \psi_e(0;\theta)s(q) - E\left[q(\theta)\psi_e(q(\theta)s(q);\theta)\right]s'(q) \\ &> r'(0) - \psi_e(0;\theta)s(0) = r'(0) - \psi_e(0;\theta) > 0\end{aligned}$$

where the last inequality follows by A3. So it pays the principal to increase $q(\theta)$; a contradiction.

Again, suppose that there is a no solution because

$$\frac{\partial \pi}{\partial q(\theta)} > 0; \text{ all } q(\tilde\theta) \quad \text{(A.9)}$$

[17] It is clear that such a contract exists, from the proof of Proposition 2 below.

for some $\tilde{\theta} \in \Theta$. Fix some $\bar{q}(.)$; and take a sequence $\{q_n(\theta)\}_{n=1}^{\infty}$ with $q_n(\tilde{\theta}) = \bar{q}(\tilde{\theta}) + \phi_n$ with $\lim_{n\to\infty} \phi_n = \infty$, and $q_n(\theta') = \bar{q}(\theta')$; $\theta' \neq \tilde{\theta}$. Note that as $q_n(.)$ is equal to $\bar{q}(.)$ a.e., then $\bar{q} = E[q_n(\theta)]$ is fixed, as is $E\left[q_n(\theta)\psi_e(q_n(\theta)s(\bar{q});\theta)\right]$. Taking the limit in (A.8), we get

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{f(\theta)}\frac{\partial \pi}{\partial q(\theta)} &= \lim_{n\to\infty} r'(q_n(\theta)) - \lim_{n\to\infty} \psi_e(q_n(\theta)s(\bar{q});\theta)s(\bar{q}) \\
&\quad + E\left[q_n(\theta)\psi_e(q_n(\theta)s(\bar{q});\theta)\right] s'(\bar{q}) \qquad \text{(A.10)}\\
&= \lim_{q\to\infty} r'(q) - \lim_{e\to\infty} \psi_e(e;\theta)s(\bar{q}) - E[\bar{q}(\theta)\psi_e(\bar{q}(\theta)s(\bar{q});\theta)]s'(\bar{q}) \\
&< 0
\end{aligned}$$

where in the last line we have used $\lim_{q\to\infty} r'(q) = 0; \lim_{e\to\infty} \psi_e(e;\theta) = \infty$ where the second limit follows directly from $\lim_{e\to\infty} c_e(e;\theta) = \infty$ in A3 and $\psi_e \geq c_e$. So, we conclude an interior solution always exists. This interior solution is characterized by first-order condition for a maximum of (11), which we obtain from (A.8) by equating the RHS to zero, dividing through by $f(\theta)$; and then writing out ψ_e in full. This gives (13) in the text.

(ii) Let the solution to (13) be $q(\theta)$. For $q(\theta)$ to be feasible in P, it must be the case that it satisfies the monotonicity condition $\bar{q}'(\theta) \leq 0$. Note that from (13), using the second-order condition, we have;

$$\text{sign } q'(\theta) = \text{sign } \frac{\partial^2 \pi}{\partial q(\theta)\partial\theta} \qquad \text{(A.11)}$$

But again from (13), we have

$$\frac{\partial^2 \pi}{\partial q(\theta)\partial\theta} = -\psi_{e\theta}[(q(\theta)s(q);\theta)s(q)f(\theta) - q(\theta)(q(\theta)s(q);\theta)s'(q)f(\theta)] \qquad \text{(A.12)}$$

so it suffices to show that $\psi_{e\theta} > 0$. Now, from (12), we have

$$\psi_{e\theta}(q(\theta)s(q);\theta) = c_{e\theta}(q(\theta)s(q);\theta) + \frac{1}{h(\theta)}c_{e\theta\theta}(q(\theta)s(q);\theta) - \frac{h'}{h^2}c_{e\theta}(q(\theta)s(q);\theta) \qquad \text{(A.13)}$$

Also, from A1, $c_{e\theta\theta} \geq 0$, and from monotone hazard rate condition, $h' \leq 0$. So, from (A.13), $\psi_{e\theta} > 0$, as required. □

Proof of Theorem 1. (i) We show that if the distribution of θ, F, has an everywhere decreasing hazard rate, then the induced distribution $\kappa = \kappa(\theta)$; has an everywhere decreasing hazard rate. This fact means that we can, without loss of generality, set $\kappa(\theta) = \theta$. In particular, we will use the fact $c_{\theta\theta e} = 0$ in what follows. First, from A1, $c_\theta = \kappa'(\theta)c(e) > 0$, so ϕ is invertible on $K = [\phi(\underline{\mu}); \phi(\bar{\mu})]$. So the induced distribution of $\phi; G(x) := F(\phi^{-1}(x))$; $x \in K$ is well-defined, as is the density $g(x) = f(\phi^{-1}(x))/\phi'(x)$. Recalling the definition of the hazard rate of G, it is sufficient to prove that $G(x)/g(x)$ is increasing in x: Then

$$\frac{d[G(x)/g(x)]}{dx} = \frac{G(x)}{f(\phi^{-1}(x))}\phi''(x) + \phi'(x)\frac{d[F(\phi^{-1}(x))/f(\phi^{-1}(x))]}{dx}$$

The first term is non-negative as $\phi'' \geq 0$, and the second term is strictly positive by the fact that F has a decreasing hazard rate and $d\phi^{-1}/dx > 0$. So, we conclude that $g(x)/G(x)$ is strictly decreasing on K, as required.

(ii) Note that $1/h(\underline{\mu}) = F(\underline{\mu})/f(\underline{\mu}) = 0$, and so from (13), we have at $\mu = \underline{\mu}$ that

$$r'(q(\underline{\mu})) = c_e(q(\underline{\mu})s(q); \underline{\mu})s(q) + E\left[q(\mu)c_e(q(\mu)s(q); \mu)\right]s'(q) + E\left[\frac{q(\mu)}{h(\mu)}c_{\mu e}(q(\mu)s(q); \mu)\right]s'(q) \tag{A.14}$$

As $s'(q) < 0$, and $c_{\mu e} > 0$, it follows from (A.14) that

$$r'(q(\underline{\mu})) < c_e(q(\underline{\mu})s(q); \underline{\mu})s(q) + E\left[q(\mu)c_e(q(\mu)s(q); \mu)\right]s'(q)$$

i.e. oversupply at $\mu = \underline{\mu}$. Also, from (13), and Definition 4, to have undersupply at $\mu = \bar{\mu}$, we must have

$$\frac{c_{\mu e}(q(\bar{\mu})s(q); \bar{\mu})}{h(\bar{\mu})}s(q) + E\left[\frac{c_{\mu e}(q(\mu)s(q); \mu)}{h(\mu)}q(\mu)\right]s'(q) > 0$$

Rearranging, and recalling $s'(q) < 0$, we get

$$-\frac{s(q)}{s'(q)}\frac{c_{\mu e}(q(\bar{\mu})s(q); \bar{\mu})}{h(\bar{\mu})} > E\left[\frac{q(\mu)}{h(\mu)}c_{\mu e}(q(\mu)s(q); \mu)\right] \tag{A.15}$$

Now note the following facts: (i) as the monotone hazard rate condition holds, $\frac{1}{h(\mu)}$ is increasing in μ; (ii) from Proposition 1, $q(\mu)$ is decreasing in μ; (iii) $c_{\mu e}(q(\mu)s(q); \mu)$ is increasing in μ as

$$\begin{aligned}\frac{dc_{\mu e}(q(\mu)s(q); \mu)}{d\mu} &= c_{\mu ee}(q(\mu)s(q); \mu)s(q)q'(\mu) + c_{\mu\mu e}(q(\mu)s(q); \mu) \\ &= c_{\mu ee}(q(\mu)s(q); \mu)s(q)q'(\mu) < 0\end{aligned}$$

where we have used $c_{\mu\mu e} = 0$ in the last line. So, from (i)-(iii), it follows that

$$E\left[\frac{c_{\mu e}(q(\mu)s(q); \mu)}{h(\mu)}\right]E[q(\mu)] > E\left[\frac{c_{\mu e}(q(\mu)s(q); \mu)q(\mu)}{h(\mu)}\right] \tag{A.16}$$

So, from (A.16), a sufficient condition for (A.15) to hold is that

$$-\frac{s(q)}{s'(q)}\frac{c_{\mu e}(q(\bar{\mu})s(q);\bar{\mu})}{h(\bar{\mu})} > E\left[\frac{c_{\mu e}(q(\mu)s(q);\mu)}{h(\mu)}\right] E\left[q(\mu)\right] \tag{A.17}$$

Rearranging (A.17), using $q = E[q(\mu)]$, gives

$$\frac{c_{\mu e}(q(\bar{\mu})s(q);\bar{\mu})/h(\bar{\mu})}{E\left[c_{\mu e}(q(\mu)s(q);\mu)/h(\mu)\right]} > -\frac{s'(q)q}{s(q)} \tag{A.18}$$

But, the LHS of (A.18) is greater than 1 by the properties of $c_{\mu e}/h$ derived above. So, it is certainly sufficient for undersupply at $\mu = \bar{\mu}$ that

$$1 \geq -\frac{s'(q)q}{s(q)} \tag{A.19}$$

As $s' = -\frac{1}{g^2}g'\circ'$, it follows that

$$\frac{-s'q}{s} = \frac{1}{g}g'\circ' q \tag{A.20}$$

But differentiation of (3), which implicitly defines $\circ$; gives

$$\circ' = \frac{1}{g + eg'(e)} \tag{A.21}$$

So, combining(A.20) and (A.21), and using $q = eg$; we get

$$\frac{-s'q}{s} = \frac{eg'(e)}{g + eg'(e)} < 1$$

so (A.19) clearly holds, as required.

(iii) The final step is to show that there is a single critical value μ' below which there is oversupply, and above which there is undersupply, it suffices to show that the difference between the perceived marginal cost (the right-hand side of (13)) and the true marginal cost (the right-hand side of (14)) is monotonically increasing in μ for a fixed $q; q(\mu)$: This difference is given in (17), and is clearly increasing in μ for fixed $q; q(\mu)$ from the monotone hazard rate condition and the properties of c. □

Proof of Theorem 2. It is easy to check, using (23) and following the proof of Theorem 1, that there will be two-way distortion if

$$\frac{-(s^a_a(q;q) + s^a_b(q;q))q}{s^a(q;q)} \leq 1 \tag{A.22}$$

First, it follows directly from differentiation of (22) that

$$\begin{aligned} s^a_a(q^a, q^b) &= -\frac{g'}{g^2}[\gamma^a_a + \sigma\gamma^b_a] \\ s^a_b(q^a, q^b) &= -\frac{g'}{g^2}[\gamma^a_b + \sigma\gamma^b_b] \end{aligned} \tag{A.23}$$

Now, using (A.23),(20),(21), and the fact that $q_a = q_b = q$ we get

$$\begin{aligned} \frac{-(s^a_a(q,q) + s^a_b(q,q))q}{s^a(q,q)} &= \frac{g'q}{g}[\gamma^a_a + \sigma\gamma^b_a + \gamma^a_b + \sigma\gamma^b_b] \\ &= \left(\frac{g'e}{g}\right)\frac{g}{D}(1+\sigma)[g + eg'(1-\sigma)] \\ &= \left(\frac{g'e}{g}\right)\left(\frac{g(1+\sigma)[g + eg'(1-\sigma)]}{g^2 + 2geg' + e^2(g')^2(1-\sigma^2)}\right) \end{aligned} \tag{A.24}$$

where $g = g(e + \sigma e)$. So, we require that the term on the RHS of (A.24) be weakly less than 1, which simplifies to

$$\left(\frac{g'e}{g}\right)\sigma \le 1 + \left(\frac{g'e}{g}\right)$$

which certainly holds. $\square$

B. Junkyard

This assumption could be relaxed in the following way without changing the statement and proof of results. Let $d(i;j)$ be a measure of the "distance" between agents i and j in the production process, with $d(i;j) = d(j;i)$, all $i;j \in I$; $d(i;i) = 0$. and normalise so that $\int_{j \in I} d(i;j)d\mu = 1$. Define the aggregate spillover for i as $\sigma_i = \int_{j \in I} d(i;j)q_j d\mu$. then (2) could be generalised to $q_i = e_i g(\sigma_i)$. Also, suppose that g is linear, so $q_j = e_j[\alpha + \beta\sigma_j]$. Then, multiplying through by $d(i;j)$ and integrating this over $j \in I$ we get

$$Q_i = \int_{j \in I} d(i;j)q_j d\mu = \sigma_i[\alpha + \beta \int_{j \in I} d(i;j)\sigma_j d\mu] = \sigma_i[\alpha + \beta\Sigma_i]$$

. Preforming the same operation again, we get

$$Q = \Sigma[\alpha + \beta\Sigma];\ Q = \int_{i \in I} d(i;j) \int_{j \in I} d(i;j)q_j d\mu;\ \Sigma = \int_{i \in I} d(i;j) \int_{j \in I} d(i;j)e_j d\mu$$

We can then invert to get $\Sigma = \gamma(Q)$

Example 1.

Let $g(e) = 1 + \sigma e$. Note that $g(0) = 1$, as required. So,

$$q = eg(e) = e + \sigma e^2$$

So,

$$e = \gamma(q) = \frac{\sqrt{1 + 4\sigma q} - 1}{2\sigma}$$

So,

$$s(q) = \frac{1}{g(\gamma(q))} = \frac{1}{1 + \sigma\frac{\sqrt{1+4\sigma q} - 1}{2\sigma}} = \frac{2}{\sqrt{1 + 4\sigma q} + 1}$$

So,

$$\frac{-qs'}{s} = \frac{(1 + 4\sigma q)^{-0.5}4\sigma q}{\sqrt{1 + 4\sigma q} + 1}$$

So, $\frac{-qs'}{s} \leq 1 \iff$

$$(1 + 4\sigma q)^{-0.5}4\sigma q \leq \sqrt{1 + 4\sigma q} + 1$$

which certainly holds, as can be checked after some rearrangement. $\square$

Here $c = \theta e$, and θ is uniformly distributed on an interval of unit length, so $F(\theta) = \theta - \underline{\theta}$; $f(\theta) = 1$. Also, $g(e) = e^{\alpha}$, then it is easily checked that $s(q) = q^{-\frac{\alpha}{1+\alpha}} = q^{-\beta}$, so

$0 < \beta < 1$. Finally, $r(q) = 2\sqrt{q}$. Then it is easily checked that $\psi(e;\theta) = (2\theta - \underline{\theta})e$; so the perceived cost function is

$$\psi = (2\theta - \underline{\theta})q(\theta)q^{-\beta}$$

Then, the condition (??) reduces to

$$1 + \beta \geq \frac{2q(2\theta - \underline{\theta})}{E[(2\theta - \underline{\theta})q(\theta)]} \tag{B.1}$$

But as $E[(2\theta - \underline{\theta})q(\theta)] \geq E[q(\theta)]\underline{\theta} = \underline{\theta}q$, then (B.1) certainly holds if

$$1 + \beta \geq \frac{2(2\bar{\theta} - \underline{\theta})}{\underline{\theta}} \Rightarrow \beta 2 + \frac{4}{\underline{\theta}}$$

Also, the first-order condition reduces to

$$(q(\theta))^{0.5} = \psi_e(q(\theta)s(q);\theta)s(q) + E\left[q(\theta)\psi_e(q(\theta)s(q);\theta)\right]s'(q)$$

Returning to the Theorem, we note the following corollary:

Corollary 3. If the external effect of aggregate effort is iso-elastic i.e. $g(e) = e^{\alpha}$, $\alpha > 0$, then there is always two-way distortion.

Proof. If $g(e) = e^{\alpha}$, then it is easily checked that $s(q) = q^{-\frac{\alpha}{1+\alpha}}$, so $-\frac{s'q}{s} = \frac{\alpha}{1+\alpha} < 1$. But as remarked above, the upper bound on $-\frac{s'q}{s}$ in the Theorem is greater than unity, so in this case, the condition on the elasticity in the theorem always holds. ■

e wil also assume directly the following properties of s :

A2. $s(0) < \infty$, $\lim_{q \to \infty} qs'(q) = 0$:

The first of these conditions says that any agent is able to produce even if all other agents do not, and the second ensures that in the aggregate, the size of the externality goes to zero as output goes to infinity. An example which satisfies these conditions is the iso-elasric case $g(e) = e^{\alpha}$: Then it is easily checked that $s(q) = q^{-\frac{\alpha}{1+\alpha}} = q^{-\beta}$, $0 < \beta < 1$ so $s(0) = \infty$, and $qs'(q) = -\beta q^{-\beta}$. [CUT Following Demski and Sappington[4], we focus on two possible equilibrium concepts for this game, dominant strategy and Nash equilibrium.

This gives rise two concepts of an incentive-compatible contract.]La¤ont(1995) has shown that in an otherwise standard principal-agent model where the agent may take some unobservable e¤ort to reduce the probability of environmental catastrophe, and where the agent is risk-averse, and has limited liability, then

Consequently, de…ne a transfer schedule as a map $t : <^2_+ \; ! \; <$ where is the transfer to agent i if he produces output i and aggregate output is q. Given a transfer schedule, the agents then play a game where the strategies are $(q_i)_{i2I}$, and payo¤s $u_i = t(q_i, q) \;¡\; q_i s(q)µ_i$. Call this the output game.

This game is non-trivial due to the spillovers, and the fact that $t(q_i, q)$ may depend non-trivially on q: This is simply the inventive-compatible payment schedule (9) written as a function of $(q_i; q)$ rather than $(µ_i; º)$. As $q^0(µ_i) < 0$, q_i maximizes (24) if and only if $µ_i$ maximizes (9): So, we can conclude that faced with payo¤ $u_i = t(q_i; q) \;¡\; q_i s(q)µ_i$ where $t(q_i; q)$ is de…ned in (24) $q_i = q(µ_i)$ maximizes u_i whatever q i.e. $q_i = q(µ_i)$ is a dominant strategy for i in the output game[18]. So, the output game replicates the outcome of the direct mechanism described above.

[18] This is quite a striking result, because (as Cooper and John[5] have shown), output games of this type without principal-agent relationships (i.e. where agents capure the full value of their output) typically have multiple equilbria, due to strategic complementarities between agents.