Risk & Sustainable Management Group

Risk and Uncertainty Working Paper: R07#6

Research supported by an Australian Research Council Federation Fellowship
http://www.arc.gov.au/grant_programs/discovery_federation.htm

Sharp and Diffuse Incentives in Contracting

By

Flavio M. Menezes

and

John Quiggin

Australian Research Council Federation Fellow, University of Queensland

Schools of Economics and Political Science
University of Queensland
Brisbane, 4072
rsmg@uq.edu.au
http://www.uq.edu.au/economics/rsmg

Sharp and Diffuse Incentives in Contracting

Flavio M. Menezes	John Quiggin
School of Economics	School of Economics
University of Queensland	University of Queensland
Email: f.menezes@uq.edu.au	Email: j.quiggin@uq.edu.au

2007

Abstract

This paper investigates the optimality of sharp incentives in contracts where output prices are set at the time of contracting but are random in nature. It shows that when prices are specified with error, schemes involving sharp incentives might result in substantial deviations from first-best output levels. The randomness of prices creates arbitrage opportunities that are exploited by agents producing phenomena such as 'cost-shifting'. Both linear and piece-wise linear contracts are shown to be subject to the possibility of arbitrage. The paper then demonstrates that incentive schemes that are arbitrage-proof exhibit 'diffuse' incentives.

JEL Classification: D81, D86.

Key words: incentives; contracts; arbitrage

1 Introduction

The optimal contract in a standard principal–agent model with a risk-averse principal and a risk-neutral agent commonly involves a linear (or, more precisely affine) contract, designed to provide the agent with sharp incentives to exert optimal effort. Typically, such contracts take the form of a payment $y(w) = A+\gamma w$, where A is a fixed fee (that could be negative), γ a positive constant, and w is an observed and measured outcome. If w is output and γ is the marginal value of output, assumed constant, then under contracts of this kind, the agent keeps all the residual gains created by her effort after paying a fixed fee.[1]

If the agent's effort is observable – or if it can be inferred from outcomes – then the Pareto optimal level of output is achieved by paying the agent the full value of his or her output. An example of this type of contract is a concession contract where the operator pays the government a fixed fee and keeps all profits resulting from the concession. This type of contract provides the agent with sharp incentives to provide the first-best level of effort.

In practice, however, many different types of contracts, involving distinct levels of incentives, coexist. For example, Walls (2003, p.32) reports on the existence of contracts in the solid waste and recycling industries where incentives vary from being very sharp (where firms retain all of the revenues from sale of materials) to nonexistent (where firms retain none of the revenues from sale of materials).

A variety of explanations has been offered as to why contracts might not have incentives as sharp as those exhibited by contracts where the agent retains the residual gains created by her efforts. A central theme of the principal–agent literature is that risk aversion on the part of agents, in combination with asymmetric information, may generate a trade-off between risk sharing and sharp incentives.[2] Alternatively, it may not be possible to measure outcomes pre-

[1] Note, however, that w may depend on random factors in addition to the agent's level of effort.

[2] For example, although under certain circumstnaces it can be shown that the optimal contract would set $\gamma = 1$ for the case of a risk-neutral agent, risk aversion can lead the optimal to be close to zero. For an excellent discussion of the trade-off between incentives and

cisely, or there might be more than one attribute associated with outcomes. These characteristics might also result in the adoption of lower-powered incentive contracts. Moreover, under some circumstances it might be approximately optimal to have a step function incentive scheme where the agent is rewarded (or penalized) by a fixed amount if performance exceeds (or falls under) a given threshold.(Holmström 1979, Mirrlees 1999)

In this paper we offer a new explanation of the prevalence of lower-powered contracts. We argue that, if prices are specified with error, schemes involving sharp incentives might result in substantial deviations from first-best output levels. This is distinct, for example, from the incentive problem identified by the multi-tasking literature (Holmström and Milgrom 1991, Baker 1992), where compensation on any subsets of tasks will result in a reallocation of activities toward those that are directly compensated and away from the uncompensated activities. The explanation that we offer for the inefficiency of sharp contracts is not related to the exclusion of some tasks from the incentive contract but rather to strategic behavior by agents faced with variability in prices.

In particular, agents may be presented with arbitrage opportunities arising from a distinction between the prices they face and the prices fixed in the incentive contract. An example is provided by the phenomenon of 'cream-skimming', which arises in a variety of public policy contexts where agents, such as hospitals or employment agencies, are rewarded on the basis of measured outputs. The typical case of cream-skimming arises where a fixed price is offered for a heterogeneous output.

For example, an employment agency may be rewarded by a payment proportional to the number of clients placed in jobs. Although this implies a fixed price for each client placed, the output is heterogeneous, because both clients and jobs are heterogeneous. Arbitrage opportunities arise if the agency can generate a supply of short-term jobs, at a unit cost lower than the incentive payment. One such case was reported in Australia following a move to competitive tendering of employment services (Webster and Harding 2000, p. 11). This and similar

risk, see Dixit (2002).

problems led to a move towards lower-powered incentive contracts.

As a second example, consider an incentive mechanism set up by a university that can accurately measure and value research performance, but is subject to error in valuing teaching. If researchers can use the returns from the incentive mechanism to pay others to perform part or all of their teaching obligations, the scheme may have unexpected and perverse effects.

The paper is organized as follows. Section 2 includes a brief survey of existing empirical evidence, and of the literature on incentive schemes. In Section 3, we present a simple model of principal–agent contracts in the presence of price uncertainty. We derive our main result, giving conditions under which any linear pricing rule will give rise to arbitrage opportunities with positive probability. We then extend our analysis to show that similar problems arise with sharp incentives as implemented by piecewise linear price contracts. In Section 4, we consider lower-powered contracts and give conditions under which such contracts are arbitrate-proof. in Section 5, we consider robust contracts based on diffuse incentives. In Section 6, we discuss the implications of our results.

2 Background

There are well-known reasons why contracts might not have incentives as sharp as those exhibited by contracts where the agent retains most of the residual gains created by her efforts. For example, if the agent is risk averse, the need to provide the agent with insurance leads to less sharp incentives. This is the well-known trade-off between risk and incentives.[3] Alternatively, it may not be possible to measure outcomes precisely or there might be more than one attribute associated with outcomes. These might also result into lower-powered incentive contracts.

The existence of sharp incentives implemented via linear contracts might also be ineffective in the context of various actions or tasks (generally referred in the literature as the multi-tasking case). If actions are substitutes, sharp

[3] See, for example, Milgrom and Roberts (1992), Chapter 7.

incentives imply that exerting more effort on one task increases the marginal cost of the substitute task.[4]

Moreover, linear contracts that reward particular outcomes in the context of jobs that involve various tasks might also lead to outcomes that are neither wanted or anticipated. For example, basing teacher evaluations only on the performance of students on standardized test will lead teachers to focus their effort on one measure of overall teaching performance, the results of the tests. This might lead to inefficient outcomes to the extent that less effort will be put into teaching general skills such as writing and problem solving.[5]

In the same vein, McKim and Hughart (2005) report on the results of a survey of staff incentive schemes used by microfinance institutions. Most such schemes consist of a weighted formula based upon three or more individual performance indicators. Incentive payments are made each month and amount to a quarter of a credit officer's fixed salary. By and large, the survey finds that staff incentives schemes are perceived as generating a positive effect on the financial performance of the institutions. However, a reported side effect was the shift in focus towards wealthier clients and clients in more urban areas. This is a cause of concern as it defeats an important rationale for the provision of microfinancing. Another unwanted consequence was an erosion of the quality of the loan portfolio as credit officers responded to the incentives facing them by offering larger loans, or by being less thorough in their evaluations of the loans.

There are other types of incentive contracts that are considered in the theoretical literature such as piece-wise linear contracts. Where there exist particular critical thresholds at which the principal is very risk averse, schemes that associate very steep incentives with movements across these particular thresholds may be useful. For example, Holmström (1979) and Mirrlees (1999) analyze contracts where the compensation of the agent is determined by a step function, conditional on achieving a certain threshold performance. In this case, if the

[4] See, for example, Holmström and Milgrom, 1991. For an excellent survey of issues in incentive schemes design, see Burgess and Ratto (2003).

[5] For different examples see, among others, Kerr (1975) and Baker, Gibbons and Murphy (1994).

agent reduces his effort even by a small amount, he faces a large penalty. Holmström and Mirrlees show that this type of contract is optimal when the output is very sensitive to the agent's effort in the neighborhood of the threshold. The intuition is that this type of scheme is optimal when agents are sufficiently concerned about the penalty they will suffer for small reductions in effort. As Dixit (2002) points out, step function schemes are also vulnerable to manipulation or gaming. In particular, there is a concern about timing issues.

Consider, for example, an environment where there are random fluctuations and, if she reports outcomes annually, the agent secures a bonus in good years but fails to do so in bad years. Now suppose the agent can choose when to report, or can shift returns from one reporting period to another. Then the rational response is to smooth reported performance, so that reported performance in most periods just exceeds the threshold. For example, it has been claimed that companies smooth their reported profits to ensure that they narrowly beat market expectations as often as possible.[6]

Below we show that both linear and nonlinear (or more specifically piece-wise linear) incentives can generate arbitrage opportunities for the agent, which when exercised can lead to undesirable outcomes for the principal. These arbitrage opportunities do not depend on enlarging the strategy space, as in the example in the previous paragraph. Instead, arbitrage arises from the randomness of the prices specified in the incentive contract.

3 Sharp Incentives and Arbitrage

An agent has available a production technology represented by an output correspondence

$$Z(\mathbf{x}) = \left\{ \mathbf{z} \in \Re^M : \mathbf{x} \in X \text{ can produce } \mathbf{z} \right\}$$

where $X \subseteq \Re^N_+$ is a compact set of inputs available to the agent and $\mathbf{z} \in \Re^M$ is a vector of net outputs. We make the following assumptions about the set $Z(\mathbf{x})$:

[6] See Courty and Marschke (2004) and, for a survey of empirical results on the effectiveness and use of incentive schemes, Prendergast (1999).

Assumption A1: (i) For each $\mathbf{x}$, $Z(\mathbf{x})$ is compact, strictly convex and contains $\mathbf{0}^M$

(ii) The technology is additive. That is, if $\mathbf{z} \in Z(\mathbf{x})$, $\mathbf{z}' \in Z(\mathbf{x}')$ and $\mathbf{x}+\mathbf{x}' \in X$, then $\mathbf{z}+\mathbf{z}' \in Z(\mathbf{x}+\mathbf{x}')$.

As in Chambers and Quiggin (2000), the outputs may be state-contingent, and realized in period 1, while inputs are committed in period 0. In this case, the dimension of the output space may be characterized as $M = K \times S$, where K is a set of commodities and S is a set of states of nature. We assume a full set of state-contingent markets, so payments to the agent are received in period 0. In effect, this is equivalent to assuming risk neutrality for the agent at the given state-claim prices.

Note that the technology permits negative outputs. The interpretation is that the contract may allow the principal to supply outputs of some goods or services to the agent, rather than *vice versa*. Where this possibility is precluded, it will be noted explicitly.

The agent's objective function is

$$y(\mathbf{z}) - g(\mathbf{x}) \tag{1}$$

where $y(\mathbf{z})$ is a payment from the principal and $g(\mathbf{x})$ represents disutility of effort, an increasing convex function. Let

$$\mathbf{x}(\mathbf{z}) = \arg\min\{g(\mathbf{x}) : \mathbf{z} \in Z(\mathbf{x})\} \tag{2}$$

be the input demand associated with output $\mathbf{z}$.

The agent chooses her level of effort to maximize her objective function (1) taking into account how $\mathbf{x}$ affects $\mathbf{z}$ and, as a result, the payment she receives from the principal.

In the particular case where the agent faces a price schedule $\mathbf{p} \in \Re^{\mathbf{M}}_{+}$, that is, a linear contract $y(\mathbf{z}) = \mathbf{pz}$, we define the agent's supply response correspondence as:

$$\mathbf{z}(\mathbf{p}) = \arg\max\{\mathbf{pz} - g(\mathbf{x}) : \mathbf{z} \in Z(\mathbf{x})\}. \tag{3}$$

The supply correspondence (3) identifies, for any vector $\mathbf{p}$, the agent's maximimal choice(s) of net output.

The principal seeks to choose the reward function $y(\mathbf{z})$ in order to maximize

$$W(\mathbf{z}) - y(\mathbf{z}) \tag{4}$$

subject to the incentive-compatibility requirement

$$\mathbf{z} \in \arg\max \{y(\mathbf{z}) - g(\mathbf{x})\}$$

and a participation constraint of the type $y(z) - g(x) \geq u^0$. We assume that $W(\mathbf{z})$ is continuous and concave so that a solution to the principal's maximization problem exists.

The first-best contract in this setting is obtained by setting

$$y(\mathbf{z}) = \mathbf{p}^*\mathbf{z} - \bar{y}$$

where

$$p_m^* = W_m(\mathbf{z}),$$

so that p_m^* is the price chosen by the principal for each net output m, which in equilibrium is equal to the marginal utility of net output m, and $\bar{y}$ is chosen to satisfy the incentive compatibility constraint. That is, the optimal contract involves the agent obtaining all the residual gains from her efforts after paying a fixed fee to the principal. We denote by $\mathbf{z}^*$ the first-best output that satisfies

$$y(\mathbf{z}^*) = \mathbf{p}^*\mathbf{z}^* - \bar{y}$$

4 The problem with stochastic variation in prices

We argue that, in practice, while the principal may wish to set some price $\hat{\mathbf{p}}$, the price actually faced by agents will be subject to some stochastic variation around $\hat{\mathbf{p}}$. The next assumption specifies a simple way to capture the stochastic variation in prices.

Assumption A2: In any actual implementation the price vector is of the form $\mathbf{p} = \hat{\mathbf{p}} + \varepsilon$ where $E[\varepsilon] = 0$ and ε has positive density on some ball around $\mathbf{0}$ with diameter δ.[7]

The principal's *feasible price-incentive problem* under Assumption A2 is to choose $\hat{\mathbf{p}}$ to maximize

$$E[W(\mathbf{z}(\mathbf{p})) - \mathbf{pz}] = E[W(\mathbf{z}(\hat{\mathbf{p}} + \varepsilon)) - (\hat{\mathbf{p}} + \varepsilon)(\mathbf{z}(\hat{\mathbf{p}}) + \varepsilon))]. \tag{5}$$

That is, for a given realization of ε, and therefore of $\mathbf{p} = \hat{\mathbf{p}} + \varepsilon$, the agent chooses $\mathbf{z}$, given the relationship between $\mathbf{z}$ and $\mathbf{x}$ and the resulting disutility generated from $\mathbf{x}$, to maximize her utility. The principal will choose $\mathbf{p}$, amongst the set of incentive-compatible prices, to maximize his expected utility.

The following definition plays a crucial role in our main argument.

Definition 1: For any $\mathbf{p}$, we define a *zero-profit net output rearrangement* for $\mathbf{p}$ as an net output vector $\mathbf{z}$ such that $\mathbf{z} \in Z\left(\mathbf{0}^N\right)$, $\mathbf{pz} = 0$. We denote by $\hat{Z}(\mathbf{p})$ the set of zero-profit net output rearrangements for $\mathbf{p}$.

The following result will be useful later.

Lemma 1 If $\mathbf{z} \in \mathbf{z}(\mathbf{p})$, and $\hat{\mathbf{z}} \in \hat{Z}(\mathbf{p})$, then $\mathbf{z} + \hat{\mathbf{z}} \in z(\mathbf{p})$.

Proof: Since $\mathbf{z} \in Z(\mathbf{x}(\mathbf{z}))$ and $\mathbf{z} \in Z\left(\mathbf{0}^N\right)$, additivity of Z implies $\mathbf{z} + \hat{\mathbf{z}} \in Z(\mathbf{x}(\mathbf{z}))$. Now, since $\mathbf{p}(\mathbf{z} + \hat{\mathbf{z}}) = \mathbf{pz}$, then $\mathbf{z} + \hat{\mathbf{z}} \in \mathbf{z}(\mathbf{p})$.

Lemma 1 establishes that for a given price schedule $\mathbf{p}$, if we add a zero-profit net output rearrangement to a vector $\mathbf{z}$ in the agent's supply response correspondence, the resulting net output vector is also in the agent's supply response correspondence.

Several features of Definition 1 are of particular interest. First, by the additivity assumption, $\hat{Z}(\mathbf{p})$ is a convex cone. Second, since $0^M \in Z\left(\mathbf{0}^N\right)$, $\hat{Z}(\mathbf{p})$ is non-empty. Of course, we are interested to investigate the circumstances under which $\hat{Z}(\mathbf{p})$ is trivial or non-trivial, that is, when $\hat{Z}(\mathbf{p})$ is or is not equal to

[7] Other details of the distribution are not crucial.

$\{\mathbf{0}^M\}$. Consider the case of a (possibly stochastic) production function technology

$$Z(\mathbf{x}) = \{\mathbf{z} : \mathbf{z} \leq f(\mathbf{x})\}$$

for some $f : X \to \Re^M$ such that $f(\mathbf{0}^N) = 0^M$. For technologies of this kind,

$$Z(\mathbf{x}) = \{\mathbf{z} : \mathbf{z} \leq \mathbf{0}^M\}$$

and $\hat{Z}(\mathbf{p}) = \{\mathbf{0}^M\}$ for any $p \in \Re^M_{++}$.

The stochastic production function or Leontief technology represents the maximally inflexible form of a state-contingent technology (Chambers and Quiggin 2000). Conversely, the more flexible the technology, the larger the set of zero-profit net output rearrangements available to the agent. In particular, consider a fully allocable technology

$$Z(\mathbf{x}) = \left\{\mathbf{z} : \sum_m g_m(z_m) \leq \mathbf{x}\right\}$$

where $g_m : \Re \to \Re^N$ is an input requirement technology.

Note that the specification of g_m allows for negative inputs and outputs. With this technology, agents can purchase some outputs z_m in the market and shift the associated inputs to other activities. For suitable price vectors p, $\hat{Z}(\mathbf{p})$ will be non-empty.

Third, since prices are non-negative, any non-trivial element of $\hat{Z}(\mathbf{p})$ must include both negative and positive elements. As noted above, we allow net output vectors to include negative elements, unless the contract structure which governs the relationship between the agent and the principal explicitly precludes this. More importantly, by virtue of Lemma 1, given a strictly positive initial output vector $\mathbf{z} \in \mathbf{z}(\mathbf{p})$ and a feasible zero-profit net output rearrangement $\hat{\mathbf{z}} \in \hat{Z}(\mathbf{p})$, $\mathbf{z} + \lambda\hat{\mathbf{z}} \in \mathbf{z}(\mathbf{p})$ for any $\lambda > 0$, and, for suitably small λ, $\mathbf{z} + \lambda\hat{\mathbf{z}}$ will be strictly positive.

Finally, we restrict attention to the case when the input vector $\mathbf{x}$ is unchanged. The definition here could be generalized to cover the case when $g(\mathbf{x})$ is unchanged, and the results derived below would, in most cases, carry through.

However, our primary concern is with the reward function $y(\mathbf{z})$ and for this purpose, an exclusive focus on changes in z is more appropriate.

We say that $(\mathbf{p}, \mathbf{x}, \mathbf{z})$ an interior solution if x is an interior point of X. Our main result is stated below.

Proposition 1 If $\hat{Z}(\mathbf{p}) \neq \{\mathbf{0}^M\}$, $(\mathbf{p}, \mathbf{x}, \mathbf{z})$ cannot be an interior solution for the principal's feasible price–incentive problem.

Proof: Let $\hat{\mathbf{z}} \neq \mathbf{0^M} \in \hat{Z}(\mathbf{p})$. There exist directions $\mathbf{d}$ such that $(\mathbf{p}+t\mathbf{d})\hat{\mathbf{z}} \neq 0$, $\forall k \neq 0$. In fact, this is true of all directions except those in a single hyperplane passing through $\mathbf{p}$ and orthogonal to $\hat{\mathbf{z}}$. In any feasible implementation of $\mathbf{p}$, therefore, there is positive probability that $(\mathbf{p}+\varepsilon)\hat{\mathbf{z}} \neq 0$. For any such ε, the agent's optimization problem allows for arbitrage. In particular, If $(\mathbf{p}+\varepsilon)\hat{\mathbf{z}} > 0$, then producing $(\mathbf{x}(\mathbf{z}(\mathbf{p})), \mathbf{z}(\mathbf{p}) + t\hat{\mathbf{z}})$, where t is such that $\mathbf{z}(\mathbf{p}) + t\hat{\mathbf{z}}$ is in the boundary of Z maximizes profits , and similarly if $(\mathbf{p}+\varepsilon)\hat{\mathbf{z}} < 0$ with $-t\hat{\mathbf{z}}$. Note that by Lemma 1, $\mathbf{z} \pm t\hat{\mathbf{z}} \in \mathbf{z}(\mathbf{p})$ as $\mathbf{z} \in \mathbf{z}(\mathbf{p})$, and $\hat{\mathbf{z}} \in \hat{Z}(\mathbf{p})$.

4.1 Piece-wise linear contracts

The results derived above extend to piece-wise linear contracts. To examine such contracts, we define a *multi-part pricing rule* as consisting of a partition of the output space $\Re^M$ into disjoint convex subsets $Z_1, Z_2, ..., Z_K$ and, for each $k = 1, ..., K$, a pricing rule $y^k : Z_k \rightarrow \Re$ implemented by a fixed payment y_0^k and a price vector $\mathbf{p}^k$ such that

$$y^k(\mathbf{z}) = y_0^k + \mathbf{p}^k\mathbf{z}. \tag{6}$$

Define an interior solution for Z_k as a pair $(\mathbf{x}, \mathbf{z})$ in the interior of $X \times Z_k$ such that $\mathbf{z} \in \arg\max\{y_0^k + \mathbf{p}^k\mathbf{z} - g(\mathbf{x}) : \mathbf{z} \in Z(\mathbf{x}) \cap Z_k\}$. Then, following the arguments given above, an interior solution for Z_k exists only if $\hat{Z}(\mathbf{p}^k) \neq \{\mathbf{0}^M\}$.

In particular, this analysis is applicable to the case when the principal requires the agent to product non-negative outputs of each output i by setting $p_i = \infty$ whenever $z_i = 0$. Letting p^1 be the price vector for the non-negative orthant, the selected output will be on the boundary of $Z_1 = \Re_+^M$ whenever $\hat{Z}\left(\mathbf{p}^1\right) \neq \{\mathbf{0}\}$.

5 A robust incentive model

This section considers a general class of incentive structures in which sharpness is sacrificed for robustness. These structures might be useful for designing contracts when the principal does not have access to a technology to monitor output in real time or when such access is costly. That is, when complex multi-part tariffs cannot be relied upon to eliminate arbitrage opportunities that might arise from random output prices.

Here we define an incentive structure as a pair $(\mathbf{p}, y_0)$ associated with a payment rule

$$y\left(\mathbf{z}; \mathbf{p}, y^0\right) = \mathbf{pz} + \mathbf{y}_0.$$

For any compact convex set $\hat{P} \subseteq \Re^{M+1}$, with typical element $(\mathbf{p}, y_0)$

$$\hat{y}\left(\mathbf{z}; \hat{P}\right) = \min_{(\mathbf{p}, y_0) \in \hat{P}} \mathbf{pz} + \mathbf{y}_0$$

we are able to find arbitrage-fee mechanisms. In particular, a sufficient condition for the absence of arbitrage is that, for some y_0^*, $(\mathbf{p}^*, y_0^*) \in P$. To see this, notice that for any feasible $\mathbf{z}$:

$$\begin{aligned} \hat{y}\left(\mathbf{z}; \hat{P}\right) &\leq \mathbf{p}^*\mathbf{z} + y_0^* \\ &\leq \mathbf{p}^*\mathbf{z}^* + y_0^*. \end{aligned}$$

Relevant examples include the case where $\hat{P}$ is the convex hull of a finite set of points or when $\hat{P}$ is a scalar multiple of the unit ball for some L^p norm. Moreover, when $M = 1$, this class of incentive structures consists of concave piecewise linear payment schedules. More generally, any concave piecewise linear payment schedule can be expressed this way.

We note that the conservatism of $\hat{P}$ with respect to any feasible output $\mathbf{z}$ may be represented by the width

$$w\left(\mathbf{z}, \hat{P}\right) = \left(\max_{(\mathbf{p}, y_0) \in \hat{P}} \mathbf{pz} + \mathbf{y}_0\right) - \left(\min_{(\mathbf{p}, y_0) \in \hat{P}} \mathbf{pz} + \mathbf{y}_0\right).$$

Thus, the optimal solution is likely to be a trade-off between precision – incentives that match actual values and keep output close to optimal – and arbitrage-avoidance, which requires setting more diffuse incentives. The optimal price set will be more diffuse, the more flexible the technology and the greater the possible error in prices. Our contribution is noting the role of flat technology and the analogy with financial arbitrage.

6 Discussion

Our main result establishes that it might not be possible to have sharp incentives as implemented by a linear contract. The reason lies in the prices that the agent faces. If these prices are random – even where the random component consists of white noise – then the agent can gain dramatically by a large deviation from first-best level of efforts. Note that even a slight difference between predicted and realized prices can lead to a very inefficient outcome under any linear contract. Our analysis complements that of Holmström and Milgrom (1991) who show that multi-tasking also leads to the non-optimality of linear contracts as the compensation on any subset of tasks results in a shift of activities towards those that are compensated and away from those that are not.

The recognition that incentives matter has lead to a revolution in the way that firms and governments deliver goods and services. At the same time, the widespread use of incentive schemes can often lead to outcomes that are either unexpected or undesirable. In this paper we examine a polar case where the production technology is flat and where output prices are random. Under these circumstances, incentive schemes can lead to extremely undesirable outcomes. In particular, agents will choose their effort to arbitrage away the profitable opportunities that arise from the differences between contract and actual prices. In the polar case that we examine, arbitrage leads to corner solutions, where the agent will produce as much output as possible under the existing technology and given the limited resources it has access to. It is not difficult to see, however, that the possibility of arbitrage distorting the agent's incentives will exist even under more general technologies.

Our analysis suggests that, in the presence of flexible technology and random prices, an optimally designed contract might need to trade-off sharpness or precision and arbitrage-avoidance. More precisely, the optimal contract price will be more diffuse, the flatter the technology and the greater the possible error in prices.

References

[1] Baker, G., 1992, "Incentive contracts and performance measurement," *Journal of Political Economy* 100(3), 598–614.

[2] Baker, G., R. Gibbons, and K. J. Murphy, 1994, "Subjective performance measures in optimal incentive contracts," *Quarterly Journal of Economics* 109, 1125–1156.

[3] Burgess, S. and M. Ratto, 2003, "The role of incentives in the public sector: Issues and evidence," CPMO Working Paper Series 03/071, Centre for Market and Public Organisation, University of Bristol.

[4] Chambers, R. G. and J. Quiggin, 2000, *Uncertainty, Production, Choice and Agency: The State-Contingent Approach*, Cambridge University Press, New York.

[5] Courty, P. and G. Marschke, 2004, "An empirical investigation of gaming responses to explicit performance incentives," *Journal of Labour Economics* 22(1), 23–56.

[6] Dixit, A., 2002, "Incentives and organisations in the public sector: An interpretative review," *Journal of Human Resources* 37(4), 696–727.

[7] Webster, E. and G. Harding, 2000, "Outsourcing public employment services: The Australian experience," *Melbourne Institution Working Paper* no.4/00.

[8] Holmström, B., 1979, "Moral hazard and observability," *Bell Journal of Economics* 10(1), 74–91.

[9] Holmström, B. and P. Milgrom, 1991, "Multi-task principal–agent analyses: Linear contracts, asset ownership and job design," *Journal of Law, Economics and Organization* 7, 24–52.

[10] Kerr, S., 1975, "On the folly of rewarding A while hoping for B," *Academy of Management Journal* 18, 769–783.

[11] McKim, A. and M. Hughart, 2005, "Staff Incentive Schemes in Practice: Findings from a Global Survey of Microfinance Institutions," *MicroFinance Network Working Paper*.

[12] Milgrom, P. and J. Roberts, 1992, *Economics, Organization and Management*, Prentice–Hall, New Jersey.

[13] Mirrlees, J. A., 1999, "The theory of moral hazard and unobservable behaviour: Part 1," *Review of Economic Studies* 66(1), 3–21.

[14] Prendergast, C., 1999, "The provision of incentives in firms", *Journal of Economic Literature* 37, 7–63.

[15] Walls, M., 2003, "How local governments structure contracts with private firms: Economic theory and evidence on solid waste and recycling contracts," *Resources for the Future Working Paper 03-62* .