INCENTIVE PLANS FOR PRODUCTIVE EFFICIENCY, INNOVATION AND LEARNING

PER J. AGRELL, PETER BOGETOFT, AND JØRGEN TIND

Abstract. In many industries where production or sales is delegated to a number of subunits, the central management faces the classical problem how to induce continous efficiency improvements, organizational learning and transfer of knowledge with a minimum of control exercised. This paper draws on recent results from Agrell, Bogetoft and Tind (1999) regarding regulatory frameworks to construct simple, yet powerful incentive schemes for decentralized production under asymmetric information. The theoretical foundation is based on principal-agent theory (cf. Laffont and Tirole, 1986) and extensions to production theory by Bogetoft (1994). The proposed incentives system is operational and makes use of available information to provide positive incentives for participation in the dynamic development of the entire organization.

1. Introduction

How many of us do not share the sudden experience of discovering a new shortcut with the latest software, a smarter way to round up a problem? A small everyday innovation that enables us to save time and effort. But, what do we do next? Sharing our illuminated state of knowledge may simply spoil its fruits, without ever harvesting them. A better alternative might be to relax for a second, or to use the freed-up time for that favorite task that never gets done? This simple reasoning illustrates one of the most challenging, complex and frustrating missions of modern management, in the presence of increasingly qualified tasks and information systems.

This paper addresses the problem of performance assessment, organizational learning, innovation, information transfer and incentives management in delegated decision making. Decentralized systems are found everywhere in private business, from the franchising copy-shops and fast-food restaurants to the world-wide offices of management consultants. In a world where constant change is the rule and stability an exception, they all face the need to adapt rapidly to changing circumstances, expectations and opportunities. Although a profit maximizing behavior may be induced by profit sharing schemes, such contracts may easily lead to suboptimal levels of organizational training, innovation and knowledge transfer. It is the role for the central management to establish incentive systems such that these globally vital functions are promoted and maintained. Indeed, such systems are commonplace, from cost-saving sharing schemes to the 'good idea' cash bonus. In

Date: July, 2000.
Key words and phrases. Data Envelopment Analysis, incentive planning, innovation.
This paper is accepted for publication in the *International Journal of Production Economics*.

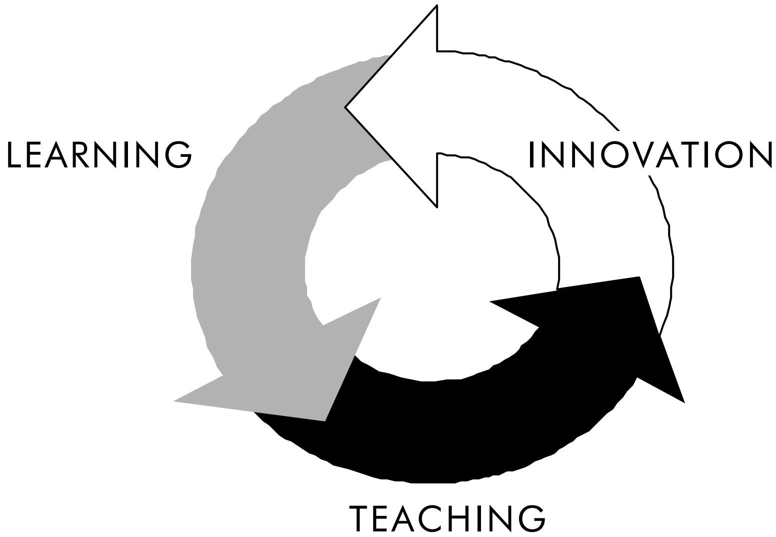

FIGURE 1. The dynamics of organizational development: learning, innovation and teaching as archetypical functions.

this paper, the proponents of organizational development and innovation find theoretically founded models to explain observed behavior and to construct incentive schemes for complex multi-task environments.

Extending the viewpoint to address organizational development and managerial control, the approach is a theoretical framework for the three archetypical roles in performance development, *learning, teaching* and *innovation* (Figure 1), naturally following each other during the life span of a project, organization or individual. For performance measurement as well as incentive systems, failure to promote all of the roles may hamper or lower the overall efficiency of the organization.

The contributions of the paper are three-fold.

The paper links the *management science literature* on productivity analysis, in particular the Data Envelopment Analysis (DEA) literature, with the general economic literature on decision making and incentive provision.

As a contribution to the *productivity analysis literature*, the distinctive feature is the relation to incentive systems in a multi-period setting and its relation to the cost-norm as a temporary measure. Previous models in this field have treated the technology as given and restricted its consideration to catch-up of inefficiency. This means that we can formally and practically evaluate the usefulness of a DEA based benchmarking from the point of inducing cost reduction beyond norm while still reducing the information rents and providing optimal incentives. Previous studies of the incentive effects of DEA-like efficiency measures include the moral hazard models in Bogetoft (1994, 1995) and the adverse selection models in Bogetoft (1997, 2000). In the adverse selection papers, which are most directly related to the present, the previous studies focus at asymmetric information about the technology.

In the *general economic literature*, the significance of information asymmetries, innovation and strategic behavior has long been recognized. There exists a considerable literature on how incentives affect a variety of management problems and the methodology for analyses of incentive problems, most notably the principal-agent model, is by now quite well-established, see e.g. Hart and Holmström (1987). For tractability, however, most applications of principal-agent models have restricted the dimensionality of the problems and assumed a fixed technology, and thus also the applicability of the results.

The outline of the paper is as follows: Section 2 defines the productivity analysis model that forms the backbone of the approach. Section 3 describes a single-period incentive system to provide the building blocks for the system. Section 4 develops gradually a two-period model, comprising elements for the ratchet effect, the innovation process and the information dissemination process, into a multi-period incentive system. The paper ends with some concluding remarks in section 5.

2. The DEA-Benchmarking Model

The idea behind the frontier benchmarking model is to take multiple inputs and outputs into account in order to compare decision making units (DMU).

To formalize the above, we assume[1] that each of n DMUs, say DMU^i transform a verifiable cost x^i into m outputs y^i.

We assume that the technological possibilities are the same for all DMU. Specifically, these possibilities may be thought of as the set T of feasible input -output combinations

$$T = \{(x, y)|x \text{ can produce } y\}$$

We shall generally assume that T satisfy

Condition 1. *Free disposability:* $(x, y) \in T, x' \geq x, 0 \leq y' \leq y \Longrightarrow (x', y') \in T$.

Condition 2. *Convexity: T is convex.*

Condition 3. *r returns to scale, $(x, y) \in T \Longrightarrow (qx, qy) \in T, \forall q \in K(r)$, where $r =$ "crs", "drs," or "vrs", and $K(crs) = \Re_0, K(drs) = [0, 1]$ and $K(vrs) = \{1\}$, respectively.*

The associated underlying cost model for a DMU is given by

$$C(y) = \min_x \{x|(x, y) \in T\}$$

Given a reference set $G = \{1, ..., n\}$ with indexes for observations of feasible production plans (x^i, y^i) the *DEA based cost norm* $C^{DEA}(.|G) : \mathbb{R}_0^m \to \mathbb{R}$ is defined as

$$\begin{array}{rll} C^{DEA}(y|G) = & \min & x \\ & x, \lambda & \\ & s.t. & x \geq \sum_{i \in G} \lambda^i x^i \\ & & y \leq \sum_{i \in G} \lambda^i y^i \\ & & \lambda \in \Gamma(r) \end{array} \tag{2.1}$$

where $\Gamma(crs) = \Re_0^n, \Gamma(drs) = \left\{\lambda \in \Re_0^n | \sum_i \lambda^i \leq 1\right\}, \Gamma(vrs) = \left\{\lambda \in \Re_0^n | \sum_i \lambda^i = 1\right\}$. We see that the DEA based cost function for any output vector gives the maximum cost of producing the output that is consistent with the observations in G.

To develop the setting into a full model we shall make a few additional assumptions.

[1]The presentation may be generalized to the multiple-input and non-controllable variable case, as in Agrell, Bogetoft and Tind (1999).

We assume that the DMUs actual cost in the planning period is the currently minimal cost $C(y)$ plus potential slack $s \in \mathbb{R}_0$ minus the effects of a possible innovation $u \in \mathbb{R}_0$ introduced in the production process, i.e.

$$x = C(y) + s - u$$

Note that both production slack and the innovation effect are summarized here as one-dimensional costs and revenues (cost-reductions). The agent (DMU) knows $C(y)$ but the principal (management) does not. The principal does, however, know the actual costs and outputs in the feasible production plans of the set G, i.e. he has information

$$(x^i, y^i) \in \mathbb{R}_0^{1\times m} \qquad i \in G$$

These plans may be the cost-production data from previous periods or from similar agents. They may also be feasible plans derived from engineering or other studies. Using the information about the production plans, the principal can infer with certainty that for the given state of technological development

$$C(y) \leq C^{DEA}(y|G) \ \ \forall y \tag{2.2}$$

This follows from the so called minimal extrapolation property of the DEA model, Banker et al. (1984). In this context with innovation, the cost function is to be seen as a temporal state-of-the-art benchmark, an upper cost bound for a given technology, not a limit for an asymptotic development. The distinction will be used later on in the dynamic setting. The principal has no more certain information about the cost structure. Formally, we let the principal's belief about the likelihood of the different cost functions be given by the probability distribution $p(.)$ on the class $\mathcal{C}$ of increasing convex r return to scale functions satisfying (2.2) and the scale assumptions of (2.1). The belief distribution represents whatever additional information the principal has and it is used to close the model as a Bayesian Game.

To model the output requirements, the demand or benefit side, let $Y \subseteq \mathbb{R}_o^m$ be a set of productions plans that are feasible for the agent and acceptable to the principal in the planning period. The aim of the principal is to minimize the costs of inducing the agent to accept employment and to select an acceptable production $y \in Y$. For simplicity, we assume that there is no difficulty in observing the fulfillment of y ,i.e., there is no moral hazard in the choice of output.

The DMU or agent on the other hand seeks to maximize his profit and slack and to minimize the effort on innovation, weighted together to give his utility

$$U(x, s, u) = (b - x) + \rho s - \gamma u = (b - x) + \rho \left[x - C(y)\right]_+ - \gamma \left[C(y) - x\right]_+$$

when $y \in Y$ is the implemented production plan, $b \in \mathbb{R}$ is the budget or the monetary transfer from the principal to the agent, x is the actual cost, and $[a]_+ = \max\{a, 0\}$ for $a \in \mathbb{R}$. Here, ρ is a fixed parameter that describes the agent's value of slack relative to profit. Since slack can only be consumed "on-the-job", we suggest that profit is more valuable and we therefore assume that $\rho \in [0, 1]$. The parameter γ denotes the agent's cost of innovation. Since the innovation we are concerned with such that it does not pay off immediately, $\gamma > 1$. The agent's reservation utility, i.e. the utility he requires to accept working for the principal, is assumed to be 0.

From a social point of view it is important which production plans are selected under which conditions. For a given cost function $C = C(.) \in \mathcal{C}$ let $x[C]$ be the cost chosen by the DMU and let $y[C]$ be the production plan that is implemented. The choice of $x[C]$ reflects the agent's strategic behavior[2], such as consumption of slack in case of $x[C] > C(y[C])$ and innovation in the case $x[C] < C(y[C])$. An outcome is said to be *cost efficient* if and only if

$$x[C] = C(y[C]) \quad \forall C \in \mathcal{C}$$

such that outputs are produced without cost slack, i.e. at minimal cost.

There is a simple but useful dual relationship between the set of possible cost functions $\mathcal{C}$ and the DEA based cost function C^{DEA}. We record this as a lemma.

Lemma 1. *We have*

$$C^{DEA}(y|G) = \max_{C \in \mathcal{C}} C(y) \quad \forall y \in \mathbb{R}_0^m$$

Proof. Simple extension of Lemma 1 in Bogetoft (2000) □

By Lemma 1, the DEA estimated cost $C^{DEA}(y|G)$ is the highest possible cost of producing y that can be claimed given the hypothesis $C \in \mathcal{C}$, i.e. given the assumption of a increasing, convex, r return to scale cost function and given the cost-production data (x^i, y^i), $i = 1, .., n$.

3. Single-period incentive system

We will now devise an incentive system for the single-period problem to induce the agent to exercise effort in catching up inefficiencies. The *single period incentive problem* (P_{SP}) therefore becomes

$$\begin{array}{llll}
\min\limits_{y, b, x(C)} & \Sigma_{c \in \mathcal{C}}\ b[x[c]]p(c) & & \\
s.t. & V(b, x[C], y) \geq 0 & \forall C \in \mathcal{C} & (IR) \\
 & V(b, x[C], y) \geq V(b, x', y) & \forall C, x' : x' \in [0, b[x']] & (IC) \\
 & b[x'] \in \mathbb{R},\ y \in Y & \forall x' \in \mathbb{R} &
\end{array}$$

with the net resulting utility function for the agent

$$V(b[.], x, y) = b[x] - x + \rho [x - C(y)]_+ - \gamma [C(y) - x]_+$$

The individual rationality constraints (IR) ensures that the whatever the underlying cost structure is, the DMU is guaranteed a positive net reward. Hence, he is willing to participate and to use the cost strategy $x[C]$. The incentive compatibility constraints (IC) ensure that this strategy is in fact the best possible strategy to use. By deviating and choosing any other cost level x' the net reward cannot increase. The principal tries to minimize the resulting expected payments to the DMU subject to these constraints.

Next, we rule out the situation that a DMU would find it optimal to innovate and consume slack simultaneously. Note that we do not require the DMU to implement at least possible cost, i.e. to be cost efficient. This however is an induced property according to the following lemma and its corollary.

[2] Since we have assumed that y is observable without noise, the choice of y is effectively at the principal's discretion. Thus, we need not consider the potential strategic element involved in selecting the output y , when the agent consciously produces an acceptable output for which the cost norm $C^{DEA}(y)$ is less precise.

Lemma 2. *There exists an optimal solution to the single-period problem* (P_{SP}) *which has neither consumption of slack, nor innovation, i.e.,*

$$s = u = 0$$

Proof. Assume that there exists a solution such that $s \geq 0$ and $u \geq 0$ and at least one of them positive. Then $x\,(C) = C\,(y) + s - u$

$$U = b - x(C) + \rho s - \gamma u = b - C\,(y) - (1-\rho)\,s - (\gamma - 1)\,u \leq b - C(y)$$

with strict inequality for $\rho < 1$. Hence, the agent prefers a solution with $s = u = 0$. □

Intuitively, slack is never strictly attractive since 1\$ of slack costs the principal 1\$ but it is only worth ρ\$ to the agent.

Corollary 1. *There exists an optimal solution to the single-period problem* (P_{SP}) *which is temporally cost efficient.*

Proof. Follows immediately from Lemma 2. □

Without loss of generality, therefore, we may restrict attention to temporally cost efficient solutions. Our next proposition characterizes the solution to this contracting problem with verifiable costs x.

Proposition 1. *An optimal solution* $(y^{SP}, b^{SP}[x], x^{SP}\,[C])$ *to the single-period contract design problem* (P_{SP}) *is given by*
$x^{SP}[C] = C(y)$ *(cost efficiency)*
$b^{SP}[x] = x + \rho[C^{DEA}(y^{SP}|G) - x]$ *(DEA-yardstick[3])*
$y^{SP} = \arg\min_{y \in Y} \rho C^{DEA}(y|G) + (1-\rho)[\sum_{C \in \mathcal{C}} C(y)p(C)]$.

Proof. The cost efficiency is given from Corollary 1 . Since innovation is not profitable to the principal, even in absence of incentive costs, the case is analogous to Bogetoft (2000). Thus, the two remaining claims draw directly on the proof in Bogetoft (2000), omitted here due to its level of detail. □

Proposition 1 shows that the optimal single-period arrangement leads to cost efficient production. This means that the agent does not consume slack since the principal can take advantage of the agent's preference for profit. According to the payment scheme above, the principal now pays the actual cost plus the fraction ρ of the amount that the agent saves compared to the DEA yardstick $C^{DEA}(y|G)$.

In our framework, the single-period incentive system induces efficiency and (costless) learning since the agent strives to keep up with the frontier. In many respects, this system is similar to the yardstick and benchmarking practices that are omnipresent in private business. The single-output case is also a well known special case of the general adverse selection problem (e.g. Laffont and Tirole, 1986). However, the optimal single-period scheme rules out innovation, since the principal will lose at least $\gamma - 1$ and the agent $\gamma - \rho$ per unit of innovation, respectively. Next, we refine the incentive system by introducing the temporal aspect to make innovation, learning and teaching profitable.

[3]This property is similar to the one found to be optimal in Bogetoft(1997). There actual costs were verifiable also but the aim was to minimize the expected costs to the principal of inducing the agents to accept employment and to minimize the production costs. The latter is not required - but derived - here.

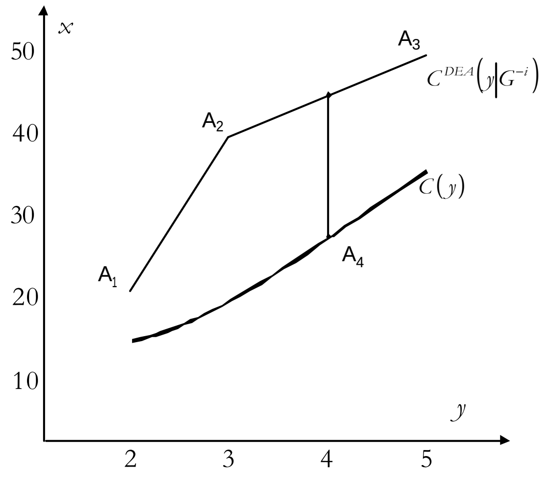

FIGURE 2. Obervations A_1, A_2, A_3, cost function $C(y)$ and the DEA cost norm $C^{DEA}(y)$ in Example 1.

Example 1. *For simplicity, assume a one-dimensional output $m = 1$, $\rho = 0.5$, variable returns to scale $r = vrs$ and three units $n = 3$ in the comparison set G. Let $(x^i, y^i) = \{(2,30), (3,40), (5,50)\}$ be the observations, forming the cost norm $C^{DEA}(y)$ in Figure 2 below. The units know the true cost function $C(y) = 10 + y^2$. A unit (A_4) is asked to produce $y = 4$. (P_{SP}) gives that the unit cannot consider producing above the cost norm, thus $x \leq C^{DEA}(y) = 0.5C^{DEA}(3) + 0.5C^{DEA}(5) = 45$. Producing at $x = 45$ would give a slack $s = 45 - C(4) = 19$ and a reimbursement $b = 45$, or $U(45, 19, 0) = (45 - 45) + \rho 19 = 9.5$. The optimal contract from Proposition 1 makes the agent indifferent between this solution and a cost-efficient solution $x = 26$, that is $b = 26 + 0.5(45 - 26) = 35.5$ and $U(26, 0, 0) = (26 + 0.5(45 - 26)) - 26 = 9.5$.*

4. Two-period Incentive System

By introducing two time periods, say 0 and 1, the multi-period incentive problem will be addressed while keeping the notation to a minimum. Throughout, the objective of the agent (DMU^i) will be to maximize the sum of single-period utilities, i.e.,

$$W(b, x) = V(b_0, x_0, y_0) + V(b_1, x_1, y_1)$$

where $b \in \mathbb{R}^2$, $x \in \mathbb{R}^2_0$, $y \in \mathbb{R}^2_0$. The objective of the principal will analogously be to minimize the total cost of production and incentives, while inducing his participation and truthfulness. To simplify the notation, assume that the production decisions y_0 and y_1 are exogenously given and that

The *two-period (dynamic) incentive problem* (P_{DP}) therefore becomes

$$\begin{array}{llll}
\min\limits_{b, x_0, x_1} & \Sigma_{C \in \mathcal{C}}\, p(C) \{b_0(x_0[C]) + b_1(x_0[C], x_1[C])\} & & \\
s.t. & W(b(x_0[C_0]), x[C]) \geq 0 & \forall C \in \mathcal{C} & (IR) \\
 & W(b(x[C]), x[C]) \geq W(b(x'), x') & \begin{array}{l}\forall C \in \mathcal{C} \\ \forall x'_t \in [0, b_t[x']]\end{array} & (IC) \\
 & b[x] \in \mathbb{R} & \forall x' \in \mathbb{R}^2, t = \{0, 1\} &
\end{array}$$

Below, we will deal with primarily three problems that are actualized by the introduction of the time dimension: the ratchet effect, the innovation incentive and the dissemination problem.

4.1. **The Ratchet Effect.** The well known tendency for a norm to be tightened after demonstrated good performance is known as the *ratchet effect*. Certainly staff would be reluctant to mount an extraordinary effort in one period if their good result would be used to establish new higher performance standards. Apparently an important and prevailing problem in performance measurement (Laffont and Tirole, 1986), it may be counteracted by excluding the studied unit from the reference set of its cost norm. Let the reimbursement for the DMU^i be $b_t\left(x_t\right)$ in period t,defined as

$$b_t\left(x_t\right) = x_t + \rho[C^{DEA}(y|G_t^{-i}) - x_t]$$

where $G_t^{-i} = \{(j,\tau)\,|j \neq i,\, 0 \leq \tau \leq t\}$, and x_t is the actual cost in period t. The cost norm $C^{DEA}(y|G_t^{-i})$ is based on $t\,(n-1)$ previous observations (x_τ^j, y_τ^j), excluding DMU^i. We state the result in a proposition.

Proposition 2. *An optimal solution to problem* (P_{DP}) *without innovation is given by*

$$x_t[C] = C(y_t)$$
$$b[x_t] = x_t + \rho[C^{DEA}(y_t|G_t^{-i}) - x_t]$$

Proof. Extension of the "exclusion principle" from Bogetoft (1994). □

4.2. **The Innovation Process.** Assume that DMU^i at the outset of period 0 has the opportunity to undertake an innovation, implying a cost reduction of $u \in \left]0, \overline{u}\right]$ beginning from period 0,at an immediately occurring cost γu, $\gamma > 1$. DMU^i faces two alternatives at $t = 0$: (i) not to innovate, (ii) to innovate. Subsequently at $t = 0$ and $t = 1$, the agent must decide (iii) to hide the innovation, or (iv) to disclose the innovation. Although the agent, e.g., may decide to undertake the innovation, hide it at $t = 0$ and reveal it at $t = 1$, the only interesting options are (1) not to innovate, (2) to innovate and hide both periods, and (3) to innovate and reveal both periods. To simplify notation, denote the current cost norms $c_0 = C\left(y_0\right)$, $C_0^{DEA} = C^{DEA}(y_0|G_0^{-i}), c_1 = C\left(y_1\right)$ and $C_1^{DEA} = C^{DEA}(y_1|G_1^{-i})$ and denote the incremental reimbursement for announced innovations with b_t^u for $t = 0, 1$.

4.2.1. *Option 1: No innovation.* Since slack is suboptimal, the decision $s_0 = s_1 = 0$, $x_0 = c_0$ and $x_1 = c_1$will imply the resulting two-period pay-off $V = \widetilde{V} = \rho\left(C_0^{DEA} - c_0\right) + \rho\left(C_1^{DEA} - c_1\right)$ for the agent and Z for the principal.

4.2.2. *Option 2: Secret innovation.* The second option would give the unit the chance to enjoy the benefits of the innovation in slack without being detected. The decision is $s_0 = s_1 = u$, $x_0 = c_0$ and $x_1 = c_1$ with total pay-off $V = \widetilde{V} + (\rho - \gamma)\,u + \rho u$ for the agent and Z for the principal.

4.2.3. *Option 3: Public innovation.* The third option, $s_0 = s_1 = 0$, $x_0 = c_0 - u$ and $x_1 = c_1 - u$ the would yield pay-off $V = \widetilde{V} + b_0^u + b_1^u - \gamma u$ for the unit and $Z - b_0^u - b_1^u + 2u$ for the principal.

This leads to the following simple observations regarding the outcome:

Proposition 3. *(i) Innovation is never undertaken if* $\gamma > 2$

(ii) Secret innovation is only a possibility if $\rho \geq \frac{\gamma}{2}$

(iii) There exist payments b_0^u, b_1^u *such that a mutually beneficial announced innovation exists for* $2\rho \leq \gamma \leq 2$

(iv) Public innovation is superior if $\rho \geq \frac{\gamma}{2}$

Proof. (i) Private innovation is never attractive since $2\rho - \gamma < 2 - \gamma < 0$.

Public innovation requires $b_0^u + b_1^u \geq \gamma u$ for the agent to accept and $b_0^u + b_1^u \leq 2u$ for the principal to be justified. I.e., for $\gamma > 2$, nor public innovation is a possibility.

(ii) To enable a secret innovation, we must have $(\rho - \gamma)\, u + \rho u \geq 0 \Longleftrightarrow 2\rho \geq \gamma \Longleftrightarrow \rho \geq \frac{\gamma}{2}$.

(iii) For $2\rho \leq \gamma$ the agent cannot threaten to innovate secretly. Hence, the principal can persuade him to innovate if $b_0^u + b_1^u - \gamma u \geq 0 \Longleftrightarrow b_0^u + b_1^u \geq \gamma u$. The principal is interested if $b_0^u + b_1^u \leq 2u$ and since $\gamma \leq 2$, there exist possible b_0^u, b_1^u values.

(iv) For $\rho > \frac{\gamma}{2}$, secret innovation is a threat. Thus, to induce public innovation the principal must provide payments b_0^u, b_1^u such that

$$b_0^u + b_1^u - \gamma u \geq 2\rho u - \gamma u \Longleftrightarrow b_0^u + b_1^u \geq 2\rho u \tag{4.1}$$

The principal is interested in public innovation only if

$$b_0^u + b_1^u \leq 2u \tag{4.2}$$

Since $\rho \leq 1$ there exist payments b_0^u, b_1^u such that both (4.1) and (4.2) are fulfilled. □

Remark 1. *The optimal innovation level* $u = \overline{u}$ *whenever innovation is undertaken.*

Proof. Follows from Proposition 3, where all benefits are directly proportional to u. □

The findings in the Proposition are summarized in Figure 3 below, depicting the acceptance regions for various options in the (γ, ρ)-space. The credible threat of secret innovation by the agent is found in the upper left corner, that is for higher values of the slack appetite ρ and lower innovation costs. Note the important delimiter $\rho = \frac{1}{2}\gamma - 1$ between the zone for costless public innovation, lower left corner, and the zone in which the agent may collect an additional benefit of $(2\rho - \gamma)\, u$ for a public innovation.

Table 1 Acceptance conditions for innovation in the two-period model.

	Acceptance, profit	
Condition	Manager	DMU
$\gamma > 2$	No, Z	No, $\widetilde{V}$
$2\rho \leq \gamma \leq 2$	Yes, $Z + (2 - \gamma)\, u$	Yes, $\widetilde{V}$
$1 < \gamma \leq 2\rho$	Yes, $Z + 2\,(1 - \rho)\, u$	Yes, $\widetilde{V} + (2\rho - \gamma)\, u$

The results indicate that there is a delicate balance between the threshold to innovate, expressed by the parameter γ, and the intensity of managerial control, as given by ρ. Without the proper level of control, even costless training or encouragement of innovation may give the subunits a strategic advantage against the management. Since the innovative ability usually comes at some cost, e.g., higher

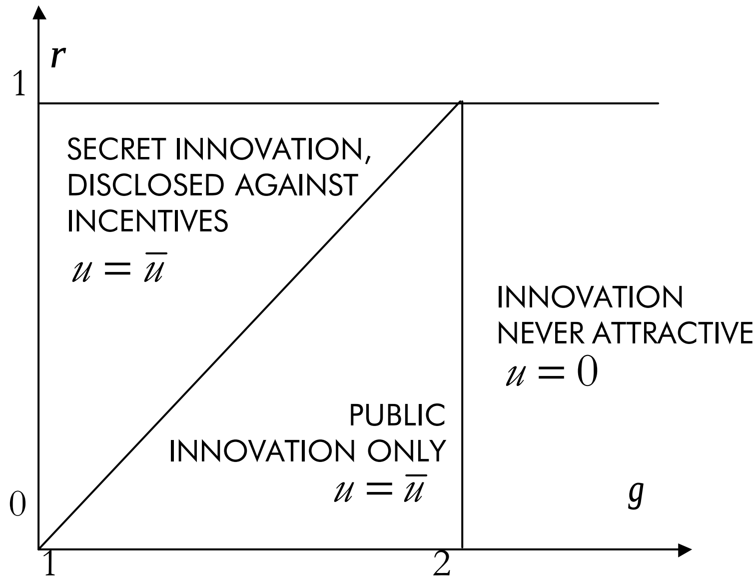

FIGURE 3. Acceptance regions and optimal innovation levels u for options 1, 2, 3 in the (γ, ρ)-space.

educational level or access to costly equipment or staff, the management may actually be facing negative returns on such investments. We state the optimal level of training (or reduction of innovation cost) in a corollary.

Corollary 2. *The optimal level of innovation cost* γ *given a valuation of slack* ρ *is* $\gamma = 2\rho$.

Proof. Follows from the optimal incentive plan b_0^u, b_1^u in Proposition 3 where the potential additional cost to avoid secret innovation is $(2\rho - \gamma)\,u$. □

As indicated in the Corollary, an agent that receives excessive training, $\gamma < 2\rho$, will use the acquired advantage to extract information rents from the principal. In an applied setting, we need creative and intelligent employees, but excesses in this regard threaten management authority. Inversely, the Corollary may also be used to determine the optimal level of control, which is related to the valuation of slack ρ.

Consider the case in Figure 4, where the rewards to the principal and the agent, respectively, are graphed as a function of the innovation cost, γ. To illustrate the effect of the ability to enjoy slack, compare $\rho_A < \rho_B$ and their respective rewards. As ρ increases, the lower the rewards for the principal for a given γ. For $\rho > \frac{3}{4}$ the agent captures more of the innovation gain than the principal, for $\frac{1}{2} \leq \rho < \frac{3}{4}$ less than the principal and for low values ($\rho < \frac{1}{2}$) the agent has no positive rent. The optimal points of training, A and B, are demarcated with circles in Figure 4. It is evident that in absence of control, that is a high ρ, an investment in innovation, that is lowering γ, risks placing the principal in a hold-up situation.

4.3. **The Dissemination Process.** The third problem results from lack of common incentive components in the organizational model. Although the current model

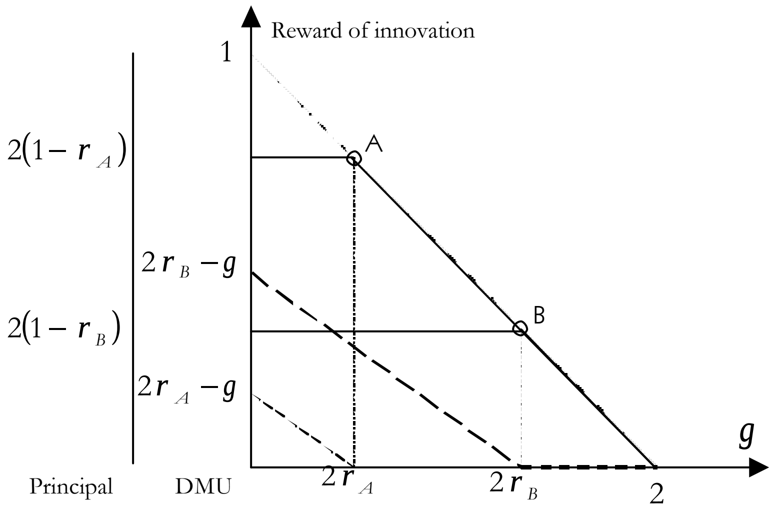

FIGURE 4. Innovation gains for principal and agent as a function of innovation cost γ. Example $\frac{1}{2} \leq \rho_A \leq \frac{3}{4}$ and $\frac{3}{4} \leq \rho_B \leq 1$.

assumes that such transfer could be done at zero cost, the construction of the cost norm discourages such transfer through an implicit ratchet effect. Thus, when unit i lowers the cost norm $C(y)$ through innovation, it increases the distance between the frontier that includes their performance $C^{DEA}(y|G)$ and the frontier used in the incentive system $C^{DEA}\left(y|G^{-i}\right)$. Since the latter is used to reimburse the unit, there are positive disincentives not to disseminate the innovation. The effect is compensated by a term b_1^d for the innovator during the transfer period 1.

For the model, assume that a unit i announcing an innovation in period 0 undertakes an observable dissemination of knowledge in period 1. Upon receiving the information, the other units have the possibility to benefit from the innovation by immediately lowering their costs. To distinguish the options, denote the reference set with dissemination by G_{1d}. Ceteris paribus, it follows that $C^{DEA}\left(y|G_0^{-i}\right) = C^{DEA}\left(y|G_1^{-i}\right) \geq C^{DEA}\left(y|G_{1d}^{-i}\right)$,i.e., dissemination can only lower costs. Releasing the ceteris paribus condition introduces the problem of disentangling dissemination and innovation. The cost norm for the $n-1$ other units may be lowered due to their own efforts rather than dissemination. However, although part of the problem is inevitable in practice, the assumption of observability saves the integrity of the model.

We state the result in the form of a proposition.

Proposition 4. *A DMU is willing to accept a contract to disseminate knowledge about an undertaken innovation iff*

$$b_1^d \geq \rho\left(C^{DEA}\left(y|G_0^{-i}\right) - C^{DEA}\left(y|G_{1d}^{-i}\right)\right)$$

Proof. The compensation awarded to the innovator in period 1 is

$$x_1 + \rho\left(C^{DEA}(y|G_1^{-i}) - x_1\right) + b_1^u + b_1^d$$

compared to the alternative compensation without dissemination

$$x_1 + \rho\left(C^{DEA}(y|G_0^{-i}) - x_1\right) + b_1^u$$

which yields the difference

$$\rho\left(C^{DEA}(y|G_{1d}^{-i}) - C^{DEA}(y|G_1^{-i})\right) + b_1^d \geq 0$$

and hence the proposition. □

The principal is willing to offer such contract since the reduction of cost for the other units only partially is reimbursed to the disseminating unit. Once again, the principal may use the fact that the agent has a preference for profit to slack. The effect is depicted in Figure 5 below where unit B discloses an innovation in period 0 and the principal observes realized cost reductions by unit A after knowledge transfer.

This problem may well occur in many settings where profit-driven subunits refuse to spend time and resources to disseminate technological advances. Hence, it is the task of the incentive system to compensate and encourage such development in the interest of the entire organization.

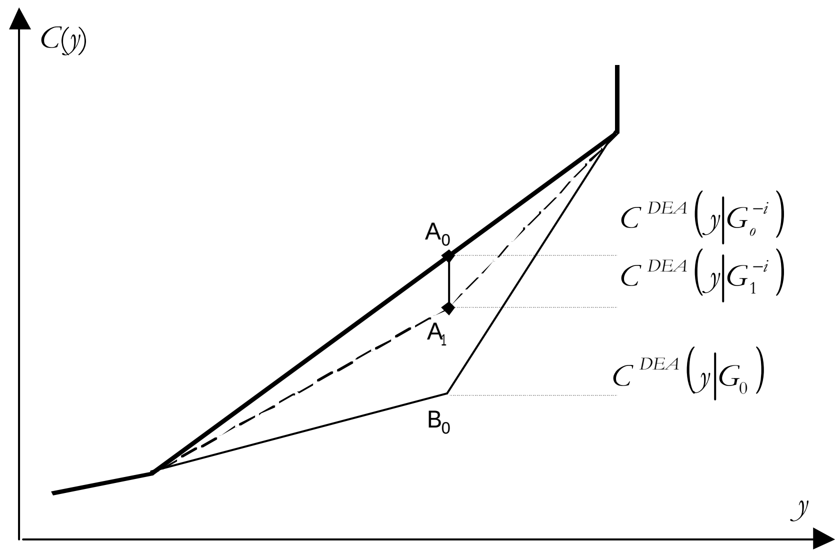

FIGURE 5. The DEA approximation to the cost curve for the single output case with B as the innovator and A subject to knowledge transfer.

4.4. **Summary and Extensions.** In this section the proposed incentive system is summarized and an extension to the multi-period case is discussed. By carefully looking at three important processes of the enterprise's adaptation to changes in the economic environment: the learning process (frontier catch-up), the innovation process (breaking the frontier) and the dissemination process (induced frontier catch-up), the following incentive structure is proposed for unit i in the two-period horizon:

$$b_0(x_0) = x_0 + \rho[C^{DEA}(y_0|G_0^{-i}) - x_0] + b_0^u$$

where the innovation premium b_0^u is awarded for an announced innovation leading to $x_1 = x_0 - \overline{u}$ at a cost $\gamma \leq 2$ according to the payment scheme

$$b_0^u = \left\{ \begin{array}{cc} \gamma\overline{u}, & \frac{1}{2}\gamma \leq \rho < 1 \\ \gamma\overline{u}, & 0 \leq \rho \leq \frac{1}{2}\gamma \end{array} \right\}$$

The second period incentive scheme becomes

$$b_1(x_1) = x_1 + \rho[C^{DEA}(y_1|G_1^{-i}) - x_1] + b_1^u + b_1^d$$

where the second-period innovation premium b_1^u is calculated as

$$b_1^u = \left\{ \begin{array}{cc} 2\rho\overline{u}, & \frac{1}{2}\gamma \leq \rho < 1 \\ 0, & 0 \leq \rho \leq \frac{1}{2}\gamma \end{array} \right\}$$

and the dissemination premium b_1^d in case of an observed knowledge transfer is given as

$$b_1^d = \rho\left(C^{DEA}\left(y_1|G_0^{-i}\right) - C^{DEA}\left(y_1|G_1^{-i}\right)\right)$$

The incentive scheme is operational, i.e., all entities except for the parameters ρ and γ may be calculated from available input-output data. The two parameters for innovation cost and managerial control may be estimated using managerial discretion or by econometric techniques from historic data. Under any circumstances, any incentive system has to be adjusted to the organizational context where it is implemented, including managerial culture, intensity of control, market situation and administrative resources.

A straight forward extension to the multi-period case is possible, using a formulation like

$$b_t(x_t) = x_t + \rho[C^{DEA}(y_t|G_t^{-i}) - x_t] + b_t^u + b_t^d(d_t)$$

where $u_t = \overline{u}$ is a period t of invention and the binary flag $d_t = 1$ iff there is dissemination in period t and $u_{t-1} > 0$, $d_t = 0$ else. By adjusting the initial data points in the reference set G_t provisions could also be made for rolling production planning and data obsolescence.

5. Conclusions

Rather than *penalizing* the information rents accrued by asymmetric information, like in the regulation models by Bogetoft (1997), Freixas et al. (1985) and others, the incentive system in private industry is established to *promote* information sharing and innovation. The subtle theoretical difference is a vague mirror image of the conceptual abyss that lies between the workings of public sector cost-orientation and the profit-orientation of private enterprises. In the latter, it is perhaps less the worry about prevailing inefficiency than the apparent suboptimality that urges the development of sophisticated incentive systems. This paper bridges the two worlds and shows how to correct common problems even for a complex decentralized production system

The optimality of the proposed system relies on the non-cooperative character of the agents' behavior, a common time-horizon and risk preference. However, these assumption are central to any comparative or yardstick incentive system and deserve more attention in a practical application.

An interesting extension to the incentive system is to include specific targets for training, for instance by focusing at local segments of the frontier. Such development could also contribute to the development of an advanced managerial compensation scheme. Building on the thoughts in the paper, it suggests that a manager raising the efficiency of horizontally placed units in the organizational chart should have a positive incentive for doing so. Such practice is yet not common, even in organizations with advanced profit sharing systems. Further extensions may also be made on the emphasis put on the three developmental phases at different hierarchical or educational levels. One may outline a system where a newcomer initially is induced to catch-up with existing best-practice (the frontier), then concentrating his efforts to break the norm (innovation), eventually to become an organizational mentor.

References

[1] Agrell, P. J., P. Bogetoft and J. Tind, *Dynamic Yardstick Regulation*. Working Paper 2000/4, Royal Agricultural University, Department of Economics and Natural Resources, Denmark, 1999.

[2] Agrell, P. J. and B. M. West, A Caveat on the Measurement of Productive Efficiency. *International Journal of Production Economics,* to appear, 2000.

[3] Banker, R.D., A. Charnes and W.W. Cooper, Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30, pp. 1078-1092, 1984.

[4] Bogetoft, P., Incentive Efficient Production Frontiers: An Agency Perspective on DEA. *Management Science*, 40, pp.959-968, 1994.

[5] Bogetoft, P., Incentives and Productivity Measurements. *International Journal of Production Economics*, 39, pp. 67-81, 1995.

[6] Bogetoft, P., DEA-Based Yardstick Competition: The Optimality of Best Practice Regulation. *Annals of Operations Research*, 73, pp. 277-298, 1997.

[7] Bogetoft, P., DEA and Activity Planning under Asymmetric Information. *Journal of Productivity Analysis*,13, pp. 7-48, 2000.

[8] Charnes, A., W.W. Cooper and E. Rhodes, Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2, pp. 429-444, 1978.

[9] Freixas, X, R. Guesnerie and J. Tirole, Planning under Asymmetric Information and the Ratchet Effect. *Review of Economic Studies*, LII, pp. 173-191, 1985.

[10] Hart, O.D. and B. Holmström, The Theory of Contracts, pp. 71-155 in T.F.Bewley, "Economic Theory — Fifth World Congress", Cambridge University Press, 1987.

[11] Laffont, J.J., and J. Tirole, The Dynamics of Incentive Contracts. *Econometrica*, 56, pp. 614-641, 1986.

[12] Shleifer, A., A Theory of Yardstick Competition. *Rand Journal of Economics* 16, pp. 319-327, 1985.

[13] Weitzman, M. The "Ratchet Principle" and Performance Incentives, *Bell Journal of Economics*, (Spring), pp. 302-308, 1980

(Per J. Agrell and Peter Bogetoft) Department of Economics, Royal Agricultural University, DK-1958 Frederiksberg C, DENMARK

E-mail address: per.agrell@kvl.dk

URL: http://www.flec.kvl.dk/pag

E-mail address: peter.bogetoft@kvl.dk

URL: http://www.flec.kvl.dk/pb

(Jørgen Tind) Department of Operations Research, University of Copenhagen, DK-2100 Copenhagen, DENMARK

E-mail address: tind@math.ku.dk

URL: http://www.math.ku.dk/or/tind