CES Working Paper Series

INTERTEMPORAL COMMON AGENCY AND ORGANIZATIONAL DESIGN: HOW MUCH DECENTRALIZATION?

Trond E. Olsen
Gaute Torsvik

Working Paper No. 37

Center for Economic Studies
University of Munich
Schackstr. 4
8000 Munich 22
Germany
Telephone: 089-2180-2748
Telefax: 089-397303

This work was completed while Torsvik was a visitor at the *Center for Economic Studies (CES),* University of Munich. Financial support from the Norwegian Research Council (P) is gratefully acknowledged.

CES Working Paper No. 37
May 1993

INTERTEMPORAL COMMON AGENCY AND ORGANIZATIONAL DESIGN: HOW MUCH DECENTRALIZATION?

Abstract

In common agency, where one agent contracts with several principals, to what extent should the principals cooperate and centralize provision of incentives? If the agency is over contract complements - where an increase in activity for one principal increases the marginal value of contracting with other principals - complete centralization seems optimal, since this internalize externalities. This intuition is known to hold true in a static setting. We show that this conclusion does not generalize to a dynamic agency situation. For although centralization provides more accurate incentives to the agent, it also aggravates the ratchet effect, and in some cases this negative effect dominates. We discuss how the optimal degree of centralization - or delegation - varies with other important variables. Our findings are related to concrete policy questions both at the level of internal organization design and to questions more related to public regulation.

Trond E. Olsen
Gaute Torsvik
The Norwegian Center for Research in Organization and Management
Rosenbergsgate 39
N-5015 Bergen
Norway

1. Introduction.

Common agency is a situation where one agent contracts with several principals, and the principals non-cooperatively provide incentives for the agent (Bernheim and Winston, 1986). If the agent has private information, the principals face a problem of mechanism design under common agency. A specific type of common agency arises when the agency is over contract complements, which is the case when an increase in activity for principal i raises the marginal utility of contracting with principal j. Such externalities arise when the agent's technology possesses economies of scope or if there are other kinds of positive spillovers in production or consumption.

There are many interesting economic problems that fit into this framework. One example is an upstream production unit, be it a firm or a subdivision of a divisionalized company, which operates under economies of scope and which contracts to deliver products to several downstream firms or other subdivisions. Another is air pollution from a single firm, when the pollution affects residents of different states or nations. A third example is when a group of firms, producing goods that are complements in consumption, use the same research or marketing agency. In such cases an important design or policy question is to what extent the principals should cooperate and centralize the provision of the agent's incentives. Given the positive externalities that exist in production or consumption, centralization seems to be a good idea. This is indeed one of the main results we have about mechanism design under a common agency: In a static setting all parties are worse off under a common agency than they would be if the principals formed a joint agency. The principals are worse off because incentives are ineffIcently low powered, and the agent is worse off since low powered incentives generate small information rents (See Stole (1990), Martimort (1990) or Fudenberg and Tirole (1991)).

This result appears to have quite strong policy implications. For example, horizontal integration in a downstream industry buying from a single upstream producer should--under economies of scope--improve efficiency, and this could be a strong argument against anti-trust measures. It is therefore important to study the robustness of these policy conclusions. Particularly, since most economic relationships last over time, it is of considerable interest to examine whether they generalize to an intertemporal setting.

The intertemporal problem is interesting because decentralised provision of incentives--or delegation of contracting authority--may mitigate the ratchet effect. This effect arises when the agent is aware that any information he reveals about his productivity now, will be exploited by the principals in the future. This occurs when contracts are incomplete in the sense that full long-term commitments are infeasible. Unless high ability types then are given very generous incentives early in the relationship, they will mimic the behaviour of less efficient types in order to capture future information rents, see e.g., Freixas et al.(1985) and Laffont-Tirole (1988). A decentralised common agency over contract complements dilutes incentives and so, under this agency form there are less information rents for the agent to wait for. This can reduce the ratchet problem. Relative to centralized contracting, decentralized provision of incentives has therefore two counteracting effects in a dynamic setting; it induces static inefficiency but, on the other hand, it gives a dynamic gain by reducing the ratchet problem.

In an earlier paper (Olsen and Torsvik 1991), we showed by a numerical example that intertemporal welfare could improve if the number of principals was increased from one to two. There we studied a specific common agency problem associated with the privatization of a publicly regulated firm, see also Laffont and Tirole (1991). In this paper we study the intertemporal effects of a common agency in a more general framework, and we focus on a different question, namely what is the optimal degree of centralization. Static inefficiency increases with the number of principals, that is, with the degree of decentralization/delegation, but so do the dynamic gains. It is therefore interesting to study which of the two effects that dominate as the number of principals increases. We prove here that under certain conditions it will never be optimal to centralize the provision of incentives completely. We also (partly) characterize how the optimum degree of centralization depends on various important parameters in this principal-agent problem.

The dynamic problem we study in this paper at a relatively general level has implications for several important economic relationships. Some further examples will show the relevance and the nature of the problem we study. First, our analysis can have interesting implications for the design and internal organization of firms. To illustrate, consider a divisionalized firm where one division delivers services with positive spillovers to other divisions. For example, suppose there is a research lab providing general-purpose products (software) that can be used by all divisions. If the research lab has private information about the tasks it is to perform, it is optimal to centralize or coordinate the incentives provided to the lab in a static context. In a dynamic setting it may be optimal to delegate decision power and move to a structure where divisions act more like independent profit centers.[1]

[1]. There is a parallel between our argument in favour of decentralisation and the "influence cost" approach (Milgrom 1988, Milgrom and Roberts 1992). That approach emphasizes that a decentralised structure

We think our model also has something to say about ownership structure --about how centralized ownership should be. The professional management of a firm will typically know more about the firm's true productivity than the owners. Owners must then put the management on an incentive scheme to align the managers' incentives with their own. The optimal scheme trades off efficiency against rent extraction. If there are several parties that independently give the management incentives to produce, incentives will end up being inefficently low powered. In a static setting the firm's value is therefore maximized if there is one owner only. Due to the dynamic effects discussed above, this needs no longer be true in a multi-period setting. So, if the founders maximize the long run value of the firm it may be optimal to split the ownership into several parts.[2]

Another example is the regulation of pollution when emissions impose costs on several parties, and the management of the polluting source has private information about abatement costs. A recent case is the transborder pollution problem generated by Russian nickel factories on the Kola peninsula.[3] There has been some political discussion about a solution to the problem where the three neighbouring Nordic countries give the firm incentives to reduce emissions. If these countries operate non-

gives the centre control of relatively less quasi rents, which in turn gives the subunits of the organization less incentives to devote resources to unproductive activities that aim at influencing the distribution of these rents. In a standard principal-agent setting, the costs associated with pooling behaviour can be interpreted as a kind of influence costs. Our argument is that in an intertemporal setting, a decentralized incentive structure generates smaller information rents, and thererfore gives subunits of the organization less incentives to undertake (wasteful) activities for the sole purpose of hiding the unit's true productivity.

[2] The costs and benefits of ownership pointed at here are clearly quite different from those considered by Grossman-Hart(1986) and Hart-Moore(1988).

[3]. Similar issues related to multi-country pollution arise along the river Rhine, see Stroebele (1991). Baron (1985) gives an analysis of multi-principal pollution regulation in a static context.

cooperatively, too little incentives will be provided. So if a once-and-for-all contract can be written, all parties would be better off if the countries formed a joint agency. But as we have seen, the non-cooperative solution has a positive effect if contracts must be made repeatedly. This naturally leads to the following question: What is the optimal degree of cooperation? Is it for example possible that the principals are better off if two of the countries form a "joint venture" and operate as one principal, so that the agent is regulated by two rather than three principals? This will reduce the static inefficiency relative to the fully non-cooperative solution. But since incentives will be more powerful, it will also be more difficult to separate high ability types from low ability types. Which of these two effect--the positive static or the negative dynamic effect--dominates, depend on central parameters of the problem such as the importance of the future relative to the current period, how likely principals think it is that the firm is of high ability, etc.

Our setting is a simple model with two periods. We assume an extreme degree of contract complementarity where the marginal cost of providing other principals with an extra unit of the good after a unit is made available for one principal, is zero. This assumption simplifies our exposition considerably. The trade-off between the dynamic gain and the static loss of common agency exists, however, for all degrees of contract complementarities. Moreover, primarily for reasons of tractability, we restrict attention to linear incentive schemes.

The next section develops the model. Then we briefly illustrate the static case in Section 3 before we turn to the dynamic model in Section 4. Some numerical examples are presented in Section 5, and Section 6 concludes. Most proofs are relegated to appendices.

2. The model.

There are $n \geq 1$ principals and one agent. We focus on a case where there is extreme "contract complementarity"; the agent produces a non-rival good for the principals. Principal i has (constant) willingness-to-pay p^i for the good, and makes each period a transfer t^i to the agent. The agent produces quantity y at cost $\psi(y,\theta)$ and has low ($\theta = \underline{\theta}$) or high ($\theta = \bar{\theta}$) productivity. The agent's productivity is private information. Output but not cost is commonly observable. The cost functions are strictly convex, and the high productivity type has strictly lower total as well as marginal costs for every positive level of output.

The per-period utilities of the agent and of principal i are, respectively;

$$t^1+\ldots+t^n - \psi(y,\theta) \quad \text{and} \quad p^i y - t^i.$$

Transfers are restricted to be linear functions of the observable y, thus

$$t^i = a^i+b^i y.$$

The principals act non-cooperatively. In a cooperative solution they would choose $(t^1,\ldots,t^n)$ jointly to maximize $\Sigma(p^i y-t^i) = Py-a-by$, where $P=\Sigma p^i$, $a=\Sigma a^i$ and $b=\Sigma b^i$. Our goal is to compare the relative merits of cooperative and non-cooperative solutions as the number of principals (n) varies. We shall therefore assume that the total willingness-to-pay P is independent of n, thus

$$p^1+\ldots+p^n = P = \text{constant, all } n \geq 1.^4$$

---

4 A second interpretation of the model is that the agent is the manager of a firm owned by n principals. The firm produces output y which is sold at a constant (e.g. competitive) price P. Only the manager knows the parameter θ in the cost function. Each owner charges a "rental" $T^i(y)=A^i+B^i y$ from the manager, whose per-period utility is thus $Py-\psi(y,\theta)-A-By$, where $A=\Sigma_i A^i$, $B=\Sigma_i B^i$. The owners' aggregate surplus, and hence the (per-period) value of the firm is then A+By. The issue we address is then how the (intertemporal) value of the firm depends on the number of owners, and further what is the optimal number of owners. By setting $a^i=-A^i$ and $b^i=P/n \cdot B^i$ this model becomes formally equivalent to the public goods model in the text.

The agent's reservation utility is assumed to be independent of type, and is normalized to zero. Provided he participates, he chooses quantity $y=y(b,\theta)$ such that the marginal cost equals the "bonus" b;

$$b=\psi'(y,\theta), \qquad b=\Sigma b^i.$$

(In order to avvoid corner solutions, we assume $\psi'(0,\theta)=0$ and $\psi'(\infty,\theta)=\infty$. We also assume $\psi''(y,\theta)$ bounded away from zero.) The quantities corresponding to low and high ability will be denoted $\underline{y}(b)$ and $\bar{y}(b)$, respectively. The IR-constraint for type θ is

$$a+by-\psi(y,\theta)\geq 0, \quad a=\Sigma a^i.$$

As a reference case, consider the <u>full information</u> solutions. The cooperative or single-agency (n=1) solution is obviously to set b=P in order to generate efficient production $y^*(\theta)=y(P,\theta)$, and then extract the full surplus by setting $a=\psi(y^*,\theta)-Py^*$. It is easily seen that the aggregate bonus under common agency (n>1) coincides with the efficient single-agency bonus P.[5] The individual fixed fees a^i are not uniquely determined, but the aggregate fee is unique, and coincides with the optimal single-agency fee. Under full information the resource allocation is therefore efficient and <u>independent of the number of principals</u>. For future reference, let $\bar{w}^*$ (respectively $\underline{w}^*$) denote the total surplus accruing to the principals when the agent has high (respectively low) ability, thus

$$\bar{w}^* = P\bar{y}^*-\bar{\psi}(\bar{y}^*) \text{ and } \underline{w}^* = P\underline{y}^*-\underline{\psi}(\underline{y}^*).$$

3. Static case.

As a prerequisite to the analysis of dynamic common agency, we consider in this section the static (one-period) case. We assume that the agent

---

[5] This holds provided all principals participate, which is a maintained assumption, see Section 3. Then principal i is (by the IR-constraint) led to choose b^i so as to maximize $p^iy - [\psi(y,\theta)-b^{-i}y-a^{-i}]$, where $b^{-i}=\Sigma_{j\neq i}b^j$, $a^{-i}=\Sigma_{j\neq i}a^j$, and $\psi'(y,\theta)=b^i+b^{-i}$. This yields $b^i=p^i$.

contracts with all principals, and that the bad type is never excluded.[6]

This type's IR-constraint will then be binding, thus

$$0 = a+b\underline{y}(b)-\underline{\psi}(\underline{y}(b)).$$

The good type will therefore collect a rent. For a given bonus b, this rent is given by

$$R(b) = b[\bar{y}(b)-\underline{y}(b)] - [\bar{\psi}(\bar{y}(b)) - \underline{\psi}(\underline{y}(b))] \tag{3.1}$$

Note that R(b) is increasing in the bonus b; the marginal information rent $R'(b) = \bar{y}-\underline{y}$ is clearly positive.

Principal i chooses (a^i,b^i) to maximize her expected surplus

$$\nu(p^i-b^i)\underline{y}(b) + (1-\nu)(p^i-b^i)\bar{y}(b)-a^i$$

subject to the $IR(\underline{\theta})$ constraint, where ν is the probability that the agent has low ability, $a=a^i+a^{-i}$ and $b=b^i+b^{-i}$. Taking account of the "IC-constraint" $(\psi'(y,\theta)=b)$ this maximization yields

$$0=(p^i-b^i)[\nu\underline{y}'+(1-\nu)\bar{y}'] - (1-\nu)[\bar{y}-\underline{y}],$$

where primes denote derivatives. The first term on the RHS of this equation is the marginal value for principal i generated by the agent's production response to an increase in the bonus b^i. The second term is the marginal cost to principal i of increased rents to the (good type of) agent. Summing over all i, we see that the equilibrium aggregate bonus $b^1+\ldots+b^n=b_n$ depends on n and is given by

---

[6] These assumptions are invoked in order to keep our arguments simple. No principal would in fact find it advantageous to offer contracts which exclude the bad type if the probability for this type is sufficiently large, or if there are large costs associated with zero output. The contract complementarity makes it advantageous for the agent to contract with all rather than with only a subset of the principals. If the good is excludable (such as computer software) each principal also has an interest in contracting with the agent. If the good is non-excludable, additional considerations of free-riding would have to be taken into account if participation is not mandatory for the principals.

$$(3.2) \qquad P-b_n = \frac{n(1-\nu)[\bar{y}(b_n)-\underline{y}(b_n)]}{\nu\underline{y}'(b_n) + (1-\nu)\bar{y}'(b_n)}$$

The aggregate fixed fee is then uniquely given by

$$(3.3) \qquad 0 = a+\underline{w}(b_n), \quad \underline{w}(b) \equiv b\underline{y}(b)-\psi(\underline{y}(b)),$$

where $\underline{w}(b)$ is the surplus produced by the bad type for a given bonus b. The individual bonuses b^i are uniquely determined, but the individual fixed fees are not. Any combination $(a^1,\dots,a^n)$ whose sum equals the aggregate fee given above will constitute (part of) an equilibrium.

The equilibrium bonus is distorted relative to the efficient bonus ($b_n<P$), and--given some regularity conditions--the degree of distortion increases with the number of principals ($b_{n+1}<b_n$). This occurs because of "multiple rent extraction" (Fudenberg and Tirole, 1991): Consider, for a given aggregate bonus b, the aggregate surplus accruing to the group of n principals; let this be denoted w(b), thus

$$w(b) = (P-b)[\nu\underline{y}(b)+(1-\nu)\bar{y}(b)] + \underline{w}(b).$$

The marginal "social" cost -- i.e. the marginal cost for the group, acting collectively -- in terms of increased information rents induced by an increase in the bonus b, is $(1-\nu)[\bar{y}-\underline{y}]$. In the non-cooperative solution, every principal takes this cost into account as a private cost. The marginal cost of increasing the agent's incentive bonus is thus exaggerated, and more so the more principals there are. Indeed, the equilibrium condition is easily seen to be of the form

$$0 = dw(b)/db - (n-1)(1-\nu)[\bar{y}(b)-\underline{y}(b)]$$

If w(b) is concave, it is thus certainly the case that the higher the number of principals, the lower is the equilibrium bonus.

The concavity condition on w(b) also ensures that an equilibrium exists for any $n \geq 1$, and that the equilibrium bonuses are unique. We shall therefore invoke this assumption in the following.[7]

The equilibrium bonus b_n depends on the probability ν that the agent has low ability, and so does of course the equilibrium surplus accruing to the principals. The total equilibrium surplus accruing to this group is

$$(3.4) \qquad w_n(\nu) = (P-b_n(\nu))[\nu \underline{y}(b_n(\nu)) + (1-\nu)\bar{y}(b_n(\nu))] + \underline{w}(b_n(\nu))$$

Letting $\nu \to 1$ for fixed n, we see that $b_n \to P$ and $w_n \to \underline{w}^*$.

While the aggregate equilibrium surplus is uniquely determined, its division between the principals is not unique. For every feasible combination of individual fixed fees $(a^1, \ldots a^n)$ there will be some division of the surplus. Assuming some equilibrium is selected, the surpluses $w_n^i(\nu)$ accruing to each principal i, $i=1,\ldots,n$ are well defined, and depend on n and ν. We shall assume that some such selection takes place, e.g., via focal points. For example, in the symmetric case $p^i = p$ for all i the individual bonuses are necessarily equal, and then the natural focal point is to set all fixed fees equal as well.

The agent's information rent is according to (3.1) given by

$$(3.5) \qquad \pi_n(\nu) = R(b_n(\nu)).$$

It follows from the properties of the equilibrium bonus $b_n(\nu)$ that the rent is strictly increasing in the probability ν, and strictly decreasing in the number of principals n. As $\nu \to 1$, the rent converges to R(P).

4. The dynamic case.

In this section we consider a two-period model of common agency, where no long-term commitments are feasible. We first consider the continuation equilibrium ensuing after an arbitrary first-period (aggregate) bonus. That

[7] One can easily check that this assumption is satisfied e.g. for a quadratic cost function.

part is just a simple adaption of the analysis given in Freixas et al. (1985)--hereafter FGT. Then we go on to characterize full intertemporal equilibria for the n-principal case, and finally to a comparison of these equilibria as n varies.

4.1 Continuation equilibria.[8]

In this subsection we consider the continuation equilibrium corresponding to a first-period bonus b. There are three types of such continuation equilibria (FGT, Proposition 4, p.181): pooling, semi-separating and fully separating. In all cases the low-productivity type produces the optimal static quantity $\underline{y}(b)$. In a pooling equilibrium the high-productivity type also produces $\underline{y}(b)$, while in a fully separating (FS) equilibrium he produces his one-period optimal output $\bar{y}(b)$. In a semi-separating (SS) equilibrium the high-productivity type randomizes between $\underline{y}(b)$ and $\bar{y}(b)$.

By pooling in the first period the high-productivity type looses some profits $\Delta(b)$ -- relative to full revelation. This loss is given by

(4.1) $$\Delta(b) = b[\bar{y}(b)-\underline{y}(b)] - [\bar{\psi}(\bar{y}(b))-\bar{\psi}(\underline{y}(b))]$$

In exchange the high-productivity type gains second-period rents $\delta\pi_n(\nu_n)$, where δ is the discount factor and ν_n is the updated probability that the firm has low productivity. There are three cases: 1) $\Delta(b)\leq\delta\pi_n(\nu)$; pooling, 2) $\delta\pi_n(\nu)<\Delta(b)<\delta\pi_n(1)$; semi-separation, 3) $\delta\pi_n(1)\leq\Delta(b)$; full separation.

In case 3, the first-period loss associated with pooling exceeds the maximal second-period gain. In case 2 there is some randomization probability x_n (of choosing $\bar{y}(b)$) such that gains and losses balance, i.e.,

---

8. The analysis in this subsection is only slightly more general than our previous one (Olsen and Torsvik, 1991). Note that our linearity restriction eliminates equilibria where also the bad type pools, see Laffont and Tirole (1988).

(4.2) $\Delta(b)=\delta\pi_n(\nu(x_n))$, $\nu(x_n)=\nu/[\nu+(1-x_n)(1-\nu)]$

Here $\nu(x_n)$ is the updated probability that the agent is a bad type, given that he has produced $\underset{\sim}{y}(b)$ in the first period. In case 1 any degree of randomization will be non-profitable for the firm.

The first-period loss for the high-productivity type associated with producing $\underset{\sim}{y}(b)$ is independent of the number of principals n. The second-period rents accruing to the high-productivity type $\delta\pi_n(\nu)$ do however depend on this number. These rents are determined by the respective static (second-period) equilibria. Since in a static context incentives are less powerful the larger is n--at least for $\nu<1$--we have $\pi_{n+1}(\nu) < \pi_n(\nu)$ for $\nu<1$, and $\pi_{n+1}(1) = \pi_n(1)$. These observations immediately yield the following result:

Lemma 1: Let P_n, SS_n and FS_n denote the sets of first-period bonuses which yield, respectively, pooling, semi-separating and fully separating continuation equilibrium for n principals, $n\geq1$. Then $P_{n+1}\subset P_n$, $SS_n\subset SS_{n+1}$, and $FS_n=FS_{n+1}$. Moreover, for a first-period bonus b in the interior of SS_n, the revelation probabilities $x_n(b)$ and $x_{n+1}(b)$ satisfy $x_n(b) < x_{n+1}(b)$. ∎

The last assertion follows directly from the relations defining these probabilities, namely $\Delta(b)=\delta\pi_n(\nu(x_n))$, $n\geq1$, and the fact that $\pi_n()$ and $\nu()$ are strictly increasing functions. The regions for two agency settings are illustrated for the "well-behaved" case -- i.e., where $\Delta(b)$ is strictly increasing -- in Figure 1 below.

(Figure 1 here.)

These considerations show that for a given first-period incentive scheme there is "more separation" (or revelation) under common agency than under single agency. The reason is of course that in the former case the agent has relatively less to loose by revealing himself early.

<u>Example.</u> Consider the following example, adapted from FGT. The cost function is quadratic;

$$\psi(y,\theta)=y^2/2\theta, \quad \theta\in(\alpha,1),$$

where $\theta=1$ for the high-productivity type, and $\theta=\alpha<1$ for the low-productivity type. For this model we have

$$y(b,\theta) = \theta b$$

$$b_n(\nu) = P[\nu\alpha+1-\nu]/[\nu\alpha+(1-\nu)K_n], \quad K_n = 1+n(1-\alpha).$$

The SS-region is the interval (b_n^*, b^{**}), where

$$b_n^* = \mu b_n(\nu), \quad b^{**} = \mu P, \quad \mu = [\delta/(1-\alpha)]^{1/2}$$

In this region the randomization probability is given by

$$x_n(b) = 1-\alpha[\nu/(1-\nu)][\mu P-b]/[bK_n-\mu P]$$

It is known that the <u>optimal</u> first-period scheme under single agency may induce the firm's types to pool, semi-separate or fully separate. (FGT, Proposition 8, p.188). We will here be particularly interested in the semi-separating case.

Before we proceed to the formal model of the intertemporal common agency equilibrium, it will be helpful to give an intuitive description of what is going on as the number of principals increases. We focus on a first-period bonus that induces semi-separation for n principals, and assume first that this bonus remains fixed as we increase the number of principals from n to n+1. The bonus will in fact not stay fixed in equilibrium, but we start out with this assumption in order to separate the various dynamic effects of decentralization. The changes in output and thus in the principals' surplus are then generated solely by the increase in the revelation probability; $x_n(b) < x_{n+1}(b)$.

A higher randomization probability has an impact on the surplus both in the first and in the second period. In the first period there is a higher probability that the low cost type produces a high output. This may in fact be disadvantageous for the principals; more separation will lower their surplus if the first period bonus is "large". The point is that an increase in production is not desirable from the principals' point of view if the bonus is higher than the marginal value of output (P). And there are parameter values such that a very high first period bonus must be given to induce separation. On the other hand, more revelation increases the first-period surplus if the bonus that induces semi-separation is lower than the marginal output value P.

The second-period impacts of a higher randomization probability are more involved. For in the second period there are two "states"; one where the low cost firm has revealed its identity and one where it has not. More revelation increases the probability of the first state, and this increases the second-period surplus. In the other state, where the low cost firm produced as if it had high costs in the first period, there are two forces at work when we increase the number of principals. One effect is the dilution of the second period bonus that follows from decentralization ($b_n(\nu)$ decreases as n increases). The other effect is that decentralization increases the revelation probability, which in turn increases the updated probability that the firm is of the high cost type, and this pushes the second period equilibrium bonus upwards ($b_n(\nu)$ increases in ν). It turns out that the last two effects exacly balance in equilibrium. So the net effect on the second period-surplus is therefore positive.

Summing up we see that for a fixed semi-separating first-period bonus the intertemporal surplus effect of increasing the number of principals consists of a first-period effect that can be either positive or negative, and a second-period effect that is positive. But, of course, to get the total effect of decentralization we must take into account the negative effect on the first-period incentives as well. In the next section these effects are studied more carefully, and we derive propositions about their relative magnitudes and thus about the total effects of decentralization.

4.2 Intertemporal common agency equilibrium.

Consider principal i, and suppose the aggregate first-period bonus $b=b^i+b^{-i}$ is in the semi-separating region ($b \in SS_n$). The surplus accruing to this principal in the first-period is then

$$(p^i-b^i)\{[\nu+(1-\nu)(1-x_n)]\underline{y}(b)+x_n(1-\nu)\bar{y}(b)\}-a^i,$$

where a^i, b^i must satisfy the IR-constraint for the low-ability type ($0 = a^i+a^{-i}+b\underline{y}-\psi(\underline{y})$), and $x_n=x_n(b)$ is the probability that the high-ability type reveals himself. The first-period effect of a marginal increase in b^i, taking the IR and "IC" constraints into account, is then given by

$$(4.3) \qquad (p^i-b^i)\{[\nu+(1-\nu)(1-x_n)]\underline{y}'+(1-\nu)x_n\bar{y}'\}$$
$$-(1-\nu)x_n[\bar{y}-\underline{y}] + (p^i-b^i)(1-\nu)[\bar{y}-\underline{y}](dx_n/db)$$

These three terms capture, respectively: (i) The direct effect associated with increased production; this effect is positive if and only if $b^i<p^i$. (ii) The direct negative effect induced by increased information rents; note that the agent's rent will increase only if the agent is good and reveals himself. (iii) The indirect effect induced by increased revelation; this effect is also positive if and only if $b^i<p^i$.

The second-period expected surplus accruing to principal i is given by

$$(4.4) \qquad [\nu+(1-\nu)(1-x_n)]w_n^i(\nu_n)+(1-\nu)x_n\bar{w}^{*i},$$

where $\nu_n=\nu(x_n)$ is the updated probability that the agent has low ability, $x_n=x_n(b)$ is the revelation probability, and $\bar{w}^{*i}$ and $w_n^i(\nu_n)$ are the second-period payoffs to principal i when the high-ability type reveals, respectively does not reveal himself in period 1. Clearly the second-period expected surplus depends on the first-period bonus b only via the revelation probability $x_n(b)$.

Let W(n,b) denote the aggregate intertemporal surplus accruing to the group of principals when the first-period aggregate bonus b is in SS_n, thus

$$(4.5)\qquad W(n,b)=W_1(n,b)+\delta W_2(n,b),$$

where

$$(4.6)\qquad W_1(n,b) = (P-b)([\nu+(1-\nu)(1-x_n(b))]\underline{y}(b) + (1-\nu)x_n(b)\bar{y}(b))+\underline{w}(b)$$

$$(4.7)\qquad W_2(n,b) = [\nu+(1-\nu)(1-x_n(b))]w_n(\nu(x_n(b)))+(1-\nu)x_n(b)\bar{w}^*$$

Differentiating principal i's second-period surplus (4.4) with respect to b^i, combining this with the marginal first-period surplus (4.3), and summing over all i=1...n, we obtain the following equilibrium condition for the aggregate first-period bonus b -- provided this bonus belongs to the semi-separating region:

$$(4.8)\qquad 0 = \frac{\partial W}{\partial b}(n,b)-(n-1)(1-\nu)x_n(b)[\bar{y}(b)-\underline{y}(b)]$$

The last term in this equation captures the first-period "common-agency effect". This effect is of course absent under single agency (n=1). It would also be absent if the group of principals acted cooperatively in the first period and non-cooperatively in the second period, i.e. if we had a situation where single agency was followed by common agency. In the latter case the optimal first-period bonus would satisfy $\partial W(n,b)/\partial b=0$. The first-period common-agency effect tends to lower the equilibrium bonus relative to this optimal bonus -- provided W(n,b) is concave in b.

Consider next the <u>pooling region</u>. For a first-period aggregate bonus in this region, both types produce $\underline{y}(b)$, and no information is revealed. Formally this corresponds to setting $x_n(b) = 0$ on the analysis above, and it is then easily seen that any equilibrium in this region must have a first-period bonus given by b = P.

Finally, in the fully separating region the good type produces $\bar{y}(b)$ with certainty. The second-period payoffs are in this region independent of the first-period incentive schemes, hence it is clear that an equilibrium first-period bonus in the FS-region must coincide with the static equilibrium bonus $b_n(\nu)$. We summarize this discussion in the following proposition.

Proposition 1: For $n \geq 1$, if the equilibrium first-period aggregate bonus $b = b(n)$ belongs to the semi-separating region (SS_n), then it is a solution of (4.8). If the equilibrium first-period bonus induces pooling (respectively full separation), then the bonus is $b = P$ (respectively $b = b_n(\nu)$). ▪

It is relatively straightforward to give sufficient conditions for existence of an n-principal SS-equilibrium. Some such conditions are given in Appendix A.

4.3 Comparisons

In this subsection we shall compare equilibria as the number of principals varies. It has already been shown that for a fixed first-period bonus, the degree of first-period revelation increases as the number of principals increases (Lemma 1). Unless the given bonus is rather large (i.e. unless $b > P$), the higher degree of separation will be shown to be beneficial for the principals; this is the positive dynamic effect associated with having more principals. But more principals also dilutes the power of first-period incentives, and thus gives rise to a negative static effect. The present subsection aims at analysing this trade-off.

Consider the intertemporal equilibrium, and in particular the surplus accruing to the group of principals. For a semi-separating first-period equilibrium bonus $b(n)$, this surplus can be written as $W(n,b(n))$, where $W(n,b)$ is given by (4.5). Variations in the number of principals n will affect the surplus "directly" via the argument n, and "indirectly" via the equilibrium bonus $b(n)$. In order to analyze the two effects jointly, it is helpful to first treat n as a continuous variable. The functions $W(\cdot)$ and $b(\cdot)$ are in fact well defined for any positive real number n, so in a formal sense it is acceptable to let n vary continuously. Doing so, we find that the equilibrium surplus function $W(n,b(n))$ generally has a positive derivative for n=1, implying that the surplus accruing to the principals is maximal for some (real) number n strictly exceeding one. To see this, differentiate the surplus function $W(n,b(n))$ to obtain

$$\frac{dW}{dn}(n,b(n)) = \frac{\partial W}{\partial n}(n,b(n)) + \frac{\partial W}{\partial b}(n,b(n))\,\frac{db}{dn}(n)$$

For n=1, the bonus $b(1)$ maximizes $W(1,b)$, hence the second term on the RHS vanishes. We will show below (Lemma 2) that $W(n,b)$ depends on n only via the revelation probability x_n, so we may write $W(n,b) = F(x_n(b),b)$. Moreover, from Lemma 1 we know that $x_n(b)$ is strictly increasing in n, so $\partial x_n/\partial n>0$. It then follows that $\partial W/\partial n = (\partial F/\partial x)(\partial x_n/\partial n)$ is positive for $n=1, b=b(1)$ <u>if</u> the intertemporal surplus under single agency is locally strictly increasing in the revelation probability [$\partial F/\partial x>0$ for $n=1$, $b=b(1)$]. We show below that this is indeed the case, and hence we have the following result.

Proposition 2: Suppose the equilibrium first-period bonus under single agency (n=1) is in the semi-separating region. Then there is some (real) number $n>1$ such that the principals' intertemporal equilibrium surplus $W(n,b(n))$ exceeds the surplus for the single principal:

$W(n,b(n)) > W(1,b(1))$, some $n>1$. ▪

To complete the proof it must first be shown that the intertemporal surplus under single agency is strictly increasing in the revelation probability (x_1). For n=1 it is well known that the principal's <u>second-period</u> surplus is strictly increasing in this probability (FGT, Lemma 3, p. 184). It is clear, e.g. from (4.6), that if (and only if) the first-period bonus is no larger than P, then the <u>first-period</u> surplus is also increasing in this probability. (It pays to have more good types produce $\bar{y}(b)$ rather than $y(b)$ if and only if $b<P$.) Hence we know that the intertemporal surplus inreases with x when $b \leq P$. So suppose the optimal bonus b(1) exceeds P. For this bonus we have $(\partial F/\partial x)(dx_1/db) + (\partial F/\partial b) = 0$. One can easily check that $\partial F/\partial b<0$ for $b>P$. Since $dx_1/db>0$, we therefore must have $\partial F/\partial x>0$ at $b=b(1)$, $x=x_1(b(1))$, which was the desired conclusion.

The final argument to complete the proof of Proposition 2 is provided by the following lemma.

Lemma 2. For a first-period aggregate bonus b in SS_k, $k \geq 1$, the equilibrium second-period bonus is the same for any number of principals n exceeding k. For such a first-period bonus b, the intertemporal surplus accruing to the principals [W(n,b)] depends on n only via the revelation probability $x_n(b)$, i.e. we may write

$$W(n,b) = F(x_n(b),b), \quad n \geq k, \quad b \in SS_k.$$

Proof: If b is in SS_k, then b is in SS_n for all $n \geq k$. For any such n, the good type will randomize with a probability x_n such that $\Delta(b)=\delta\pi_n(\nu(x_n))$, where $\nu(x_n)=\nu/[\nu+(1-\nu)(1-x_n)]$ is the aposteriori probability that the agent is bad, and $\pi_n(\ )$ is the second-period rent. But this rent depends only on the second-period equilibrium bonus; $\pi_n(\nu(x_n))=R(b_n(\nu(x_n)))$, see (3.5). Hence the equilibrium second-period bonus b_n is determined by $\Delta(b)=\delta R(b_n)$, and is therefore independent of n. The second statement in the lemma then follows directly from the first. ■

Remark. The observation in Lemma 2 that, for a fixed first-period bonus, the second-period equilibrium bonus is independent of n, is simple but quite important. One interesting implication is that <u>the principals' second-period equilibrium surplus increases as n increases</u> -- provided the first-period bonus is fixed and induces semi-separation. For, as n increases, the revelation probability $x_n(b)$ increases while the second-period bonus remains unaffected, and this in turn increases the expected second-period surplus accruing to the principals. This is so because for every good type that separates, second-period production efficiency is enhanced, and the principals save rent.[9] ■

The proof of Proposition 2 above exploits the fact that to a first-order approximation, the "indirect" effect associated with an increase in the number (n) of principals is negligible for n=1 [$\partial W/\partial b=0$ at $n=1$, $b=b(1)$]. For n>1 this is not so. We know that there is then a (first-period) common-agency effect which inflates the principals' perceived bonus costs, leading to an equilibrium first-period bonus which is "too small" in the sense that the principals' aggregate surplus is increasing in b at the equilibrium point, i.e. $\partial W/\partial b>0$ at $b=b(n)$, $n>1$, see (4.8). For n>1 the "indirect" effect is thus negative, at least in well-behaved cases. This negative effect, which arises because of the first-period static common-agency effect, then works against the positive "direct" effect caused by the agent's dynamic response to common agency. At an interior optimum, the two effects must of course balance.

In order to examine whether there is in fact an interior optimum, we consider now the asymptotic behaviour of the principals' aggregate equilibrium surplus W(n,b(n)) as n increases without bounds. We then find the following. First, for a fixed bonus in the first period, the dynamic common-agency effect induces in the limit full separation in that period. However, the static common-agency effect drives the first-period bonus b(n)

---

[9] One may further note that, since the second-period bonus B(b) is independent of n, so is the agent's second-period rent. For a fixed first-period bonus b in SS_k, all parties' second-period payoffs therefore (weakly) increase as n increases. Note also that for a fixed b in SS_k, the agent's intertemporal rent is fixed, and is in particular independent of the number of principals.

asymptotically to zero, such that in the limit no production is carried out in the first period. For a similar reason, there is in the limit neither any production in the second period if the agent has not been revealed as a high-ability type. Hence there is asymptotically only production if the agent has high ability, and then only in the second period -- after the agent's type has been revealed. The principals' aggregate intertemporal surplus therefore approaches the value $\delta(1-\nu)\bar{w}^*$ as the number of principals goes to infinity. If $w_1(\nu)\geq(1-\nu)\bar{w}^*$, a single principal can obtain a payoff exceeding $\delta(1-\nu)\bar{w}^*$ by offering a pooling contract in period one. Hence there must then be some finite number of principals which maximizes the aggregate intertemporal surplus.[10]

Proposition 3: For the "well-behaved case" the following holds. The equilibrium first-period bonus $b(n)\to 0$, and the principals' surplus $W(n,b(n))$ satisfies $W(n,b(n))\to\delta(1-\nu)\bar{w}^*$ as $n\to\infty$. Hence, if $w_1(\nu)\geq(1-\nu)\bar{w}^*$, the principals' aggregate surplus is maximal for some finite n. ▪

The proof is given in Appendix B. It follows from this result and Proposition 2 that there is some finite real number $n^*>1$ which maximizes the surplus function $W(n,b(n))$. Restricting n to its natural domain of positive integers, we can for the parametric example given in Section 4.1 prove the following:

---

[10] Note that $w_1(\nu)\geq(1-\nu)\bar{w}^*$ is precisely the condition which makes it non-optimal to exclude the bad type in static single agency. Our maintained assumption is that such exclusion is not allowed. If it were allowed, the conclusion in Propositon 3 would still hold if we invoked the stronger assumption $\underline{w}^*>(1-\nu)\bar{w}^*$. To see this, note that for n sufficiently large the bad type would certainly be excluded in period 2, and therefore also in period 1, so the principals' payoff would be $(1+\delta)(1-\nu)\bar{w}^*$. Since $w_1(\nu)\geq \underline{w}^*$, a single principal would not exclude in period 2. Suppose she neither excludes in period 1, and offers a bonus b=P. Her first-period payoff is then $\underline{w}^*$, and the second-period payoff is no smaller than $w_1(\nu)$. A single principal can therefore obtain at least $(1+\delta)\underline{w}^*$, which exceeds the asymptotic multi-principal payoff $(1+\delta)(1-\nu)\bar{w}^*$.

Proposition 4: For the quadratic cost function the following holds: For any integer n>1, there exist parameter values (for ν, δ and α) such that the principals' aggregate equilibrium surplus is maximal for a number of principals which is finite and at least equal to n. ▪

This result, which is proved in Appendix C, shows that any (finite) number of principals may do better than a single principal. We find by simulations -- reported in the next section -- that the optimal number of principals may be one, two, three etc., depending on the parameters.

The numerical examples indicate that, within some range, the optimal number of principals is an increasing function of the discount factor δ. One can easily verify that for δ sufficiently small, single agency (n=1) is indeed optimal--see Proposition 6 below. Intuitively, the reason is that for δ small, equilibria are fully separating, and the (static) first-period effects then dominate.

It is perhaps less intuitive that for δ sufficiently large, single agency returns to be the optimal agency form. It is well known--and intuitively reasonable--that when the weight attached to the second period is sufficiently large, the single-agency equilibrium is a pooling one. But for given δ, semi-separating equilibria can always be induced by increasing the number of principals. The numerical examples indicate that doing so is beneficial for the principals if δ is not too large. However, when the second-period weight δ becomes "very large", all the multi-principal SS-equilibria are apparently dominated by the single-principal pooling equilibrium. For quadratic cost functions (see the example in Section 4.1) we can prove that (at least for some parameter combinations) single agency is indeed optimal when δ is sufficiently large. The reason is that a highly decentralized structure (many principals) has a serious negative effect on the second-period surplus, because the bonus that is given in that period is extremely low when δ is high. (This is of course precisely the reason why the high ability agent is willing to randomize). The negative effect of the low powered bonus dominates the positive revelation effect for high δ's. This result is stated as Proposition 5, the proof of which is given in Appendix D.

Proposition 5: Suppose costs are quadratic. If (for given probability ν) the two types' productivities are sufficiently close, or if (for given productivities) the probability ν is sufficiently large, the following then holds: For sufficiently large δ single agency (n=1) is optimal, and the equilibrium is a pooling equilibrium. ▪

Our final result is concerned with the polar case of small discount factors.

Proposition 6: Suppose $\Delta'(b)>0$. Then the (optimal) single-agency equilibrium is fully separating if the discount factor δ is sufficiently small. In that case single agency (n=1) is also optimal. ▪

Proof: Let δ be sufficiently small such that the optimal static bonus $b_1(\nu)$ induces full separation (FS), i.e., $b^{**}<b_1(\nu)$ -- see Figure 1. Since a first-period bonus of magnitude P then also is in the FS-region, pooling cannot be an equilibrium, for any $n\geq 1$. Consider, for $n\geq 1$, a bonus b which is semi-separating. Since b<P, more separation will now increase both the first-period and the second-period surplus accruing to the principals. The intertemporal surplus is therefore bounded above by $w_n(\nu)+\delta[\nu \underline{w}^*+(1-\nu)\bar{w}^*]$. Moreover, this value can be attained if (and only if) the optimal static bonus $b_n(\nu)$ induces full separation. So is indeed the case for n=1, and the result then follows from $w_n(\nu)<w_1(\nu)$, $n>1$. ■

5. Numerical examples.

The discount factor δ, which captures the relative importance of the second period, is a key parameter in the study of the trade-off between the static losses and the dynamic gains associated with an increasing number of principals. The effects of varying this parameter are here illustrated by simulations done for the quadratic cost example of Section 4.1. Table 1 below reports the results.

(Table 1 here)

The table illustrates that it is optimal to form a single agency if the second period is of relatively little importance compared to the first period (for low values of δ); $n^*=1$ for $\delta=0,5$. We can also see that as δ increases so does the optimal number of principals; $n^*=2$ for $\delta=1$, $n^*=4$ for $\delta=2$, etc. For even higher values of δ single agency returns to be the best agency form; $n^*=1$ for $\delta=10$. These are the results we typically find, irrespective of what values we choose for the other parameters (ν and α).

To get some intuition for why n^* increases with δ, recall that for SS-equilibria the second-period surplus increases with n. For the first-period surplus, an increase in the number of principals has two effects. First there will be more revelation, which has an ambiguous effect on the principals' surplus (the first-period effect depends on whether the bonus is higher or lower than the marginal value of output P) and, second, the bonus given in the first period will be comparatively low, which has a negative effect on the first-period surplus. As the second period becomes more important, that is, as δ increases, more weight is put on the second-period gain induced by having more principals. This gives the intuition for why an increase of the number of principals from n=1 to n=2 enhances welfare when $\delta=1$ but lowers welfare when $\delta=0,5$.

6. Concluding remarks

A decentralized provision of incentives generates inefficiency if the agency is over contract complements. The reason is that decentralization leads to low powered incentives. This is well known to be true in a static context. In this paper we ask if centralization is optimal also in a dynamic model. And if not, how the optimum degree of decentralization--or delegation--varies with other exogenous variables. We were motivated to study these questions by the observation that the dilution of incentives that follows from a decentralized agency has a positive dynamic effect: It makes it less tempting for the high-productivity type to hide his identity early in the relationship, that is, it mitigates the ratchet effect.

We have shown that in a dynamic framework there are cases where it is never optimal to completely centralize the provision of incentives. The static conclusion does therefore not generalize to a dynamic framework. Our study also gives some interesting indications about how the optimum degree of centralization varies with other parameters. Of particular interest is the parameter that captures the importance of the future relative to the current period. We find that it is optimal to make incentive provision highly centralized if the future is considered to be of either very high or very low importance. In the intermediate case, where the future and the present are of more equal importance, it appears that the optimum degree of delegation increases as the weight attached to the future increases.

In this paper we have used a specific model with an extreme degree of contract complementarity--the agent produces a public good--and we consider only linear contracts. It is an open question whether the insights from this model are robust in the sense that they are valid also in more general settings.

Appendix A: Existence of SS-equilibria.

Recall that in the "well-behaved" case, the semi-separating region SS_n is an interval (b_n^*, b^{**}). Let the function f(n,b) be defined by the RHS of the equilibrium condition (4.8) for b in (b_n^*, b^{**}), and extend f() to the closed interval by taking limits. We then have the following result.

Proposition A. Suppose the principals are symmetric; $p^i = P/n$, $i=1,\ldots,n$. Suppose further that $\Delta(b)$ is strictly increasing, $b_n^* \leq P$ and that

(A.1) $f(n,b) > (<) 0$ as $b < (>) b(n)$, $\quad b \in [b_n^*, b^{**}]$

for some $b(n) \in (b_n^*, b^{**})$. Then an intertemporal equilibrium exists, and the first-period aggregate bonus is unique and given by $b(n) \in SS_n$. ∎

Proof: The definition of f(n,b) as the RHS of (4.8) yields

(A.3) $$f(n,b) = (P-b)([\nu+(1-\nu)(1-x_n)]\underline{y}'+(1-\nu)x_n\bar{y}')$$
$$- n(1-\nu)x_n[\bar{y}-\underline{y}]$$
$$+(P-b)(1-\nu)[\bar{y}-\underline{y}](dx_n/db)$$
$$+ \delta(1-\nu)(\bar{w}^*-w_n(\nu_n)+\nu_n \frac{dw_n}{d\nu_n}(\nu_n)) \frac{dx_n}{db},$$

where we have used the fact that $\nu_n = \nu/[\nu+(1-\nu)(1-x_n)]$ satisfies $[\nu+(1-\nu)(1-x_n)](d\nu_n/dx_n) = (1-\nu)\nu_n$.

For ease of notation we will in this proof drop -- except where absolutely necessary -- references to the index n. In particular, we will use $\hat{b}$ to denote the solution to f(b)=0, thus $\hat{b}$ is the candidate equilibrium bonus. From the expression for f( ) above we have

(A.4) $f(b) = (P-b)g(b) + h(b)$,

where

$$g(b) = ([\nu+(1-\nu)(1-x_n)]\underline{y}'+(1-\nu)x_n\bar{y}') + (1-\nu)[\bar{y}-\underline{y}](dx_n/db)$$

is positive for b in SS_n, and h(b) consists of the remaining terms on the RHS of (A.3).

Consider $\hat{b}\in SS_n$ defined by $f(\hat{b})=0$. We shall verify that $\hat{b}^i = \hat{b}/n$, $\hat{a}^i = \underline{w}(\hat{b})/n$, $i=1,\ldots,n$, is the first-period part of an intertemporal equilibrium. The corresponding second-period strategies are those that divide the second-period aggregate surplus equally among the principals.

Consider principal i, and suppose she contemplates a deviation a^i,b^i from $\hat{a}^i$, $\hat{b}^i$. For $b^i+\hat{b}^{-i}\in SS_n$, where $\hat{b}^{-i}=(n-1)\hat{b}/n$, the marginal value to principal i of increasing b^i is then according to (4.3)-(4.4) given by

$$(p^i-b^i)g(b^i+\hat{b}^{-i})+h(b^i+\hat{b}^{-i})/n=G(b^i)$$

Since $p^i=P/n$ it follows that

$$nG(b^i)-f(b^i+\hat{b}^{-i}) = (\hat{b}^{-i}-(n-1)b^i)g(b^i+\hat{b}^{-i}),\ b^i+\hat{b}^{-i}\in SS_n.$$

Since $g(b)>0$ and $\hat{b}^{-i}=(n-1)\hat{b}/n$, it follows that $nG(b^i)<(>)f(b^i+\hat{b}^{-i})$ for $b^i<(>)\hat{b}/n$. From the assumptions regarding $f(\cdot)$ we then see that $b^i=\hat{b}/n$ is optimal for principal i, provided b^i is restricted to $b^i+\hat{b}^{-i}\in SS_n$.

One can easily show that (A.1) implies $b^{**}>b_n(\nu)$, and that principal i's marginal value therefore is negative for $b^i+\hat{b}^{-i}>b^{**}$ (in the FS region). Finally, the marginal value is easily seen to be positive in the pooling region. Hence no deviation from $\hat{b}^i=\hat{b}/n$ can be profitable for principal i, and the equilibrium is verified. ■

The following corollary shows that Proposition A is not vacuous. (A proof is available from the authors.)

Corollary. In the quadratic case, $f(b^{**}-)<0$ and $b_n^*\leq P$ if and only if

$$\text{(A.5)}\qquad \frac{\alpha\nu(1+1/n)+1-\nu}{\alpha\nu[1+(1+\alpha)/2n] + (1-\nu)K_n} < \mu \leq \frac{\nu\alpha+(1-\nu)K_n}{\nu\alpha+1-\nu}\ .$$

If (A.5) holds, the conditions of Proposition A also hold, and hence there is an n-principal intertemporal equilibrium. The unique first-period equilibrium bonus induces semi-separation.

Appendix B: Proof of Proposition 2.

Note that in the well-behaved case, the lower boundary b_n^* for the SS_n region converges to zero as $n \to \infty$. For fixed $b < b^{**}$, let k be sufficiently large that $b \in SS_n$ for all $n \geq k$.

We shall first prove that for any positive $b < b^{**}$, the aggregate surplus W(n,b) satisfies

$$\text{(B.1)} \qquad W(n,b) \to (P-b)[\nu \underline{y}(b) + (1-\nu)\bar{y}(b)] + \underline{w}(b) \\ + \delta\{(1-\nu)\bar{w}^* + \nu[(P-B(b))\underline{y}(B(b)) + \underline{w}(B(b))]\},$$

where B(b) is the second-period equilibrium bonus, which according to Lemma 2 is independent of n. Indeed, this bonus is given by $\Delta(b) = \delta R(B(b))$. It must satisfy

$$\text{(B.2)} \qquad b_n(\nu_n(b)) = B(b),$$

where $b_n(\ )$ is the static equilibrium bonus, $\nu_n(b) \equiv \nu(x_n(b))$ is the posterior probability that the agent is bad, and $x_n(b)$ is the revelation probability for a good type. Letting $n \to \infty$, it follows from the formula (3.2) for the static equilibrium bonus that $\nu_n(b) \to 1$ and

$$n[1-\nu_n(b)] \to [P-B(b)]\underline{y}'(B(b))/[\bar{y}(B(b)) - \underline{y}(B(b))] \equiv g(b)$$

Since $1-\nu_n = (1-\nu)(1-x_n)/[\nu + (1-\nu)(1-x_n)]$, we then have $x_n(b) \to 1$ and

$$n[1-x_n(b)] \to g(b)\nu/(1-\nu).$$

Note from (4.7) and (B.2) that the principals' second-period surplus can be written as $W_2(n,b) = F_2(x_n(b),b)$, where

$$\text{(B.3)} \qquad F_2(x,b) = [\nu + (1-\nu)(1-x)]\{(P-B(b))[\nu(x)\underline{y}(B(b)) \\ + (1-\nu(x))\bar{y}(B(b))] + \underline{w}(B(b))\} + (1-\nu)x\bar{w}^*.$$

The formula (B.1) for the limit of W(n,b) now follows immediately from (4.5), (4.6) and (B.3), and the fact that $\nu_n \to 1$, $x_n \to 1$.

Next consider the function f(n,b) defined in (A.3). Note first that the term $(1-\nu)(\bar{w}^*-w_n(\nu_n)+\nu_n(dw_n/d\nu_n))$ is the derivative of the principals' second-period surplus w.r.t. the revelation probability, and can thus according to (B.3) be written as

(B.4) $\partial F_2/\partial x = (1-\nu)(\bar{w}^*-\underline{w}(B(b))-(P-B(b))\bar{y}(B(b))) > 0.$

In particular, this term is independent of n.

The equilibrium condition for the first-period bonus takes the form $f(n,b)=0$, where $b=b(n)$ belongs to (b_n^*,b^{**}). For n sufficiently large we certainly have $f(n,b_n^*)>0$ -- since $b_n^*\to 0$ and thus $b_n^*<P$ -- and $f(n,b^{**})<0$-- since the term $n[\bar{y}(b^{**})-\underline{y}(b^{**})]$ dominates. It is relatively straightforward to show that, for any $\varepsilon>0$ $(\varepsilon<b^{**})$, then for n sufficiently large $f(n,b)<0$, all $b>\varepsilon$. From this it is clear that $b(n)\to 0$ as $n\to\infty$.■

<u>Appendix C: Proof of Proposition 4</u>:

The proof is based on five claims. These are stated without proofs; details can be obtained from the authors. The basic idea is to compare values when the two types' productivities are close, i.e. when α is close to one. We choose parameter values such that $b_n^* < P < b_k^* < b^{**}$, all $k<n$. A first-period bonus $b=P$ will then induce pooling if and only if the number of principals (k) is smaller than n, otherwise this bonus will induce semi-separation. We show that the parameters can be chosen such that the intertemporal equilibrium is a pooling one for $k<n$ and a semi-separating one for $k=n$, and that the surplus for the latter dominates.

If the inequalities above (e.g., $P<b^{**}=\mu P$) are to hold for arbitrarily similar types ($1-\alpha$ arbitrarily small), then we must allow for arbitrarily small discount factors. In the first four steps we invoke the following assumption:

Assumption (AD): Suppose $\delta=\delta(\alpha)$ is such that $\mu=[\delta/(1-\alpha)]^{1/2}$ satisfies $[1-\mu]/[1-\alpha]\to\mu'$ as $\alpha\to1$.■

In the last step (Step 6) we choose, for given ν, $\delta=\delta(\alpha)$ and the limit μ' so as to obtain our desired conclusion. We normalize so that P=1 throughout.

Step 1: Given (AD), the SS-boundaries b_n^* and b^{**} satisfy

$$\text{(i)} \qquad \frac{1-b^{**}}{1-\alpha}\to\mu', \qquad \frac{1-b_n^*}{1-\alpha}\to\mu'+n(1-\nu)$$

Moreover, if $b(\alpha)\in[b_n^*,b^{**}]$, all $\alpha\in(\alpha_0,1)$, and

$$[1-b(\alpha)]/[1-\alpha]\to\beta \quad \text{as} \quad \alpha\to1,$$

then,

$$\text{(ii)} \qquad f(n,b(\alpha))/(1-\alpha)\to G_n(\phi),$$

where $\phi=1+(\mu'-\beta)/n$, $\phi\in[\nu,1]$, and

$$G_n(\phi) = [\mu'+n(1-\phi)](1+\frac{\nu}{n\phi^2})-n(1-\frac{\nu}{\phi})+\frac{\nu}{2n\phi^2}$$

In particular,

$$\text{(iii)} \qquad f(n,b^{**})/(1-\alpha)\to G_n(1), \qquad f(n,b_n^*)/(1-\alpha)\to G_n(\nu).$$

Step 2: Given (AD), suppose there is a sequence (α_j), $\alpha_j\to1$, such that for every α_j there is an n-principal SS-equilibrium. Then there must exist $\phi\in[\nu,1]$ such that $G_n(\phi)=0$.

Step 3: Consider n=1. Given (AD), suppose that for every $\alpha\in(\alpha_0,1)$ there is a (single agency) pooling equilibrium. Then the intertemporal surplus $\omega^P(\alpha)$ satisfies

$$\lim[\bar{w}^*-\omega^P(\alpha)]/[1-\alpha]^2 = \mu'+1/2.$$

Step 4: Given (AD), suppose $G_n(\phi)=0$ has a unique solution $\phi=\phi_n$ on $[\nu,1]$. Suppose also that for every $\alpha\in(\alpha_0,1)$ there is an n-principal SS-equilibrium. Let $b(n,\alpha)$ and $\omega(n,\alpha)$ denote the equilibrium bonus and surplus, respectively. Then

$$\lim \frac{1-b(n,\alpha)}{1-\alpha} = \mu'+n(1-\phi_n) \equiv \beta_n$$

$$\lim \frac{\bar{w}^*-\omega(n,\alpha)}{(1-\alpha)^2} = \frac{\nu}{2\phi_n} + \mu' + \frac{1}{2}\beta_n^2 - \beta_n(1-\frac{\nu}{\phi_n})$$

Step 5: For given $n>1$, there is $\nu_0=\nu_0(n)>0$ such that for $\nu<\nu_0$ the following holds: For $\mu'=-n(1-\phi_n)$, where $\phi_n\in(\nu,1)$ is the unique solution to $(\phi_n-\nu)\phi_n = \nu/2n^2$, we have

(i) $G_n(\phi)=0$ has a unique solution $\phi=\phi_n$ on $[\nu,1]$, and $G_n(\nu)>0>G_n(1)$.

(ii) $G_k(\phi)<0$ for all $\phi\in[\nu,1]$, all $k=1,\ldots,n-1$.

(iii) $\mu'[1+\nu/n]-n(1-\nu)+\nu/2n < 0$.

(iv) $\mu'+k(1-\nu) < 0$, all $k=1,\ldots,n-1$.

Proof of Proposition 4: For given $n>1$, let $\nu<\nu_0$ and $\mu'=-n(1-\phi_n)$ be given as in Step 5. For $0<\alpha\leq 1$ define $\mu(\alpha)$ and $\delta(\alpha)$ by

$$\mu(\alpha)=1-(1-\alpha)\mu', \quad \delta(\alpha)=(1-\alpha)\mu(\alpha)^2.$$

Note that by choosing α' sufficiently close to 1 we can be assured that the assumptions of Proposition 3 holds, for any $\alpha\in(\alpha',1)$. We will now show that there is α_0 $(\alpha'<\alpha_0<1)$ such that for every $\alpha\in(\alpha_0,1)$ it holds:

For $\delta=\delta(\alpha)$ we have:

(i) An n-principal equilibrium exists, and any such equilibrium is semi-separating.

(ii) For k<n any k-principal equilibrium is pooling.

(iii) The n-principal surplus strictly exceeds the k-principal surplus, for every k<n.

To prove statement (i), note from Step 1(i) that $\mu'<0$ and $\mu'+n(1-\nu)>0$ implies $b_n^*<P(-1)<b^{**}$ for all α sufficiently close to 1. Moreover, Step 1(iii) and Step 5(i) implies $f(n,b_n^*)>0>f(n,b^{**})$ for all such α. These conditions are sufficient to guarantee that there is, for any such α, an n-principal SS-equilibrium, with a unique first-period equilibrium bonus (see Appendix A). (There can be no pooling or fully separating equilibria.)

Next, Step 2 and Step 5(ii) imply that there is some $\alpha''<1$ such that for $\alpha\in(\alpha'',1)$ there is no k-principal SS-equilibrium, $k\leq n-1$. Since $\mu'+k(1-\nu)<0$, $k\leq n-1$, (see Step 5(iv)), it follows from Step 1(i) that $b_k^*>P(-1)$. Only pooling equilibria are therefore feasible for $k\leq n-1$. This proves statement (ii) in Step 6.

Finally, note that a k-principal pooling surplus is dominated by the single-agency pooling surplus $\omega^P(\alpha)$. From Step 3 and 4 it now follows that for $\alpha\in(\alpha_0,1)$, some $\alpha_0<1$, the value difference $\omega(n,\alpha)-\omega^P(\alpha)$ has the same sign as

$$(1-\nu/\phi_n)/2-\beta_n^2/2+\beta_n(1-\nu/\phi_n),$$

where $\beta_n=\mu'+n(1-\phi_n)$. Since by construction $\mu'=-n(1-\phi_n)$--so $\beta_n=0$--and $\phi_n>\nu$, the conclusion in statement (iii) then follows. ■

Appendix D: Proof of Proposition 5:

Define $\bar{b}$ and $\bar{w}$ by

(D.1) $\bar{b}=P(1-\alpha)\nu\alpha/[(2-\alpha)(\nu\alpha+1-\nu)]$

(D.2) $\bar{w}=(1-\nu)\bar{w}^*+\nu\alpha\bar{b}(P-\bar{b}/2)$

Note that $\bar{b}\to 0$ and $\bar{w}\to(1-\nu)\bar{w}^*$ as $\alpha\to 1$. Note further that the static single agency surplus $w_1(\nu)\to\bar{w}^*$ as $\alpha\to 1$. So, for given ν there is $\alpha_0<1$ such that

(D.3) $\quad w_1(\nu) > \tilde{w}$ for all $\alpha \in (\alpha_0, 1)$.

One sees similarly that for given α, there is ν_0 such that $w_1(\nu) > \tilde{w}$ for $\nu_0 < \nu < 1$. Fix α, ν such that the inequality in (D.3) holds.

For δ sufficiently large it is known that the single agency equilibrium is a pooling one (FGT). The equilibrium surplus is thus given by

(D.4) $\quad \underline{w}(P) + \delta w_1(\nu) \equiv W_1^{Pool}$

Consider $n > 1$. Since $b^{**} = \mu P > b_n(\nu)$ for δ (and hence μ) sufficiently large, all equilibria are then either pooling or semi-separating. If the equilibrium is pooling, the aggregate surplus accruing to the principals is definitely smaller than the single-agency pooling equilibrium in (D.4).

If the equilibrium is semi-separating, the principals' surplus is

$$W(n, b(n)) = F(x_n(b(n)), b(n)),$$

where $F(x,b) = F_1(x,b) + \delta F_2(x,b)$, and $F_1(\ ), F_2(\ )$ are the first- and second-period surpluses, respectively (cfr. Lemma 2 and equation (B.3)). Since $F(x,b)$ is linear in x, we have

(D.5) $\quad F(x,b) \leq \max(\sup F(1,b), \sup F(0,b))$,

where the supremum is w.r.t. $b \in [0, b^{**}]$. The interval $[0, b^{**}]$ contains SS_n, all n, and the proof is therefore complete if the maximum in (D.5) is--for δ sufficiently large--smaller than the single-agency pooling surplus W_1^{Pool}.

It is easily seen that $\sup F(0,b) < W_1^{Pool}$. So consider $\sup F(1,b)$. In the quadratic cost case $F(1,b)$ is also quadratic, and for δ -- and hence μ-- sufficiently large, there is b_μ in $(P, \mu P)$ such that $\sup F(1,b) = F(1, b_\mu)$. Moreover, $b_\mu/\mu \to \tilde{b}$ as $\mu \to \infty$, where $\tilde{b}$ is given by (D.1), and this implies

$$F(1, b_\mu)/\delta \to \tilde{w} \quad \text{as} \quad \delta \to \infty,$$

where $\tilde{w}$ is given by (D.2). The inequality $\tilde{w} < w_1(\nu)$ in (D.3) and the fact that $W_1^{Pool}/\delta \to w_1(\nu)$ then implies that for δ sufficiently large we have $W_1^{Pool} > W(n, b(n))$ for any n which yields an SS-equilibrium. Single agency is thus optimal.

REFERENCES

Baron, D. (1985): "Noncooperative Regulation of a Nonlocalized Externality," 16 Rand Journal of Economics 553-68.

Bernheim, D., and M. Whinston. (1986): "Common Agency," 54 Econometrica 923- 42.

Freixas, X., R. Guesnerie and J. Tirole (1985): "Planning under Incomplete Information and the Ratchet Effect", Review of Economic Studies LII, 173-191.

Fudenberg, D. and J. Tirole (1991): Game Theory, Cambridge, MIT Press.

Grossman, S. and O.Hart (1986): "The costs and benefits of ownership: A Theory of Vertical and Lateral Integration". Journal of Political Economy, 94, 691-719.

Hart, O. and J. Moore (1990): "Property Rights and the Nature of the Firm", Journal of Political Economy.

Laffont, J.-J. and J. Tirole (1988): "The Dynamics of Incentive Contracts", Econometrica 56, 1153-1175.

Laffont. J.J. and J. Tirole (1991): "Privatization and Incentives", Journal of Law, Economics and Organization, 7, 84-105.

Martimort, D. (1990): "Multiple Principals and Asymmetric Information". Mimeo. Université de Toulouse.

Milgrom, P. (1988): "Employment Contracts, Influence Activities and Efficient Organization Design, Journal of Political Economy.

Milgrom, P. & J. Roberts (1992): "Economics, Organization and Management" Englewood Cliffs, N.J: Prentice Hall, 1992.

Olsen, T.E. and G. Torsvik (1991): "The Ratchet Effect in Common Agency: Implications for Regulation and Privatization", LOS-centre WP 91/36, forthcoming in Journal of Law, Economics and Organization.

Stole, L. (1990): "Mechanism Design under Common Agency". Mimeo. MIT.

Ströbele, W.J. (1991): "Institutional Arrangement for Transfrontier River Pollution" in H. Siebert (ed.) Environmental scarcity: the international dimension.

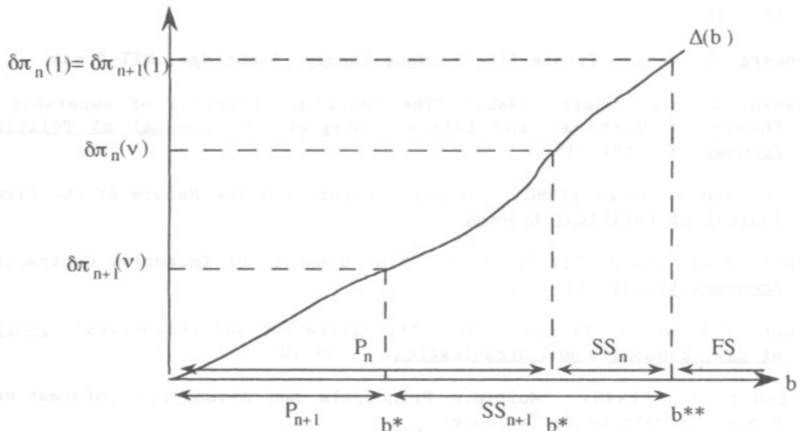

figure 1

Pooling, semi-separation and full separation for n and n + 1 principals.

Table 1: Optimal number of principals for different discount factors and for parameter values $\nu = 0,5$, $\alpha = 0,4$ and $P = 10$.

	n = 1	n = 2	n = 3	n = 4	n = 5	n = 6	n = 10
$\delta = 0,5$	**40,3** **(8,4ss)**	40,2 (7ss)	39 (6ss)				
$\delta = 1$	50,8 (10,3ss)	**53,2** **(8,5ss)**	53 (7,3ss)	51,8 (6,4ss)			
$\delta = 2$	69 (10 p)	75,5 (10,9ss)	77,5 (8,3ss)	**77,6** **(8,3ss)**	76,5 (7,4 ss)		
$\delta = 4$	118 (10 p)	114,8 (14,5ss)	121,7 (12,5ss)	122.6 (11ss)	**123,3** **(9,9ss)**	122,9 (9ss)	
$\delta = 10$	**265** **(10 p)**	251,9 (10 p)	233.6 (18,7ss)	242,2 (16,5ss)	248,4 (14,9ss)	250 (13,5ss)	248,9 (10ss)

The upper number in each row is the equilibrium surplus accruing to the principals under different agency forms. The number in parentheses is the first period bonus; SS indicates semi-separation and p indicates that it induces pooling. The figures in bold are written in the column with the optimal number of principals (n*).

CES Working Paper Series

01 Richard A. Musgrave, Social Contract, Taxation and the Standing of Deadweight Loss, May 1991

02 David E. Wildasin, Income Redistribution and Migration, June 1991

03 Henning Bohn, On Testing the Sustainability of Government Deficits in a Stochastic Environment, June 1991

04 Mark Armstrong, Ray Rees and John Vickers, Optimal Regulatory Lag under Price Cap Regulation, June 1991

05 Dominique Demougin and Aloysius Siow, Careers in Ongoing Hierarchies, June 1991

06 Peter Birch Sørensen, Human Capital Investment, Government and Endogenous Growth, July 1991

07 Syed Ahsan, Tax Policy in a Model of Leisure, Savings, and Asset Behaviour, August 1991

08 Hans–Werner Sinn, Privatization in East Germany, August 1991

09 Dominique Demougin and Gerhard Illing, Regulation of Environmental Quality under Asymmetric Information, November 1991

10 Jürg Niehans, Relinking German Economics to the Main Stream: Heinrich von Stackelberg, December 1991

11 Charles H. Berry, David F. Bradford and James R. Hines, Jr., Arm's Length Pricing: Some Economic Perspectives, December 1991

12 Marc Nerlove, Assaf Razin, Efraim Sadka and Robert K. von Weizsäcker, Comprehensive Income Taxation, Investments in Human and Physical Capital, and Productivity, January 1992

13 Tapan Biswas, Efficiency and Consistency in Group Decisions, March 1992

14 Kai A. Konrad and Kjell Erik Lommerud, Relative Standing Comparisons, Risk Taking and Safety Regulations, June 1992

15 Michael Burda and Michael Funke, Trade Unions, Wages and Structural Adjustment in the New German States, June 1992

16 Dominique Demougin and Hans–Werner Sinn, Privatization, Risk–Taking and the Communist Firm, June 1992

17 John Piggott and John Whalley, Economic Impacts of Carbon Reduction Schemes: Some General Equilibrium Estimates from a Simple Global Model, June 1992

18 Yaffa Machnes and Adi Schnytzer, Why hasn't the Collective Farm Disappeared?, August 1992

19 Harris Schlesinger, Changes in Background Risk and Risk Taking Behavior, August 1992

20 Roger H. Gordon, Do Publicly Traded Corporations Act in the Public Interest?, August 1992

21 Roger H. Gordon, Privatization: Notes on the Macroeconomic Consequences, August 1992

22 Neil A. Doherty and Harris Schlesinger, Insurance Markets with Noisy Loss Distributions, August 1992

23 Roger H. Gordon, Fiscal Policy during the Transition in Eastern Europe, September 1992

24 Giancarlo Gandolfo and Pier Carlo Padoan, The Dynamics of Capital Liberalization: A Macroeconometric Analysis, September 1992

25 Roger H. Gordon and Joosung Jun, Taxes and the Form of Ownership of Foreign Corporate Equity, October 1992

26 Gaute Torsvik and Trond E. Olsen, Irreversible Investments, Uncertainty, and the Ramsey Policy, October 1992

27 Robert S. Chirinko, Business Fixed Investment Spending: A Critical Survey of Modeling Strategies, Empirical Results, and Policy Implications, November 1992

28 Kai A. Konrad and Kjell Erik Lommerud, Non–Cooperative Families, November 1992

29 Michael Funke and Dirk Willenbockel, Die Auswirkungen des "Standortsicherungsgesetzes" auf die Kapitalakkumulation – Wirtschaftstheoretische Anmerkungen zu einer wirtschaftspolitischen Diskussion, January 1993

30 Michelle White, Corporate Bankruptcy as a Filtering Device, February 1993

31 Thomas Mayer, In Defence of Serious Economics: A Review of Terence Hutchison; Changing Aims in Economics, April 1993

32 Thomas Mayer, How Much do Micro–Foundations Matter?, April 1993

33 Christian Thimann and Marcel Thum, Investing in the East: Waiting and Learning, April 1993

34 Jonas Agell and Kjell Erik Lommerud, Egalitarianism and Growth, April 1993

35 Peter Kuhn, The Economics of Relative Rewards: Pattern Bargaining, May 1993

36 Thomas Mayer, Indexed Bonds and Heterogeneous Agents, May 1993

37 Trond E. Olsen and Gaute Torsvik, Intertemporal Common Agency and Organizational Design: How much Decentralization?, May 1993

38 Henry Tulkens and Philippe vanden Eeckaut, Non–Parametric Efficiency, Progress and Regress Measures for Panel Data: Methodological Aspects, May 1993