A General Theory of Delegated Contracting And Internal Control

Wolf Gick *†

New Version: April 14, 2011

Abstract

This paper studies a framework of delegated contracting with a top principal, an intermediary with subcontracting power, and a productive agent who can be of a continuum of types. The intermediary is hired to forward a screening contract to an interval of agent types determined by the principal. Different from the literature, the paper uses a continuous-type setup, with novel findings on the origin and size of the intermediary's rent. Specifically, (1) the intermediary's rent (loss of control, agency cost) is typically lower than in discrete-type frameworks (Faure-Grimaud and Martimort, 2001) where the rent is determined through the span of type difference between the highest and the lowest agent type in the regime, and that (2) measures of internal control furthermore reduce the intermediary's rent that she can reap to fulfill the task (delegation proofness). The framework is generally suitable for a wide range of control and auditing concepts (marginal deterrence, endogenous and exogenous punishment, maximum punishment principle) and the first in the literature where control is applied to a setting with a continuum of agent types. The problem studied in the paper is typical for vertical relationships such as supply chain management and public procurement, when incentive alignment with the intermediary through contract design is important. It furthermore relates to auctions with costly participation where the auctioneer has discretion to exclude a nonzero measure of buyer types. It also relates to findings in efficient tax literature under asymmetric information.

JEL classification: D23, D73, D82, L51.

Keywords: Delegated Contracting, Vertical Hierarchies, Loss of Control, Agency Costs.

Paper submitted to the 22nd International Conference on Game Theory at Stony Brook University, July 2011

*Center for European Studies, Harvard University, ph. (617) 495 4303 x244, email: gick@fas.harvard.edu

†School of Economics, Georgia Institute of Technology, 221 Bobby Dodd Way, Atlanta, GA 30332. I am truly indebted to Philippe Aghion, Jesse Bull, Gorkem Celik, Bob Gibbons, Werner Güth, Dilip Mookherjee, Thomas Jeitschko, Simon Loertscher, Guillaume Roger, Emilson Silva, Jean Tirole, and my discussant Liad Wagman for helpful discussions who all helped to shape this paper into its current form. Earlier versionshave been presented to the Research Seminar at LMU Munich, the ESI Spring 2010 Workshop at MPI Jena, the Midwest Fall 2010 Economic Theory Meetings, the 2011 IIOC Conference as well as to department seminars at Georgia Tech and Florida International University. All errors are mine.

1 Introduction

Williamson's (1967) seminal paper on the origins of a "loss of control" across successive vertical layers of a hierarchy has opened several avenues for research. A first literature on moral hazard and supervision, initiated by Calvo and Wellisz (1978), has analyzed under which conditions efficient performance monitoring will involve times when a productive worker will not be "checked" by the supervisor. Qian (1994), by adding a "loss of control" feature in the Calvo and Wellisz's model, has reached a general result in a model where the span of control (i.e. the number of subordinates per supervisor) and wages are determined endogenously. Under certain circumstances, higher effort levels and higher wages at upper tiers in the hierarchy may outweigh the loss of control. This, to a large extent, contrasts with Mirrlees' (1976) findings, namely that "middle managers should get more than workers."[1]

While Williamson's (1970) analysis on firm size was based on communication losses following losses in supervision that cannot be eliminated because of the underlying limits of supervision, the nature of the information loss across vertical layers has remained an unsolved issue in models with moral hazard.[2] The difference comes with the introduction of the revelation principle in settings for more than one agent. Once the revelation principle is established with all downstream players, their information rents of other than productive players can be quantified: a supervisor gets paid the rent for truthfully executing his task in a Grand Contract. This all follows, in spirit, Laffont's (1990) early analysis who called for an analysis of information rents at different stages of the contracting game. When downstream players have an informational advantage, such rents can be the result of limited contract design options of the top principal.

An important step forward in analyzing the "loss of control" has been made by McAfee and McMillan (1995) who show that, if an intermediary is protected by limited liability

[1] Mirrlees (1976), p. 130.

[2] Later contributions have shed important light on issues of adverse selection and hidden information in supervision settings. See e.g. Faure-Grimaud, Laffont and Martimort (2002 and 2003) in settings with a Grand Contract, which I adopt here as well. Still, as Mookherjee (2006) observes in his overview paper, this literature has been unable to generally explain why delegation may dominate centralization.

and endowed with sub-contracting power, delegated contracting is plagued by a double-marginalization of rents. The intermediary is then in a position to extract information rents from both the downstream agent and the top principal.[3]

The present paper develops a model in which the intermediary's rent is determined endogenously, in a setting where the productive agent's type space is continuous. The "loss of control" that I study is typical for two types of real-world organizations. First of all, it occurs as a control problems inside business firms. The literature on corporate re-engineering takes up this issue.[4] Little has been said in the incentive and contract design literature about *how* to reduce this loss of control.

Consider a multinational firm with a value chain. The CEO has to set up operations with a division manager who has institutional knowledge and therefore delegates the task of contracting with important input providers to the division manager. By doing so, he endows her with sub-contracting power. This form of decentralization, as Horngren et al. (2003) point it out from a cost accounting perspective, typically comes with suboptimal decision making practices that are ameliorated by the use of management control systems and performance budgeting. Such internal forms of control to restore congruence between top-layers and middle management of the hierarchy are commonly used. In other words, large firms will typically rely on some benchmarks for the design of performance evaluation systems.[5]

Secondly, the loss of control across vertical layers is also a widespread phenomenon in public agencies where intermediate layers of the hierarchy have sub-contracting power. In both cases there exist forms of internal control such as bookkeeping and performance budgeting systems that are designed to improve efficiency, including internal control.[6]

[3]Mookherjee and Tsumagari (2004) study a model with a productive intermediary (or prime supplier) who underproduces to minimize the downstream agents information rent. In the setting studied here, the intermediary has no productive task.

[4]See Brynjolfsson and Hitt (2000).

[5]See in particular the case study on reaching supply chain efficiency (HBS, 2000). Furthermore, the setting relates to classic case of Lincoln Electric (HBS, 1982), designing contracts relies on stability. Supply chain management that involves extensive outsourcing may at the same time require adjustments toward the adaptation of internal control systems, which is, in some multinational firms, not always viable. I thank Christine Ries for pointing this out.

[6]Government auditing standards (see e.g. USGAO 2010) and management accounting systems aim at limit the discretion of intermediate managers or bureaucrats.

1.1 Relationship to the Literature

While focusing on typical problems of contract design and control in vertical hierarchies, the paper relates to the larger strand of delegation and supervision. Traditionally, most papers on delegation deal with some shifts in the *relative* advantage between centralized and decentralized organizational forms. As Che and Kim (2006) argue, delegation of contracting authority is difficult to justify in such settings. In turn, models that incorporate collusion such as Faure-Grimaud, Laffont and Martimort (2003), Mookherjee and Tsumagari (2004) and Celik (2009) do not necessarily show an advantage of decentralization in the case of collusion.[7]

This paper is not on collusion, nor does it directly focus on a comparison of delegation which centralization. Given the setup, there is no general trade-off to expect from specialization and the reduction of information processing costs through having an intermediary offering the contract. Instead, the paper characterizes a delegation benchmark and so explains the emergence of a loss of control under delegated contracting in a one one-agent setting with a continuum of types. More generally, it shows that delegated contracting can be improved through internal control.

A last previous theory of delegated contracting that has built on McAfee and McMillan (1995) is Faure-Grimaud and Martimort (2001, FGM hereafter). These authors assume a discrete type setting in which the downstream agent can be of three possible types, with the special twist that the screening contract should be only given to two types. Information, rents, and communication become intertwined in the following way: the intermediary is hired because of her ability to costlessly filter out the unwanted third type of agent but to offer a Baron-Myerson (BM) style contract to the two remaining types, which she still cannot distinguish. While the great merit of their contribution is to shift the analysis from the simple moral hazard in supervision problem into the task of forwarding a screening contract to the downstream (productive) agent, there are several issues that require more attention. First, In

[7]Che and Kim (2006) show in particular that delegation cannot be more justifiable in the presence of collusion than in its absence. My paper does not focus on a setting with two or more productive agent in the Marschak-Radner tradition such as Mookherjee and Tsumagari (2004) and Cella (2005).

Gick (2008) I have extended the FGM setting by applying a viable auditing scheme to the contract, showing that the top principal can always reduce the loss of control that emerges from the intermediary's discretion.

Second, and more importantly, I show that the assumption of a *discrete* type space where two out of three types are to being given a contract will lead to an overstatement of the intermediary's rent. The task of the present paper is to study a general framework in which a BM-style contract should be offered to a productive agent who is of a continuum of types. I thus study internal control by providing a general framework. As I will show, a theory of delegated contracting will lead to quite different results when analyzed in a continuous-type framework. As known from related fields, discrete-type settings may show extreme results when e.g. one more agent is added.[8] A next advantage of an extension toward a setting with a continuum of agent types is the option to determine some key parameters endogenously. This not only relaxes the assumptions made in FGM, it also replaces the assumption of a risk-averse intermediary in their discrete type setting, which drives a wedge between possible contract pairs in favor of a more intuitive treatment.

Lastly, the paper offers an auditing scheme for a continuous type setting. Auditing has so far been studied for productive agents at the bottom of a hierarchy, and are typically restricted to a treatment with two discrete types only.[9] By applying an internal control scheme between the two top layers of a hierarchy, the paper extends the scope of auditing to intermediate players and provides a new analysis.

The remainder of the paper is organized as follows. Section 2 presents the model, section 3 offers an auditing scheme, section 4 concludes.

[8]From a theory perspective, continuous-type models are seen as being based on a less restrictive setup, permitting a variety of types. More generally, continuous type model lead to quite distinct results compared to simple discrete-type settings. See e.g. the comparison by Armstrong and Rochet (1999) of discrete versus continuous type models of multidimensional screening.

[9]Border and Sobel (1987) provide an analysis of the set of binding incentives constraints for a two-player setting with a finite number of agent types. With more than two discrete types, the incentive problem becomes intractable.

2 Model

2.1 Players, preferences and payoffs

There is a principal P, an intermediary I, and an agent A.[10] The agent produces a quantity q of output at a marginal cost θ, which is his private information. θ is drawn from continuous distribution on the support $[\underline{\theta}, \overline{\theta}]$, with a c.d.f. of $F(\cdot)$ and a p.d.f. of $f(\cdot)$ that is positive for all θ.

The distribution is common knowledge; I furthermore require that the monotone hazard rate condition holds for the distribution, that is, $\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right)$ is assumed to be nonnegative, and the distribution is well defined and differentiable nearly everywhere over the entire interval. The agent is risk-neutral and has a utility function $U = t - \theta q$, where t is the monetary transfer he receives from the intermediary. The agent accepts to produce as long as he gets his reservation utility exogenously normalized to zero.

The intermediary is hired to forward a screening contract to the productive agent. The rents for this contract are included in the budget s. Specifically, the principal requests the intermediary to offer this screening contract to all types of agents in the interval $[\underline{\theta}, \hat{\theta}]$, which would maximize his surplus. The intermediary is risk-neutral but protected by limited liability below zero wealth. She has preferences of $V = s - t$. In other words: subtracting the transfer to the agent, t, from the intermediary's budget s yields the intermediary's income.

The principal has no access to a productive agent and thus cannot offer a contract himself. His gross surplus is $S(q)$, with $S(0) = 0$, $S'(q) > 0$ and $S"(q) < 0$. To ensure positive production levels, I assume that the Inada conditions hold, that is $R'(0) = +\infty$ and $R'(+\infty) = 0$. The principal's net surplus is simply his gross surplus minus the budget paid to the downstream hierarchy $S(q) - s$.

[10]As usual in this literature, I use male pronouns for principal and agent, and a female for the intermediary.

2.2 Timing

The contracting game has the following extensive form:

- ($t = 0$) Agent learns its type θ. Intermediary learns the agent's type only if $\theta > \hat{\theta}$.
- ($t = 1$) P offers a Grand-Contract to I specifying output targets and transfers
- ($t = 2$) Intermediary accepts or rejects.
- ($t = 3$) Subcontracting stage: intermediary offers a sub-contract to agent.
- ($t = 4$) Agent accepts or rejects.
- ($t = 5$) Production and transfers take place. Outputs are observed. Game ends.

2.3 Agent's constraints

To find the optimal contract for the intermediary as a principal who optimally should forward a screening contract over the range $[\underline{\theta}, \hat{\theta}]$ with $\hat{\theta} < \overline{\theta}$, I first reduce the setup to a two-layer vertical hierarchy consisting of a principal and the productive agent. This step is easy to motivate: as in FGM and related settings, the rent of the intermediary is a function of the rent of the productive agent. To find this rent, I characterize the optimal contract for the productive agent over the interval $[\underline{\theta}, \hat{\theta}]$.[11]

As derived in the appendix, the closed-form expression of the principal's program is:

$$\underset{\{U(\cdot),q(\cdot)\}}{max} \int_{\underline{\theta}}^{\hat{\theta}} \Big(S(q(\theta)) - \Big(\theta + \frac{F(\theta)}{f(\theta)}\Big) q(\theta) q(\theta) \Big) f(\theta) d\theta. \tag{P}$$

with the known solution of[12]

[11] This is a benchmark known in the literature; for an exposition see Appendix 1.

[12] See e.g. Laffont and Martimort, 2002.

$$S'(q^{SB}(\theta)) = \theta + \frac{F(\theta)}{f(\theta)}, \tag{10}$$

which entails no downward distortion for the most efficient type $\underline{\theta}$ and a decreasing schedule of outputs q for all other types. This output distortion is second-best optimal.

2.4 Grand Contract

Based on this contract, I now characterize the optimal contract in a vertical hierarchy through the optimal Grand Contract, which is carried out sequentially. The goal is to find the optimal solution and to check if additional distortions become optimal when an intermediary is added to the hierarchy.

Definition 1 ***Grand-Contract.*** A Grand-Contract $GC = s(q)$ is a direct truthful mechanism that satisfies all rents of all players in the hierarchy, with the revelation principle applying a the sub-contracting stage.

Without loss of generality, I restrict attention to direct truthful mechanisms (t, q) between I and A that include budgeting between P and I with I receiving a payment $s - t$. To see this, define the contract from the intermediary's perspective, replacing transfers as functions of output targets q, and expressing the expected utility of the intermediary when designing the screening contract for the agent:

$$\underset{\{(U,q)\}}{max} \int_{\underline{\theta}}^{\hat{\theta}} V(s(q)) - \theta q - U(\theta) f(\theta) d\theta. \tag{GC$_{SC}$}$$

2.4.1 Delegation Proofness

Should the intermediary be incentivized to offer a screening contract for all types in $\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right)$, the optimal sub-contract needs to implement optimal output targets $q(\theta)$, fulfilling the following incentive constraints between intermediary and agent:

$$s(q(\theta)) - \theta(q(\theta)) \geq s(q(\theta')) - \theta(q(\theta')) \quad for\ all\ \theta \leq \theta'.$$

This implies that for more efficient agent types, I needs to receive an at least weakly higher payment scheme, and that this must hold for all output targets in the contract.

Similarly to the *Revelation Principle* for the simple two-player Principal-Agent model, I now apply the *Delegation Proofness Principle* that implements a truth-telling contract across the hierarchy, with optimal output targets and information rents paid to each player: take any GC such that the optimal sub-contract recommends production $\tilde{q}(\theta)$ and a production target of zero for all types of agents with higher marginal costs than envisaged by the principal. It must then be possible to establish a delegation proof contract as a direct mechanism defining output targets $q = \tilde{q}(\theta)$ and budgets $s = s(\tilde{q}(\theta))$ such that the intermediary truthfully reveals the agent's type to the principal. If not, the envisaged output target $\tilde{q}(\theta)$ would not have been optimal in the first place. This can be summarized in the following statement using contracts as pairs of output targets and budgets.

Definition 2 ***Delegation Proof Grand-Contract.*** *A delegation-proof Grand-Contract for a continuum of agent types consists of a menu $\{(s, q)\}$ for all agent types, satisfying the intermediary's incentive constraints for each intermediary-agent coalition with $t = \theta q$ being the transfer paid to the agent, and $s - t$ the part of the budget that I keeps for herself:*

$$s - \theta q \geq s' - \theta(q'),$$

where $s' > s$ and $q' > q$ denote budgets and outputs that are suboptimal for types $\theta' > \theta$.

Furthermore, the intermediary is protected by limited liability against outcomes below zero, with her rent being $V(\theta)$ being

$$V(\theta) \geq 0. \tag{LL}$$

For an agent type of $\hat{\theta}$ there is no rent to be offered by I. That is, would P want that I offers a contract only to the $\hat{\theta}$-type of agent, there would be no need to include any rent. For all lower types, the intermediary's rent may become positive.

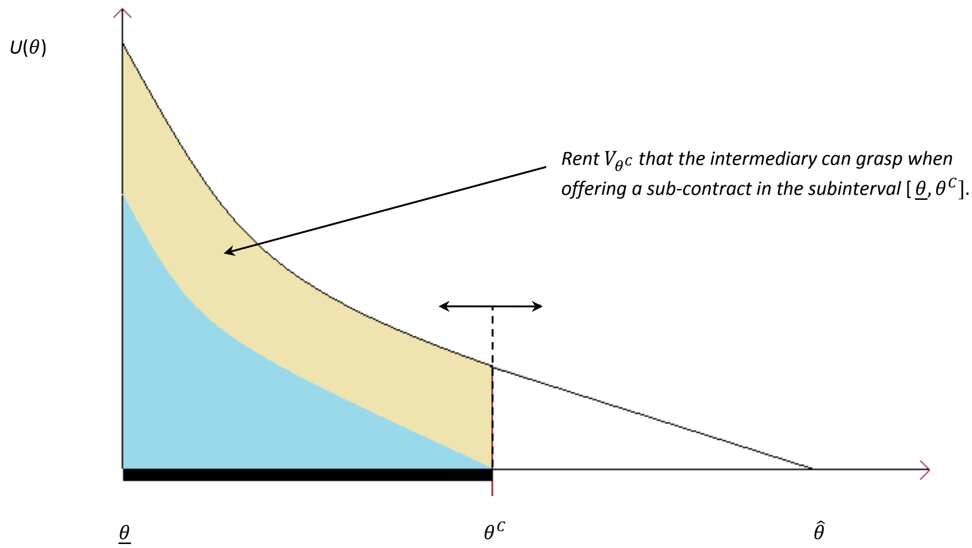
2.4.2 Incentive Constraint of Intermediation

In a (real) two-type setting as in FGM, the intermediary's rent follows from her option to "gamble" and to offer a contract to the most efficient type only. Specifically, she can offer a shutdown contract to the most efficient agent in the discrete type setting, reaping the entire information rent of this type with a probability that is exogenously given. If the intermediary loses the gamble, no contract exists and no production occurs. If she succeeds, she pockets the entire rent she is hired to include in the contract design.

The results of the continuous type setting differ from FGM in two important aspects. First, the intermediary cannot offer a contract to a close to zero mass of agent types on the left side. Second, since the mass between the most and the least efficient agent type is positive everywhere for any nonzero subset of types, the rent analysis for the intermediary is built on an entirely different concept.

The optimal contract can be characterized using the following steps. Observe first, the intermediary requires a strictly positive rent for truthfully forwarding a BM-style contract to all agent types in the interval $[\underline{\theta}, \hat{\theta}]$ except for the highest type. If not, the contract would not be delegation proof and the intermediary could do better to cut off a positive mass of agent types and offer a BM-style screening contract to a subinterval from the most efficient type up to a cutoff type $[\underline{\theta}, \theta^C]$ that maximizes her rent.
I characterize the origin of the intermediary's rent (loss of control, agency cost) for any $\theta^C \in [\underline{\theta}, \hat{\theta}]$:

$$E(V_{\theta^C}) = \int_{\underline{\theta}}^{\hat{\theta}} U(\theta)d\theta - \int_{\theta^C}^{\hat{\theta}} U(\theta)d\theta - \int_{\underline{\theta}}^{\theta^C} U(\theta)d\theta \tag{11}$$



Proposition 3 *The intermediary chooses a cutoff value $\theta^C \in [\underline{\theta}, \hat{\theta}]$ that maximizes her rent $E(V^*_{\theta^C})$. The principal includes $E(V^*_{\theta^C})$ in the Grand Contract to reach delegation proofness. Given single-peaked densities that satisfy monotone hazard rate property, there is always an interior solution for $E(V^*_{\theta^C})$.*

The above illustration (*Fig. 1*) sketches the optimal cutoff that the intermediary chooses to maximize her rent. The findings are summarized in the following proposition.

Proposition 4 *Because of her subcontracting power, the intermediary can offer any subcontract to A over a subset of types. For any leftbound subset of the contracting space, $[\underline{\theta}, \theta^C]$, with $\theta^C < \hat{\theta}$ she can reap a strictly positive rent. If the contract was accepted, she was able to pocket parts of the information rent included in the budget for herself. If not, no contract exists and the intermediary misreports the type being in the interval $[\hat{\theta}, \bar{\theta}]$.*

As a last exercise, I proceed similar to Section 2.3 where we derived (P") and first set up the Grand-Contract over rents and outputs:

$$\max_{\{U(\cdot),V(\cdot),q(\cdot)\}} \int_{\underline{\theta}}^{\hat{\theta}} \Big(S(q(\theta)) - \theta q(\theta) - U(\theta) - V(\theta)\Big) f(\theta) d\theta. \tag{GC}$$

Proposition 5 ***Optimal Delegation-Proof Grand-Contract and the Span of Contracting***. *The optimal contract entails additional distortions compared to second-best outputs, with*

$$S'_{GC}(q^{SB}(\theta)) = \theta + \frac{F(\theta)}{f(\theta)} + \frac{F(\theta)}{f(\theta)} - \frac{F(\theta)}{f(\theta)}\bigg|_{\theta^C}^{\hat{\theta}} - \frac{F(\theta)}{f(\theta)}\bigg|_{\underline{\theta}}^{\theta^C}. \tag{12}$$

Note that the last three parts of the sum follow from the additional rent that the delegation proof Grand-Contract includes. This increases the slope of the principal's surplus function and reduces efficiency. The proof is given in the appendix.

In particular, the result points out the fact that more generally, an intermediary's rent under delegated contracting *in a continuous type setting* can be expected to be lower than characterized in the discrete type setting of FGM. The rent is lower for two reasons. In their model, the intermediary's rent is a function of the lowest type probability times the information rent of the lowest type based on the span between lowest and the middle type $\hat{\theta}$, that is, the highest type that should be given a contract. First of all, with a continuum of agent types, requesting such a rent scheme would involve basing the rent on a close to zero measure of types, which cannot be optimal.

Secondly, and related, the claim of the agent to misrepresent the type space, and to so request a rent, will be based on a lower span between the highest and the lowest type θ^C and $\underline{\theta}$. Thus, the intermediary cannot claim a rent based on the entire span of contracting but only on a significantly smaller subset.

3 Auditing

This section shows that it is possible to audit a contract under a continuous type setting, to restore efficiency to some degree. I assume that the principal has access to a costly audit technology and can commit to audit the intermediary whenever a contract exists. The intermediary is now imposed a penalty P^s with probability $\underline{\wp}$ if the examination of the written subcontract detects an irregularity in the contract offer, while with probability $1-\underline{\wp}$ she keeps her information rent as before. This simple control scheme does not need to involve probabilistic auditing; as long as the principal commands a costly but fully revealing examination technology, it is sufficient for the principal to examine the contract whenever output was observable. The argument is similar as in Gick (2008) where I have provided an auditing scheme to the FGM contract, and it is based on the same assumptions.

$$\wp\Big(E(V(\theta)^*) - P^s\Big) + (1-\wp)E(V(\theta)^*)$$

With endogenous punishment, $E(V(\theta))$ reduces to

$$E(V(\theta)^A) = (1-\wp)\Big(\int_{\underline{\theta}}^{\hat{\theta}} U(\theta)d\theta - \int_{\theta^C}^{\hat{\theta}} U(\theta)d\theta - \int_{\underline{\theta}}^{\theta^C} U(\theta)d\theta\Big) < E(V(\theta)^*). \qquad (13)$$

The principal's problem under auditing changes to:

$$\max_{\{U(\cdot),V(\cdot),q(\cdot),c(\wp)\}} \int_{\underline{\theta}}^{\hat{\theta}} (S(q(\theta)) - \theta q(\theta) - U(\theta) - V(\theta))f(\theta)d\theta$$

Proposition 6 *Under auditing, the optimal downward distortion in the solution to the delegation proof Grand-Contract is reduced. The principal chooses an audit probability $\wp$ equal to the marginal cost of auditing and the optimal Grand-Contract now entails:*

$$S'_A(q^{SB}(\theta)) = \theta + \frac{F(\theta)}{f(\theta)} + (1-\wp)\Big(\frac{F(\theta)}{f(\theta)} - \frac{F(\theta)}{f(\theta)}\Big|_{\theta^C}^{\hat{\theta}} - \frac{F(\theta)}{f(\theta)}\Big|_{\underline{\theta}}^{\theta^C}\Big) < S'_{GC}(q^{SB}(\theta)). \qquad (14)$$

The proof is analog to the one of *Proposition 5* (utilizing (14)) and omitted.

As a result, *Proposition 6* shows that auditing increases efficiency and reduces the optimal distortion in the delegation proof Grand-Contract. It decreases the slope of the surplus function for all θ and improves efficiency of the Grand Contract. Note that the setup is extendable to additional forms of auditing with more complex structures, including external audits, and audits to which the principal cannot commit.[13]

4 Concluding remarks

The paper has presented a generalized treatment of delegated contracting and internal control. The central findings of the paper are the following. First, it extends the findings of the previous literature on delegated contracting and organizational diseconomies of scale in the following three ways:

• The origins of agency costs or a loss of control are no longer directly based on the span of control as in FGM, that is, the type difference between the highest and lowest cost type. Instead, as this paper has shown, in a continuous type setting the intermediary can only grasp *parts* of the rent that the top principal includes to be forwarded to the productive agent.[14]

• The continuous-type model furthermore shows that the intermediary will always cut off high-cost types of agents from receiving a contract. While this option of the intermediary is the source of inefficiency, she does not have an incentive to exclude *efficient* types in the regime. Note also that the paper endogenously derives the rent of the intermediary.

• There exists a simple auditing scheme that always reduces inefficiency. Auditing schemes of this kind are typically used together with budgeting techniques in firms to enhance efficiency. The novel contribution of this paper is that it offers an auditing scheme which works with continuous types of agents. Adverse selection models with an auditing scheme for the productive

[13]See e.g. Khalil (1997) for an overview of other possible forms of auditing in BM-style contracts.

[14]In FGM, the rent is the prior of the efficient agent's type times the entire virtual cost $\Delta\theta\bar{q}$. In this way, the findings of my paper show that vertical hierarchies, in a more general setting, show a lower loss of control compared to the result of discrete type models.

types typically restrict their attention to discrete types only.

The results shed new light on the specific issues of agency costs that typically arise in large multinational firms and in public bureaucracies. Seen in a continuous type framework, the nature of the "loss of control" is quite different, and the setting itself limits the options of the intermediary of requiring rents up to the level of the downstream agent's rent. That is, a top principal should be aware of losses in the hierarchy, but adding a next layer to a hierarchy does clearly not imply doubling information rents as modeled in the literature. In addition, the paper has shown that even if the top principal has no access to the bottom tier, he generally has some leeway to reduce the loss through auditing.

As already stressed, the present paper is not about collusion. Still, collusion constitutes a borderline case that may be worth extending: the intermediary offers a take-it-or-leave-it subcontract and never leaves any rent to the productive agent beyond what a standard screening contract would encompass. Collusion does not arise in the model because the intermediary does not observe a partition of the contracting space that would endow him with additional information.[15]

Several extensions are possible that were either mentioned or implied in the above setup. While this paper has served the task to lift the literature of delegated contracting onto a new level of generality, it could, firstly, be an interesting task to separate the intermediary's observation space from the contracting space, to so open the setting toward collusion with a subset of agent types close to θ^C.

Similarly, there are a series of additional auditing forms that become available based on the findings of this paper. These, as already mentioned, can involve auditing without commitment, external auditing, as well as the use of specific penalties that may be typical for public bureaucracies.

Lastly, and from a theory perspective, two last comparisons and extensions may be worth exploring. First, the findings of this paper on the intermediary's rent are reminiscent of the

[15] Celik's (2008) paper is, to my knowledge, the only contribution so far using a continuum of agent types in a setting with principal, insurer and agent where the principal facing a budget constraint. The vertical structure is different to mine, and there is no endogenosly determined information rent of the insurer.

literature on optimal income tax schedules under asymmetric information. Mirrlees (1971 and 1997), as well as Seade (1982) and Guesnerie and Seade (1982) show properties of an optimal tax for different, continuous income "types" that resembles the endogenously determined information rent.[16] Second, and as mentioned at the beginning, the setup between top principal, intermediary and agent is similar to an auction environment with costly participation involving with a seller, auctioneer and a continuum of buyers. If the auctioneer has discretion to exclude a nonzero subset of buyer types, e.g. by designing access in a way that the seller would find suboptimal, the optimal contract will require additional information rents for the auctioneer.[17] This, as well, is left for future research.

5 Bibliography

Brynjolfsson, E., Hitt, L. M., 2000, Beyond Computation: Information Technology, Organizational Transformation and Business Performance, Journal of Economic Perspectives 14(4), 23-48.

Bull, J., Watson, J., 2004, Evidence Disclosure and Verifiability, Journal of Economic Theory, 118, 1-31.

Calvo, G., Wellisz, S., 1978. Supervision, Loss of Control, and the Optimum Size of the Firm. Journal of Political Economy, 86(5), 94352.

Celik, G., 2008. Counter Marginalization of Information Rents: Implementing Negatively Correlated Compensation Schemes for Colluding Parties. B.E. Journal of Theoretical Economics (Contributions): 8(1), Article 3.

Celik, G., 2009. Mechanism design with collusive supervision. Journal of Economic Theory 144, 69-95.

[16]I thank Emilson Silva for pointing this out.

[17]See Menezes and Monteiro (2000) for an overview as well as McAfee and McMillan (1987a,b). Celik and Yilankaya (2009) study an optimal auction with stochastic bidder participation and endogenous cutoffs.

Celik, G., Yilankaya, O., 2009. Optimal Auctions with Simultaneous and Costly Participation, mimeo, available at http://sites.google.com/site/gorkemcelikswebsite/

Cella, M., 2005. Monitoring Subcontracting in a Suppliers' Hierarchy, mimeo, Discussion Paper Series, Department of Economics, Oxford University.

Faure-Grimaud, A., Laffont, J.-J, Martimort, D., 2000. A Theory of Supervision with Endogenous Transaction Costs, in: Annales of Economics and Finance 1, 231-263.

Faure-Grimaud, A., Laffont, J.-J, Martimort, D., 2002. Risk Averse Supervisors and the Efficiency of Collusion, in: Contributions to Theoretical Economics (The B.E. Journals in Theoretical Economics) 2(1), Article 5.

Faure-Grimaud, A., Laffont, J.-J, Martimort, D., 2003. Collusion, Delegation and Supervision with Soft Information, in: Review of Economic Studies 70, 253-279.

Gick, W., 2008. Delegated Contracting, Information, and Internal Control. Economics Letters 101, 179-83.

Guesnerie, R., Seade, J., 1982. Nonlinear Pricing In a Finite Economy. Journal of Public Economics 17, 157-179.

Hammer, M., Champy, J., 1993. Reengineering the corporation: A manifesto for business revolution. New York: Harper Business.

HBS, 1992, Harvard Business School Case "The Lincoln Electric Company." 9-376-028 Rev. of July 29, 1983.

HBS, 2000, Harvard Business School Case "Aligning Incentives for Supply Chain Efficiency." 9-600-110.

Horngren, C.T., Datar, S.M., Foster, G., 2003. Cost accounting: a managerial emphasis. Pearson, Upper Saddle River, NJ.

Khalil, F., 1997. Auditing without Commitment. Rand Journal of Economics 28(4), 629-640.

Laffont, J.-J., 1990. Analysis of Hidden Gaming in a Three-Level Hierarchy. Journal of Law, Economics, and Organization 6, 2, 301-324.

Laffont, J.-J., Martimort, D., 2002. The Theory of Incentives: The Principal-Agent Model, Princeton: Princeton University Press.

McAfee, R. P., McMillan, J., 1987a. Auction with entry. Economics Letters 23, 343347

McAfee, R. P., McMillan, J., 1987b. Auctions with a stochastic number of bidders. Journal of Economic Theory 43, 119.

McAfee, R. P., McMillan, J., 1995. Organizational Diseconomies of Scale. Journal of Economics and Management Strategy 4(3), 399-426.

Melumad, N. D., Mookerjee, D., and Reichelstein, S., 1992. A Theory of Responsibility Centers. Journal of Accounting and Economics 15, 445-84.

Melumad, N. D., Mookerjee, D., and Reichelstein, S., 1997. Contract Complexity, Incentives, and the Value of Delegation. Journal of Economics and Management Strategy 6(2), 257-89.

Melumad, N. D., Reichelstein, S., 1989. Value of Communication in Agencies. Journal of Economic Theory 47(2), 334-68.

Menezes, F. M., Monteiro P.K. 2000. Auctions with endogenous participation. Review of Economic Design 5, 71-89.

Mookerjee, D., 2006. Delegation, Hierarchies, and Incentives: A Mechanism Design Perspective. Journal of Economic Literature 44, 367-90.

Mirrlees, J. A. 1976. The optimal structure of incentives and authority within an organization. Bell Journal of Economics 7(1), 105-131.

Mirrlees, J. A. 1971. An exploration in the Theory of Optimum Income Taxation. Review of Economic Studies 38, 175-208.

Mookerjee, D., 2006. Delegation, Hierarchies, and Incentives: A Mechanism Design Perspective. Journal of Economic Literature 44, 367-90.

Mookerjee, D., Tsumarari, M., 2004. The organization of supplier networks: effects of delegation and intermediation. Econometrica 72, 1179- 1219.

Qian, Y., 1994. Incentives and Loss of Control in an Optimal Hierarchy. Review of Economic Studies 61, 527-544.

Seade, J., 1977. On the shape of optimal tax schedules. Journal of Public Economics 7, 203-235.

Tirole, J., 1986. Hierarchies and Bureaucracies: On the Role of Collusion in Organizations. Journal of Law, Economics and Organization 2, 181-224.

USGAO (2010). Government Auditing Standards - 2010 Exposure Draft, United States Government Accountability Office, available at http://www.gao.gov/new.items/d10853g.pdf

Williamson, O. E., 1967. Hierarchical Control and Optimal Firm Size. Journal of Political Economy 75(2), 123-38.

6 Appendix

6.1 Appendix 1: Characterizing the two-player contract for the type interval $[\underline{\theta}, \hat{\theta}]$

To a large extent, my exposition follows Laffont and Martimort (2002) (LM hereafter) at this point, except that the contracting type space is reduced to a subinterval, permitting the typical and "garbled" information structure as in FGM, in which the intermediary always has the option to misrepresent a type within $[\underline{\theta}, \hat{\theta}]$ as a type outside this interval. For the continuous type case in my model we have $F(\theta)\Big|_{\underline{\theta}}^{\hat{\theta}} < F(\theta)\Big|_{\underline{\theta}}^{\bar{\theta}} = 1$. This follows from the assumption implicitly made in FGM that the reason to hire the intermediary to "filter out" an unwanted agent of a higher type.

To characterize the continuous-type 2-player contract I restrict attention to direct revelation mechanisms $q(\tilde{\theta}), t(\tilde{\theta})$ for which

$$t(\theta) - \theta q(\theta) \geq t(\tilde{\theta}) - \theta q(\tilde{\theta}). \tag{1}$$

Truthful revelation contracts need to satisfy the following first order condition for the truthful response $\tilde{\theta}$ to become optimal:

$$\dot{t}(\tilde{\theta}) = \theta \dot{q}(\tilde{\theta}) = 0 \tag{2}$$

Should this hold for all θ in the type space, it must be the case that

$$\dot{t}(\theta) = \theta \dot{q}(\theta) = 0, \tag{3}$$

or, in other words, for reporting any θ different from the true type, it is nonincreasing in his rent.

From the second-order condition,

$$\ddot{t}(\tilde{\theta})\Big|_{\tilde{\theta}=\theta} - \theta \ddot{q}(\tilde{\theta})\Big|_{\tilde{\theta}=\theta} \leq 0, \tag{4}$$

or

$$\ddot{t}(\theta) - \theta \ddot{q}(\theta) \leq 0. \tag{5}$$

Differentiating (3) permits to rewrite (5) into

$$-\dot{q}(\theta) \geq 0, \tag{5'}$$

where (3) and (5') ensure that the agent does not misreport his type locally.

In a next step, it can be shown that the relevant local constraint (3) also holds globally, that is, for all types besides adjacent types, one can use (3) to replace the R.H.S. of (1) by:

$$t(\theta) - t(\tilde{\theta}) = \int_{\tilde{\theta}}^{\theta} \tau \dot{q}(\tau) d\tau = \theta q(\theta) - \tilde{\theta} q(\tilde{\theta}) - \int_{\tilde{\theta}}^{\theta} q(\tau) d\tau \tag{6}$$

Isolating the terms without tilde and τ leads to:

$$t(\theta) - \theta q(\theta) = t(\tilde{\theta}) = t(\tilde{\theta}) - \theta q(\tilde{\theta}) + (\theta - \tilde{\theta})q(\tilde{\theta}) - \int_{\tilde{\theta}}^{\theta} q(\tau)d\tau \tag{7}$$

The term right of the plus sign is nonnegative, which implies that the local incentive constraint (2) holds for all θ.

We now use rent notation for setting up the optimal contract, with $U(\theta) = t(\theta) - \theta q(\theta)$. Substituting in (2), we have

$$\dot{U} = -q(\theta). \tag{IC}$$

The principal's optimization problem can now be expressed over rents and outputs in the following form for the contracting space $[\underline{\theta}, \hat{\theta}]$:

$$\underset{\{U(\cdot),q(\cdot)\}}{max} \int_{\underline{\theta}}^{\hat{\theta}} \Big(S(q(\theta)) - \theta q(\theta) - U(\theta) \Big) f(\theta) d\theta \tag{P}$$

s.t. (IC), a nondecreasing schedule of outputs (M) and a nonnegative rent for all agents (IR):

$$\dot{U} = -q(\theta). \tag{IC}$$

$$\dot{q}(\theta) \leq 0 \tag{M}$$

$$U(\theta) \geq 0. \tag{IR}$$

Because of (IC), (IR) simplifies to

$$U(\hat{\theta}) \geq 0, \tag{IR}$$

which, as in the discrete type case, becomes binding: the least efficient type of agent is given no rent. Because of the earlier stated monotone hazard rate property $\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right)$, condition (M) is always fulfilled by the optimal contract and will be slack in the optimum.

Solving (IC) yields

$$U(\hat{\theta}) - U(\theta) = -\int_{\theta}^{\tilde{\theta}} q(\tau)d\tau. \tag{8}$$

Last, because of (IR) $U(\theta) = 0$ and

$$U(\theta) = \int_{\theta}^{\tilde{\theta}} q(\tau)d\tau. \tag{9}$$

Replacing $U(\theta)$ in (P) by its R.H.S. permits us to express the principal's program in a reduced form:

$$\max_{\{U(\cdot),q(\cdot)\}} \int_{\underline{\theta}}^{\hat{\theta}} \Big(S(q(\theta)) - \theta q(\theta) - \int_{\theta}^{\tilde{\theta}} q(\tau)d\tau\Big) f(\theta)d\theta. \tag{P'}$$

Note that $\int_{\underline{\theta}}^{\hat{\theta}} \Big(\int_{\theta}^{\tilde{\theta}} q(\tau)d\tau\Big) f(\theta)d\theta$ is the expected rent of the agent of type θ, or $E(U(\theta))$.

Using integration by parts, this expression simplifies into

$$\int_{\underline{\theta}}^{\hat{\theta}} \left(\int_{\theta}^{\hat{\theta}} q(\tau)d\tau\right) f(\theta)d\theta = \int_{\underline{\theta}}^{\hat{\theta}} \left(\int_{\theta}^{\hat{\theta}} q(\tau)d\tau\right) dF(\theta)$$

$$\int_{\underline{\theta}}^{\hat{\theta}} \left(\int_{\theta}^{\hat{\theta}} q(\tau)d\tau\right) f(\theta)d\theta = \left(\int_{\theta}^{\hat{\theta}} q(\tau)d\tau\right) dF(\theta)\Big|_{\underline{\theta}}^{\hat{\theta}} - \int_{\underline{\theta}}^{\hat{\theta}} F(\theta)(-q(\theta))d\theta$$

$$\int_{\underline{\theta}}^{\hat{\theta}} \left(\int_{\theta}^{\hat{\theta}} q(\tau)d\tau\right) f(\theta)d\theta = \int_{\underline{\theta}}^{\hat{\theta}} F(\theta)q(\theta)d\theta,$$

which yields

$$U(\theta) = \int_{\underline{\theta}}^{\hat{\theta}} \left(\int_{\theta}^{\hat{\theta}} q(\tau)d\tau\right) f(\theta)d\theta = \int_{\underline{\theta}}^{\hat{\theta}} \frac{F(\theta)}{f(\theta)} q(\theta) f\theta d\theta.$$ [18]

[18]Note in particular that this holds because the density at the left margin must be zero, or $F(\theta) = 0$, and that the IC constraint was obtained by differentiating $\frac{d\left(\int_{\theta}^{\hat{\theta}} q(\tau)d\tau\right)}{d\theta} = -q(\theta)$.

6.2 Appendix 2: Proof of Proposition 5.

Re-expressing $E(V(\theta))$ from (11) into

$$\int_{\underline{\theta}}^{\hat{\theta}} U(\theta)d\theta - \int_{\theta^C}^{\hat{\theta}} U(\theta)d\theta - \int_{\underline{\theta}}^{\theta^C} U(\theta)d\theta = \int_{\underline{\theta}}^{\hat{\theta}} \frac{F(\theta)}{f(\theta)}q(\theta)f\theta d\theta - \int_{\theta^C}^{\hat{\theta}} \frac{F(\theta)}{f(\theta)}q(\theta)f\theta d\theta - \int_{\underline{\theta}}^{\theta^C} \frac{F(\theta)}{f(\theta)}q(\theta)f\theta d\theta$$

which gives the required result. ■