IFN Working Paper No. 1136, 2016

A Theory of Delegated Contracting

Wolfgang Gick

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

A Theory of Delegated Contracting

Wolfgang Gick*

IFN - Research Institute of Industrial Economics
Grevgatan 34 - 2 fl
Box 55665, SE-102 15 Stockholm
wolfgang.gick@ifn.se

Free University of Bozen/Bolzano
School of Economics
Universitätsplatz 1
I-39100 Bozen, Italy
wolfgang.gick@unibz.it

Abstract

Delegated contracting describes a widely observable agency mode where a top principal, who has no direct access to a productive downstream agent, hires an intermediary to forward a sub-contract with specified output targets and payments. The principal makes the payment to the intermediary contingent on production taking place; the intermediary is protected by limited liability and paid a bonus.

I characterize the optimal grand-contract with a continuum of agent types by using optimal control techniques with a scrap value function. Delegation proofness is reached through paying the intermediary what she could obtain by deviating. This rent is shown to be convex and increasing in the contracting space. There is internal verification of the ex-post state to reach compliance. The principal uses cutoff structures instead of additional output distortions. A leftbound incentive alignment principle between principal and intermediary applies. The paper so delivers a general analysis of the loss of control in vertical hierarchies.

JEL classification: D23, D73, L51.

Keywords: Delegated contracting, vertical hierarchies, internal verification, adverse selection, limited liability, cutoff structures, leftbound incentive alignment.

*I thank Philippe Aghion, Jesse Bull, Jacques Crémer, Bob Gibbons, Thomas Gresik, Werner Güth, Florian Herold, Friederike Hertweck, Yuriy Kaniovskyi, Dilip Mookherjee, Andrea Prat, Marco Sahm, David Sappington and Philipp Weinschenk for helpful suggestions, furthermore participants at the Econometric Society European Meeting 2012, the 2011 IIOC Conference in Boston, the 2011 International Game Theory Conference at Stony Brook, the Midwest Fall 2010 Economic Theory Meetings, the 2010 MPI Jena Spring Strategy Workshop, department seminars at Harvard, IFN Stockholm, LMU Munich, FAU Erlangen-Nuremberg, Free University of Bozen/Bolzano, Georgia Institute of Technology, and Florida International University. Particular thanks to Jean Tirole for many early discussions and to Thomas Tangerås for recent suggestions that helped bringing this paper into its current form. All errors are mine. Long-term research support by the Center for European Studies at Harvard University is gratefully acknowledged.

1 Introduction

Successfully managing global value chains requires the proper design of contracts with manufacturers abroad. To carry out production of a specific component, a multinational firm's top management will typically appoint an internal division manager (intermediary) who possesses some skills for the task at hand such as the capability to detect some possibly unwanted types of manufacturers.

Endowing an intermediary with sub-contracting power remains a double-sided sword. Even when budgeting is used, the optimal design of such an agency mode is far from trivial. It is known from the literature on cost accounting that vertical hierarchies are plagued by suboptimal decision making practices that follow from delegation.[1] This property is inherent to many vertical hierarchies: on one side, the intermediary's access to information occurs naturally when approaching the agent with a contract offer – one of the reasons why the principal may want to hire her. On the other, delegating this task makes it difficult to control her subsequent actions. Information itself is difficult to verify for the principal, and even the use of internal control and auditing schemes will not fully restore incentives.[2] At the end, the principal may well be left without production of the component if the contract is not properly designed.

The value chain example carries over to many other real-world agencies. Consider a public agency with an internally located bureaucrat who is commissioned to offer a pre-defined contract to a firm. Clerks are typically endowed with *some* knowledge about possible contractors' costs. Moreover, public agencies are required to use internal verification schemes to control the bureaucrat or its division in charge of the task, at the same time insuring the clerk through paying her a wage independent of the level of production realized *ex post*.

[1] See Horngren et al. (2003), Gjesdal (1981) as well as case studies such as HBS (1992 and 2000).

[2] In Gick (2008) I have ventured into an analysis of internal control to limit such discretion, building on the top principal's option to verify the intermediary's contract offer at some cost.

Corresponding research to the present paper is the study of optimal auction design when participation is costly and endogenously determined. There, an auctioneer may e.g. have discretion to cut off a specific subset of bidder types. This makes a higher knock-down price more likely, permitting the auctioneer to require a higher payment for carrying out the auction, which the owner has to pay for not being left with the item unsold.[3]

Taking a step back reveals how closely the problem at hand relates to the classical incentives and control theme already tackled in Holmström's early (1977) work. Like Williamson (1976) and Calvo and Wellisz (1987), Holmström considers a moral hazard setup, yet connecting to screening contracts when comparing decentralization and centralization with an eye on optimal control structures. While different in detail, his two-dimensional control problem is akin to the setup with verification used in the present paper.[4]

In the narrow sense, the literature on delegated contracting is best categorized in Mookherjee's (2006) concept of the "Delegation to the Middleman (DM)" , which in his description comprises two distinct papers: one on intermediated contracting by Faure-Grimaud and Martimort (2001, FGM hereafter) and one on organizational diseconomies of scale, written by McAfee and McMillan (1995). Neither of these papers offers a clear-cut overall advantage for having an intermediary in the regime. As in the present paper, they also assume that the top principal cannot perform the task himself, looking for the optimal contract to limit the additionally emerging agency costs for having an intermediary carrying out the sub-contract.

To briefly give an intuition why my paper reaches different results it should suffice to motivate upfront that in FGM, the contracting space for the agent is binary,[5] while my setup uses a continuous type space for the agent to study the principal's

[3]See Celik and Yilankaya (2009), as well as Menezes and Monteiro (2000) on auction design under costly participation. I expand on the parallel contributions of this literature in Section 3.

[4]See Holmström's (1977) Ph.D. thesis, ch. 2.4 as well as Holmström (1984), issues to which I further refer in Section 3.

[5]Throughout the paper, I refer to FGM when speaking of the binary case.

options to limit output distortions. Due to its assumptions, FGM prevents itself from studying cutoff structures. The two types envisaged in their contract are fixed and the sub-contract needs to be carried out with exactly two types, with the intermediary's rent being a function of the marginal cost difference of the two existing types of agents that carry out production.

My paper also differs from McAfee and McMillan (1995). While they do encompass a setup with a continuum of agent types, their conclusion is that more layers always imply more distortions. In their paper, costs consist of both production costs and information rents; yet, they exclude a "measure-zero case" where marginal information costs in fact do not stack up along the hierarchy.[6]

My first result is that in an agency setup with internal verification the emerging fixed costs lead to a grand-contract with cutoff structures. It is already known from other strands of the contracting literature that models with continuous types lead to equilibrium outcomes that are different from those with discrete types. [7] To my knowledge, delegated contracting so far has not been studied in such an extended framework. It is the purpose of this paper to characterize the optimal cutoff structures following from participation constraints. Fixed costs emerge naturally in agencies when an intermediary is present, and my first result shows that the optimal contract does not entail further output distortions over second-best. Related to this result is the emergence of bunching over an interval close to the most *inefficient* type of agent in the contracting space.

A second result is the emerging information rent that directly follows from the internally limited options of the intermediary to deviate. This deviation is directional in that it is generally optimal for her to threaten the principal to cut off a set of rightbound agent types, a result following from a leftbound incentive alignment that

[6]See McAfee and McMillan (1995, p. 413). There, the authors explicitly distinguish between these two costs, adding the notion of "marginal virtual costs", with the result that in all but one case there are additional output distortions. My argument is different from theirs.

[7]See Armstrong and Sappington (2007), who point out that for multi-dimensional private information, the qualitative properties of the optimal regulatory policy may change significantly when the agent's private information becomes continuous. See also Armstrong (1999).

exists between principal and intermediary.

A third result is the relative optimality of using a bonus payment together with pursuing internal verification. Paying the intermediary a bonus is a natural way to design incentives, a payment scheme that lines up well with the use of verification adopted in this paper as the existence of a bonus eases compliance by the intermediary in an intuitive way. My view thus supports forms of real-world delegation contracting modes that encompass only a subset of more efficient input producers (or those that are of higher quality) compared to the setup without an intermediary.

I organize the remainder of my paper as follows. The following section presents the model and specifies the intermediary's *ex-ante* rent and his participation constraint as a function of the agent's cost type as well as an internal verification scheme. Section three offers a detailed discussion of the literature, Section four concludes. Proofs are given in the appendix.

2 The model

2.1 Primitives

There is a top principal $\mathcal{P}$ (he), an intermediary $\mathcal{I}$ (she), and a downstream agent $\mathcal{A}$. All players are risk-neutral, and the intermediary is protected by limited liability below zero wealth. The agent has a utility function $U(q(\theta), t(\theta), \theta)$, in its functional form written linearly as $U = t - \theta q$, where t is the monetary transfer he receives from the intermediary to produce an output target of q, and θ is the agent's privately known marginal costs or simply his type. The agent knows his type when accepting the contract and has a reservation utility normalized to zero. As is typical for the contracting literature,[8] I assume a concave surplus function for $\mathcal{P}$ of $S(q)$. Both the intermediary and top principal know the agent's type distribution, which has a continuum density and a strictly positive type set of $\Theta = [\underline{\theta}, \bar{\theta}]$. Monotone hazard rate

[8]See e.g. Laffont and Martimort (2002).

condition holds, with $\frac{d}{d\theta}\left(\frac{F(\theta)}{f(\theta)}\right) \geq 0$. $F(\theta)$ is the c.d.f. and $f(\theta)$ the p.d.f. . The distribution is well-defined and differentiable nearly everywhere. Besides the continuous types inside the contracting space, there exists one single and highly inefficient type of agent $\theta^{\infty} > \bar{\theta}$ who appears with a probability ε.[9] Should the intermediary not learn whether the agent is of type θ^{∞}, he will appear in Θ with probability $1 - \varepsilon$.[10]

2.2 Timing

The timeline of the contracting game reads as follows:

- $(t = 0)$ Agent learns its type θ. All players know that the intermediary will learn the agent's type at the sub-contract offer stage $(t = 3)$ if $\theta = \theta^{\infty}$.

- $(t = 1)$ $\mathcal{P}$ offers a delegation-proof grand-contract with internal verification to $\mathcal{I}$, specifying output targets and transfers to the agent as well as the intermediary's wage.

- $(t = 2)$ Intermediary accepts or rejects the grand-contract.

- $(t = 3)$ Intermediary offers a sub-contract.

- $(t = 4)$ Agent accepts or rejects.

- $(t = 5)$ Production and transfers take place: $\mathcal{A}$ is paid contingent on type-contingent output being delivered, $\mathcal{I}$ receives her wage, and $\mathcal{P}$ receives the output. Game ends.

[9] As in FGM, this type is just chosen high enough, with $\theta^{\infty} << \infty$ such contracting is too costly for the principal.

[10] See FGM.

2.3 Constraints and rents

The intermediary is contracted *ex ante*, that is, neither principal nor intermediary know the type of agent at the time the grand-contract is offered. However, both know the agent's type distribution. The intermediary is fully protected by limited liability. In addition, she may deviate from the sub-contract at the time of contracting as this is not observable. As in FGM, deviations by the intermediary are limited: I exclude random contracts. Moreover, type reports cannot be forged; consequently output targets and transfers in terms of the sub-contract, that is, $\{t(\theta), q(\theta)\}$ are those designed by the principal when a sub-contract is carried out. The proof for the principal that the intermediary did his job is the fact that the agent accepted and output is produced and delivered to the principal.

To perform the rent analysis, I proceed as in McAfee and McMillan (1995) by starting with the standard two-player contract under adverse selection. Setting this problem up in standard textbook fashion yields

$$\max_{\{(U(\cdot),q(\cdot))\}} \int_{\underline{\theta}}^{\bar{\theta}} \left(S(q(\theta)) - \theta q(\theta) - U(\theta)\right) f(\theta) d\theta.$$

I now consider possible deviations. When forwarding the correct production targets and transfers to all types, she can capture parts of the virtual cost designed to be forwarded to the agent, namely $U(\theta)$ herself by deviating from the pre-designed subcontract. In FGM, this deviation is reached by a "gamble" through which, in case of appearance of the discrete $\underline{\theta}$-type, the intermediary would pocket the (discrete-type) virtual costs of $U(\underline{\theta}) = \Delta\theta\bar{q}$ with a given probability ν, whenever the type is not the most inefficient one.

In my paper, to forward the screening contract to all interior types between $\underline{\theta}$ and $\bar{\theta}$, the principal designs a budget s from which the intermediary pays a transfer t to the agent and verification occurs *ex post*. As a helpful illustration, consider the budget

s being set high enough to cover the contractual costs for the most efficient type $\underline{\theta}$, that is the real costs of output θq plus the virtual costs that are required to invoke the revelation principle for the screening contract. Whenever the intermediary reaches a deviation from the contract, in that $s - t > 0$ or simply $t < s$, she earns a positive rent. To set up the principal's problem I have

$$\max_{\{\cdot\}}(1-\varepsilon)\int_{\underline{\theta}}^{\overline{\theta}} \left(\left(S(q(\theta) - \theta q(\theta) - U(\theta)\right)dF(\theta) + \varepsilon\left(s(q^{\infty}) - t(\theta^{\infty})\right)\right)$$

where $S(q(\theta)) - \theta q(\theta)$ is the difference between the principal's surplus and the true production costs for each type θ. By the very nature of the setup, the condition that no production occurs for the θ^{∞}-type are trivially fulfilled. Still, the intermediary, at $t = 3$, knowing the virtual costs for each agent in the contracting space, will only forward the correct optimal screening contract to all agent types in the contracting space when she is being paid a rent equivalent to her highest possible deviation. This rent is limited and endogenously determined by the type space and the virtual costs. While at $t = 2$, the intermediary has no additional information about types in the contracting interval compared to the principal himself, she can always obtain a positive rent as long as the type space is nonzero. The intermediary so derives a rent that is increasing and strictly convex in the type distribution.[11] In what follows, I determine the intermediary's equilibrium rent $v^{DP}(\bar{\theta})$ to derive her participation constraint (IPC) that satisfies delegation proofness.

[11]Since it is generally assumed that the leftmost type $\underline{\theta}$ is fixed, I henceforth will, with a slight abuse of notation, denote this function as being depending on the highest type $\bar{\theta}$, which is determined by the principal in equilibrium.

2.4 Delegation Proofness

Proposition 1 Intermediary's space-dependent preferences. *When performing his task correctly, $\mathcal{I}$ is paid a fixed equilibrium rent $v^*(\bar{\theta})$ found at the maximum value of θ^{C*} that yields the highest rent given $\bar{\theta}$. The maximum expected equilibrium rent that $\mathcal{I}$ can obtain by cutting off a subset of type is found as follows:*

$$v^{DP} = \underset{\theta^C}{argmax}\Big(\int_{\underline{\theta}}^{\bar{\theta}}\Big(\int_{\theta}^{\bar{\theta}} q(\tau)d\tau\Big)f(\theta)d\theta - \int_{\theta^{C*}}^{\bar{\theta}}\Big(\int_{\theta}^{\bar{\theta}} q(\tau)d\tau\Big)f(\theta)d\theta - \int_{\underline{\theta}}^{\theta^{C*}}\Big(\int_{\theta}^{\theta^{C*}} q(\tau)d\tau\Big)f(\theta)d\theta\Big).$$

While the intermediary may use cutoffs to modify the type distribution to her own ends, the deviation itself is endogenously limited as a higher rent gained through more types being cut off comes with a lower probability of having an agent accept the contract.

To illustrate this trade-off, consider *Figure 1* below. The first term of the *IPC* constraint is $E(U(\theta))$, corresponding to the entire triangle in the graph. The second term is $E(U(\theta))|_{\theta^C}^{\bar{\theta}}$, which corresponds to the triangle on the right. The third and last part is the rent paid to the downstream agent or $E(U(\theta))|_{\underline{\theta}}^{\theta^C}$ should only types up to θ^C be offered a contract, which is exactly her deviation. It corresponds to the blue triangle. Note that all three parts of the sum are positive for any $\underline{\theta} < \theta^C < \bar{\theta}$. For any well-behaved type distribution Θ there is always an interior solution.[12]

Proposition 2 *The intermediary can only obtain a rent by cutting off a rightbound set of agent types.*

Corrollary 1. Leftbound Incentive Alignment Principle. *Incentives between $\mathcal{P}$ and $\mathcal{I}$ are aligned at the left end of the type distribution: it is much less costly for the principal to reach delegation proofness for a leftbound subset of the type space than for the entire type space.*

---

[12]I have $\frac{d\left(\int_{\theta}^{\bar{\theta}} q(\tau)d\tau\right)}{d\theta} = -q(\theta)$, and the expected rent is $\int_{\underline{\theta}}^{\bar{\theta}}\Big(\int_{\theta}^{\bar{\theta}} q(\tau)d\tau\Big)f(\theta)d\theta = E(U(\theta))$.

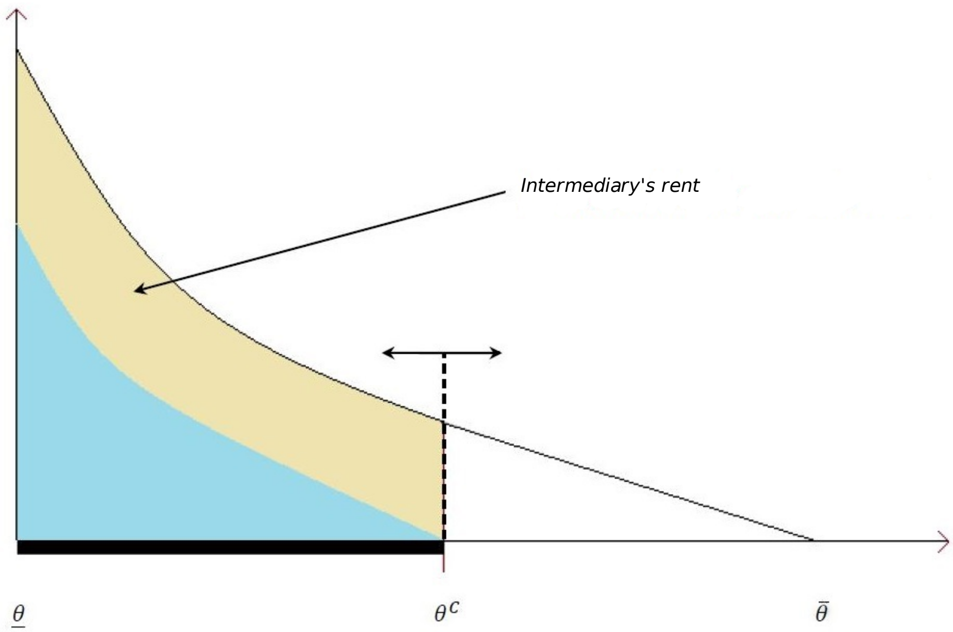

Figure 1: *Determining the rent payment $v^{DP}(\theta^C)$ for delegation proofness.*

2.5 Internal verification

Note that this setup encompasses full insurance for the intermediary for production states. She is paid the maximum rent she could reap by deviating as shown in the previous subsection, and receives this amount as a bonus payment whenever production takes. Note however that without further verification there is no guarantee she would definitely comply with the rules set by the principal to forward the pre-designed screening sub-contract. It is possible that the intermediary could, even while paid to not deviate, exactly pursue the same deviation to get an extra rent, possibly of the same size of the bonus payment. Even if one excludes such outright deception, she could still engage in other, secondary deviations.

More specifically, the intermediary could increase her rent by e.g. offering a one-size-fits-all contract, which is tailored to meet partipation of some interior type, say θ^C, which would make this type break even while offering a strictly positive payoff to all more efficient types that of course would accept. With this deviation, the intermediary could claim to have paid the virtual costs to enable acceptance for the θ^C-type, while in fact cutting off all rightbound types. In this way, the intermediary could capture additional, though smaller, rents.

FGM's model is not affected by this deviation, thus constituting a costly verification model itself, as a benefit of reducing the contracting space to two types.[13] Binary models in fact limit the rent for the intermediary to the virtual costs of the screening contract, there uniquely $\Delta\theta\bar{q}$. This is easy to see as in FGM there are only two production targets: $\underline{q}$ and $\bar{q}$. A realized output of $\bar{q}$ would then indicate that either a screening contract was offered correctly, or that the intermediary could have deviated by offering a one-size-quantity contract, which would be accepted by both the $\underline{\theta}$ and $\overline{\theta}$ type of agents. In FGM, this secondary deviation would not lead to any additional rent for $\mathcal{I}$.

Instead, the model presented here needs state verification at the end of the contracting game. Such a form of internal control is generally found in agencies. It is well known that multi-divisional firms regularly gather information about their subunits, and public agencies are required by law to permit access to accounting offices that perform verification.[14].

In what follows, I adopt a simple verification scheme that fits the setup with an intermediary. Verification involves costs of size $\mathcal{V}$ that are independent of the result detected. Similar to the regulatory setup in Baron (1984) I assume that in case of a detected deviation the principal can inflict the verification costs on the intermediary and punish her by cancelling the bonus payment. I so define an indicator function $\Im$ as follows:

$$\Im(\Theta, \Theta^P) = \begin{cases} 1 & \text{if } \Theta = \Theta^P \\ 0 & \text{else,} \end{cases}$$

where Θ^P stands for the sub-contracting space prescribed by $\mathcal{P}$.

Truthfully forwarding the sub-contract according to the principal's design is an

[13]Relatedly, and based on this property, there is scope in FGM there is scope for an internal control scheme, which I have studied in Gick (2008). Because of the binary structure, auditing is optimally carried out only if one state is reported.

[14]See e.g. GAO (2011)

equilibrium as long as

$$(1-\wp)v^{DP} - \wp \cdot \mathcal{V} \geq 0 \tag{CV}$$

where (CV) denotes compliance under verification. Rearranging shows that the intermediary forwards the contract truthfully as long as $v^{DP} \geq \frac{\wp}{1-\wp} \cdot \mathcal{V}$. $\mathcal{P}$ bears the fixed costs of verification $\mathcal{V}$. In equilibrium, $\mathcal{I}$ chooses to not deviate and receives the bonus payment as long as (CV) is fulfilled. No additional output distortions occur for types in the contracting space.[15]

The principal commits to always perform verification. The following equilibrium exists.[16]

$$\max_{\{q(\theta),U(\theta),\bar{\theta},\mathcal{V},\phi\}} (1-\wp)\int_{\underline{\theta}}^{\bar{\theta}}\Big(S(q(\theta)) - \theta q(\theta) - U(\theta)\Big)f(\theta)d\theta +$$
$$\wp\Big(\int_{\underline{\theta}}^{\bar{\theta}}\Big(S(q(\theta)) - \theta q(\theta) - U(\theta)\Big)f(\theta)d\theta\Big) - v^{DP} - \mathcal{V}(1 - \Im(\Theta, \Theta^P)),$$

2.6 Grand-contract

Setting up the optimal control problem with a scrap value function $\phi(q(\bar{\theta}^*), \bar{\theta}^*)$ leads to the following expression:

$$\max_{\{q(\theta),U(\theta),\bar{\theta},\mathcal{V},\phi\}} (1-\wp)\Big(\int_{\underline{\theta}}^{\bar{\theta}}\Big(S(q(\theta)) - \theta q(\theta) - U(\theta)\Big)f(\theta)d\theta +$$
$$\wp\Big(\int_{\underline{\theta}}^{\bar{\theta}}\Big(S(q(\theta)) - \theta q(\theta) - U(\theta)\Big)f(\theta)d\theta\Big) - \phi(q(\bar{\theta}^*), \bar{\theta}^*) - V,$$

With (IPC) binding and $\wp$ given, satisfying (CV) and limited liability, the grand-contract is characterized as follows.

[15]For simplicity and without loss of generality I assume that there exists a verification technology that delivers the required probability $\wp$ by paying costs of $\mathcal{V}$. For further discussions that compare verification costs with court costs see Baron (1984). There is no limit to extend the setup, in the sense of Baron and Besanko (1984), to comprise a convex cost function of verification, with $\wp$ becoming part of the maximization problem, with the marginal cost of verification being set equal to $\wp$ in equilibrium.

[16]This all comes with a slight abuse of notation, as the ex-ante probability for the contract to happen is $(1-\varepsilon)$ in equilibrium. Trivially, as the designed output target to the extremely inefficient type is zero, the surplus for that state is zero as well, with a probability of ε, and the intermediary receives no bonus payment.

$$\max_{\{q(\theta),U(\theta),\bar{\theta},\mathcal{V},\phi\}} \int_{\underline{\theta}}^{\bar{\theta}} \Big(S(q(\theta)) - \theta q(\theta) - U(\theta) \Big) f(\theta) d\theta - \phi(q(\bar{\theta}^*), \bar{\theta}^*) - \mathcal{V},$$

s.t.

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

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

$$U \geq 0 \text{ for all } \theta \text{ in } \Theta,$$

$$\dot{q}(\theta) = y(\theta) \tag{ME}$$

and

$$y(\theta) \leq 0.$$

This leads to the following proposition.

Proposition 3 GRAND-CONTRACT. *The grand-contract entails the following results:*

• *delegation proofness applies and the intermediary offers the sub-contract truthfully,*

• *there is some bunching left of the optimal cutoff* $\bar{\theta}^*$*,*

• *no additional output distortions apply compared to second-best.*

3 Contribution to the literature

With the grand-contract exhibiting cutoff structures, my paper differs in several aspects from the literature. In the narrower sense, my paper generalizes the setup by permitting the intermediary to require a payment based on the type distribution. Given the distribution, she may threaten the top principal by cutting off a right-bound set of agent types. As I have shown, this may, under the given structures, enable the top principal to refrain from the use of additional output distortions for agent types.

In turn, FGM has an intermediary who is faced with only two types in the contracting space, and these types cannot be altered. Her informational advantage follows her discretion to possibly *not* include a fixed intermediate cost type in the sub-contract. The intermediary can therefore play a gamble to offer a contract to the most efficient type only. When successful, she obtains a rent based on the entire type difference times the output target for the intermediate type (that is, the virtual costs $\Delta\theta\bar{q}$, in FGM notation). In other words, by offering a shut-down contract to only one, namely the most efficient type, she may be lucky to obtain a rent of $\nu\Delta\theta\bar{q}$, where ν represents the probability to encounter the most efficient agent, contingent on no type being outside the contracting space. In the FGM setup, the principal will third-best reduce $\bar{q}$, which corresponds to the mainstream argument that adding more layers leads to higher agency costs and adds more distortions.

While the intermediary has no productive task, FGM's result hinges on a strict definition of the intermediary's risk-averse preferences, based on $\Delta\theta\bar{q}$. When the intermediary gradually becomes risk neutral, she has more incentives to gamble. In turn, when her degree of risk aversion would increase, she would be paid the same rent in all states, converging to my setup with continuous agent types. However, for very conceptual reasons, I do not express this through a risk-averse utility function such as used for a productive agent. Instead, I consider the intermediary an intermediated principal in the sense of McAfee and McMillan (1995). I reach a strictly positive rent to be paid to her, a rent that strictly exceeds what she could solely require from limited liability. The result is a state-independent bonus payment that the intermediary is paid when production takes place.[17]

Another result of my paper is that once the intermediary has freedom to modify a type distribution to her own ends, she obtains a rent that is convex and increasing in the type span, but the deviation itself shows an endogenous limitation as a higher rent gained through a given cutoff structure is traded off by a lower probability of having an agent accept in equilibrium. Given this property, my paper differs clearly from McAfee and McMillan (1995) as the double marginalization problem turns out to disappear in equilibrium. Yet, there is some bunching close to the optimal right-bound cutoff $\bar{\theta}^*$.[18]

Within the field of auction design, work by Chen (2013) constitutes the closest parallel to my paper. Chen assumes that both sides incur participation costs. There, the mediator delegates the auction design to the seller, with the mediator-optimal mechanism exhibiting cutoff structures as in the present paper. In both my and Chen's paper there is too little ex-ante participation in equilibrium: here, the top principal would like production to take place with even more inefficient types of agents could he perform the task himself, while in Chen's (2013) paper the seller expects to attract fewer bidders than in a seller-optimal auction.

In the broader sense, my paper also relates to the literature on endogenously emerging participation constraints as amply discussed in Jullien (2000). My three-player model necessarily displays countervailing incentives, a feature studied in famous work by Lewis and Sappington (1989) as well as Maggi and Rodriguez (1995a,b), where participation constraints may become binding at an interior type, a

[17]See also Kim (1997) on the optimality of bonus contracts under limited liability.

[18]While different in detail, my model is not unrelated to Celik's (2008) concept of a counter-marginalization of information rents, by which a top principal can offset additional distortions that result from the general leftboundedness of incentives. In Celik's paper, there is an insurer, and distortions are offset by creating an incentive reversal. In my paper, the intermediary requires full insurance, which makes to cutoff optimal.

parallel to the IPC constraint in my paper. It was, notably, the merit of Jullien (2000) to provide a generalization for such setups. In particular, Jullien has argued that participation constraints in the sense of a “forgone opportunity” may involve non-monotonic information rents, precisely when speaking of fixed trading costs. Such costs are akin to the agency costs that occur once an intermediary is in the regime. In the present paper, it is therefore the intermediary who has an incentive to overstate the agent’s type to so increase her rent, with the top principal responding by mitigating this rent through imposing a specific cutoff structure.

While closing this specific lacuna in the field of vertical hierarchies, my paper also opens up new avenues for future research. So far, the literature does not explicitly study delegation together with contract bargaining. It could be rewarding to study issues like bargaining power and how this power varies with different contracting setups under limited liability and/or risk aversion. For example, Inderst (2002) shows in an adverse selection framework where the agent may play with some rents, that contractual distortions vanish as the bargaining power of the agent *increases*, a property in line with my rent construction for the intermediary. The more general view traces back to Pitchford (1998) who shows that the distribution of bargaining power clearly has an impact on the optimal contract. Similar models are found in Demougin and Helm (2006), a paper comparing three different but equivalent bargaining scenarios between one principal and one agent, using the standard screening contract that invokes the revelation principle as a benchmark for comparison. Li et al. (2013) study contract bargaining under moral hazard and risk aversion.

A last direction of research that encompasses mechanisms without invoking the revelation principle is work on optimal report management. Ronen an Yaari (2003) depart from the revelation principle as well, presenting a theory of optimal contracts with partial suppression of messages.

4 Concluding remarks

This paper has offered a new framework for the analysis of delegated contracting. My main result is that this agency mode permits an overall very intuitive way to respond to increased agency costs that emerge from having and intermediary in the regime. In my general setup that encompasses a continuum of agent types, the top principal has the option to make use of cutoff structures to restore delegation proofness, an instrument quite different from the literature so far.

By departing from a design that responds to higher rents by imposing additional output distortions, my paper draws a new picture on how agencies handle the loss of control. This handling in fact corresponds to the way of how multidimensional firms use internal verification schemes and bonus payments. As the intermediary's options to reap a rent are endogenously limited by the type span, the principal will optimally recur to a contract design that makes use of the existing joint (leftbound) alignment of incentives between principal and intermediary, resulting in optimal cutoff structures.

It seems worthwhile to point out that my generalized framework differs to a large extent from the literature. With its conclusion to not impose additional output distortions to mitigate the loss of control but instead by a reduced contracting space, it is therefore much closer in spirit to Williamson's (1967) and Calvo and Wellisz' (1978) findings, namely that agency costs are driven by the "span of control", which in their setup originates from the range of *tasks* given to the intermediary. The larger this range, the larger is the loss of control. My setup with a continuum of agent types is akin to their view. As I have shown, the so emerging agency costs do, under internal verification, not carry over into output distortions. In this way, I give a more complete picture of delegated contracting. It is my hope that by bringing such properties to the attention of the reader will enhance our understanding of this specific agency mode.

Appendix

Proof of Proposition 1.

Consider a type space with $\underline{\theta} < \theta_1 < \theta^C < \theta_2 < \bar{\theta}$ and the given conditional beliefs. The proof comes in two parts.

• Assume first the intermediary cuts off a leftbound set of types, offering the sub-contract only to the types in $\{\theta_1, \bar{\theta}\}$. If the contract is accepted, the intermediary has to pay the corresponding rent. If not, whenever the type that occurs is in $\{\underline{\theta}, \theta_1\}$, no contract is accepted. Still, for any type in $\{\theta_1, \bar{\theta}\}$, the intermediary needs to pay the full information rent. While the forwarded rent is the same for any type who would accept a contract, the intermediary cannot gain from this action. Would she threaten the principal to forward the contract only to rightbound types $\{\theta_1, \bar{\theta}\}$, she could not reap any rent.

• Assume instead that the intermediary cuts off a rightbound set of types, offering the sub-contract only to the types in $\{\underline{\theta}, \theta_2\}$. This is a credible threat: not offering the contract to types in the sub-interval $\{\theta_2, \bar{\theta}\}$ permits her to save on information rent included in the sub-contract for any type in $\{\underline{\theta}, \theta_2\}$. ■

Sketch of Proof of Proposition 2. Proposition 2 follows directly from the two conditions in the setup for the standard screening model, namely $\dot{U}(\theta) = -q(\theta)$ and $\dot{q}(\theta) \geq 0$. ■

Proof of Proposition 3.

State variable is q and y is control. This leads to a Hamiltonian of

$$\mathcal{H} = \Big(S(q) - \Big(\theta + \frac{F(\theta)}{f(\theta)}\Big)q\Big)f(\theta) + \mu y.$$

There are two transversality conditions:

$$\mu(\bar{\theta}^*) = \frac{\partial \phi(q(\bar{\theta}^*), \bar{\theta}^*)}{\partial q(\theta)}. \tag{TC1}$$

The second transversality condition reads

$$H[q^*, \mu^*, y^*, \dot{\mu}] - \mu(\bar{\theta}^*) - \frac{\partial \phi(q(\bar{\theta}^*), \bar{\theta}^*)}{\partial \theta} = 0, \tag{TC2}$$

or

$$\Big(S(q(\bar{\theta}^*)) - (\bar{\theta}^* + \frac{F(\bar{\theta}^*)}{f(\bar{\theta}^*)})q(\bar{\theta}^*)\Big)f(\theta) + \mu(\bar{\theta}^*)y(\bar{\theta}^*) - \frac{\partial \phi(q(\bar{\theta}^*), \bar{\theta}^*)}{\partial \theta} = 0.$$

Using the Pontryagin principle I get

$$\dot{\mu}(\theta) = -\frac{\partial H}{\partial q} = -\Big(S'(q(\bar{\theta})) - \Big(\theta + \frac{F(\theta)}{f(\theta)}\Big)\Big)f(\theta).$$

With (TC1) positive, I have $y = 0$ and $\dot{q} = 0$. As long as $\dot{q} = 0$, it follows from complementary slackness that $\mu = 0$.

Complementary slackness requires that

$$\frac{\partial \mathcal{H}}{\partial y} \geq 0; \quad y \leq 0 \quad \frac{\partial \mathcal{H}}{\partial y} \cdot y = 0.$$

Whenever $y = 0$, it follows that $\frac{\partial \mathcal{H}}{\partial y} = \mu = 0$. Whenever q is decreasing, I must have that $\mu = 0$. But then, $\dot{\mu} = 0$ and $S' = \theta + \frac{F(\theta)}{f(\theta)}$. It is also known that $\mu(\bar{\theta}^*) > 0$. Thus, before $\bar{\theta}^*$ is reached, $\dot{\mu}$ must remain positive. I get $-[S'(\theta + \frac{F(\theta)}{f(\theta)}]f(\theta) > 0$. Maximizing w.r.t. to $q(\cdot)$ and using $y(\theta) \leq 0$ yields $\mu(\theta) \geq 0$, with $y(\theta) = 0$ for $\mu(\theta) > 0$. At $\bar{\theta}^*, \mu = 0$, which coincides with type $\bar{\theta}^*$ not being paid any information rent, and $y = 0$.

Solving leads to

$$\pi(\bar{\theta}) - \frac{\partial\phi(q(\bar{\theta}), \bar{\theta}^*)}{\partial(\bar(\theta)} \geq 0; \; q(\underline{\theta}) - q(\bar{\theta}) \geq 0; \; \Big(\pi(\bar{\theta}) - \frac{\partial\phi(q(\bar{\theta}), \bar{\theta}^*)}{\partial(\bar{\theta})}\Big)(q(\underline{\theta}) - q(\bar{\theta})) = 0.$$

where $\pi(\bar{\theta}) = \frac{\partial\phi}{\partial q(\bar{\theta})}$.

The result has an optimal value $\bar{\theta}^*$ that cuts off leaving exactly the scrap value equal to the intermediary's participation constraint $v(\bar{\theta}^*)$. Whenever $q(\underline{\theta})$ is free, the initial value is free and $\mu(\underline{\theta}) = 0$ and constant, with $q(\underline{\theta})$ falling. In addition, $\dot{\mu}(\underline{\theta}) = 0$ implies that $S'(q(\underline{\theta})) = 0$, so there is no distortion at the top.[19] ■

[19]See also Seierstad and Sydsaeter (1977 and 1987). A web appendix that contains additional technical details is in preparation.

Bibliography

Armstrong, M., (1999). Optimal Regulation with Unknown Demand and Cost Functions, in: *Journal of Economic Theory, 84* (2), 196-215.

Armstrong, M., Sappington, D.E.M., (2007). Recent Developments in the Theory of Regulation, in: Armstrong, M., and Porter, R. (eds.) *Handbook of Industrial Organization,* Amsterdam (North Holland) vol. 3, ch. 27, pp. 1557-1700.

Baron, D.P., (1984). Regulatory Strategies Under Asymmetric Information, in: Boyer, M., and Kihlstrom, R. (eds.) *Bayesian Models in Economic Theory,* Amsterdam (North Holland), pp. 155-180.

Baron, D.P., Besanko, D., (1984). Regulation, asymmetric information, and auditing, *RAND Journal of Economics, 15* (4), 447–470.

Calvo, G.A., Wellisz, S., (1978). Supervision, Loss of Control, and the Optimum Size of the Firm, in: *Journal of Political Economy 86*(5), 943-952.

Cassar, G., Gibson, B., (2008). Budgets, Internal Reports, and Manager Forecast Accuracy, in: *Contemporary Accounting Research 25*(3), 707–737.

Chen, Y.-J., (2013). Optimal mediated auctions with endogenous participation, in: *Decision Support Systems 54* (3), 1302–1315.

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

Celik, G., Yilankaya, O., (2009). Optimal Auctions with Simultaneous and Costly Participation, in: *B.E. Journal of Theoretical Economics (Advances) 9(1)*, Article 24.

Demougin, D. M., and Garvie, D. A., (1991). Contract design with correlated information under limited liability, in: *RAND Journal of Economics 22*, 477-489.

Demougin, D. M., and Helm, D., (2006). Moral Hazard and Bargaining Power, in: *German Economic Review 7* (4), 463-470.

Faure-Grimaud, A., Martimort, D., (2001). On Some Agency Costs of Intermediated Contracting, in: *Economics Letters 71*, 75-82.

GAO (United States Government Accountability Office) (2011). *Government Auditing Standards,* 2011 Revision, Washington, D.C., available at http://gao.gov/assets/590/587281.pdf

Gick, W. (2008). Delegated Contracting, Information, and Internal Control, in: *Economics Letters 101,* 179-83.

Gjesdal, F. (1981). Accounting for stewardship, in: *Journal of Accounting Research 19* (1), 208-231.

Harvard Business School (HBS) (1992). *The Lincoln Electric Company.* Harvard Business School Case 9-376-028.

Harvard Business School (HBS) (2000). *Aligning Incentives for Supply Chain Efficiency.* Harvard Business School Case 9-600-110.

Holmström, B. (1977). *On Incentives and Control in Organizations.* Unpublished Ph.D. thesis, Graduate School of Business, Stanford University.

Holmström, B. (1984). On the theory of delegation, in: Boyer, M., and Kihlstrom, R. (eds.) *Bayesian Models in Economic Theory,* Amsterdam (North Holland), pp. 115-141.

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

Inderst, R., (2002). Contract design and bargaining power, in: *Economics Letters 73,* 171-176.

Jullien, B., (2000). Participation Constraints in Adverse Selection Models, in: *Journal of Economic Theory 93,* 1-47.

Kim, S. K., (1997). Limited Liability and Bonus Contracts, in: *Journal of Economics and Management Strategy 6*(4), 899-913.

Laffont, J.-J., (1990). Analysis of Hidden Gaming in a Three-Level Hierarchy, in: *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.

Lewis, T., Sappington, D.E.M., (1989). Countervailing Incentives in Agency Problems, in: *Journal of Economic Theory 49,* 294-313.

Li, S., Hao, X., and Yao, D., (2013). Contract Bargaining with a Risk-Averse Agent, in: *B.E. Journal of Theoretical Economics (Contributions): 13(1)*, 285-301.

Maggi, G., Rodriguez-Clare, A., (1995a). On Countervailing Incentives, in: *Journal of Economic Theory 66,* 238-263.

Maggi, G., Rodriguez-Clare, A., (1995b). Costly distortion of information in agency problems, in: *RAND Journal of Economics 26,* 675-689.

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

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

Mookherjee, D., (2006). Delegation, Hierarchies, and Incentives: A Mechanism Design Perspective, in: *Journal of Economic Literature 44,* 367-390.

Pitchford, R., (1998). Moral hazard and limited liability: The real effects of contract bargaining, in: *Economics Letters 68*(2), 251-259.

Ronen, T., and Yaari, V., (2002). On the Tension between Revelation and Earnings Management: A Reconsideration of the Revelation Principle, in: *Journal of Accounting, Auditing & Finance* 17(4), 273-294.

Seierstad, A., Sydsaeter, K., (1977). Sufficient Conditions in Optimal Control Theory, in: *International Economic Review 18,* 367-391.

Seierstad, A., Sydsaeter, K., (1987). Optimal Control Theory with Economic Applications, Amsterdam: North Holland.

Williamson, O. E., (1967). Hierarchical Control and Optimum Firm Size, in: *Journal of Political Economy 75* (2), 123-138.