Controlling versus enabling*

Andrei Hagiu[†] and Julian Wright[‡]

May 23, 2016

Abstract

We study the optimal allocation of control rights over decisions that can be either made by a principal or an agent. The choice is determined by the need to balance double-sided moral hazard problems arising from investments that only the agent can make and investments that only the principal can make, while at the same time minimizing revenue-sharing distortions in the decisions that either party can make. We show control rights should all be given to the party that obtains the higher revenue share. We explore how changes in moral hazard, the contractibility of decisions, and spillovers across multiple agents affect this choice.

JEL classification: D4, L1, L5

Keywords: control rights, decision authority, employment, independent contractors, organizational theory.

1 Introduction

A key decision for many firms is whether to control the provision of services to customers by employing workers or whether to enable independent contractors to take control of service provision. This decision has been relevant in some industries for a long time—such as professional service firms and professionals, manufacturers and sales agents, and franchisors and franchisees. However, it has

*We thank Oliver Hart, Justin Johnson, Bruno Jullien, Marc Rysman, Steve Tadelis, Jean Tirole, Eric Van den Steen, Birger Wernerfelt, participants in the Platform Strategy Research Symposium at Boston University, the Platforms Workshop at the National University of Singapore, the Conference on the Economics of Intellectual Property, Software and the Internet at the Toulouse School of Economics, and participants in seminars at Boston University, the Federal Trade Commission, Georgetown University, Harvard University, MIT, the National University of Singapore, Northwestern University, University of California San Diego, and University of Southern California for their helpful comments. Andy Lee and Bo Shen provided excellent research assistance. Julian Wright gratefully acknowledges research funding from the Singapore Ministry of Education Academic Research Fund Tier 1 Grant No. R122000215-112.

[†]Harvard Business School, Boston, MA 02163, E-mail: ahagiu@hbs.edu

[‡]Department of Economics, National University of Singapore, Singapore 117570, E-mail: jwright@nus.edu.sg

become more prominent in recent times, reflecting that in a rapidly increasing number of service industries (e.g. consulting, education, home services, legal, outsourcing, staffing, and taxi services), online platforms have emerged to take advantage of information, communication, and remote collaboration technologies to enable independent professionals to directly connect with customers (e.g. Coursera, Gerson Lehrman Group, Hourly Nerd, Task Rabbit, Uber, and Upwork). These firms typically differ from their more traditional counterparts by letting professionals control some or all of the relevant decision rights, such as prices, expenditure on equipment, and promotion of the professionals' skills and services. This contrast motivates our theoretical study of a principal's choice of whether to keep decision authority or to grant it to an agent.

In these settings, the revenues generated by the principal typically depend on both its ongoing investments as well as those made by the agent. When neither the principal's nor the agent's investments are contractible, joint production calls for some sharing of revenues between the principal and the agent to help balance the resulting double-sided moral hazard problem. At the same time, there are other non-contractible decisions, which also affect revenues, but can be controlled by either the principal or the agent. In this paper, we study the optimal allocation of control rights over these transferable decision variables, taking into account the underlying double-sided moral hazard problem.

To do so, we develop a model that contains three types of non-contractible decisions: two sets of costly and non-transferable investment decisions—one for the principal and one for the agent—and a set of transferable decisions, each element of which can be controlled by either the principal or the agent. In our model, given any allocation of control rights over transferable decisions between the principal and the agent, linear contracts, in which the agent receives a share of the revenue it generates as well as a fixed payment (that could be negative), are optimal. Since revenues must be shared between the principal and the agents to incentivize their respective non-transferable investments, no allocation of control rights will achieve the first-best.

Our analysis yields three main sets of results. First, we show that control rights over the transferable decisions should all be given to the party—principal or agent—that obtains the higher revenue share in equilibrium. In other words, low-powered incentives (i.e. control over the transferable decisions) should be aligned with high-powered incentives (i.e. higher share of revenues) in order to minimize revenue-sharing distortions. A key implication of this result is that the principal only has to choose between two modes of organization: the $\mathcal{P}$-mode, in which the principal keeps control over all transferable decisions (this can be interpreted as employment), and the $\mathcal{A}$-mode, in which the principal gives control over all transferable decisions to the agent (this can be interpreted as independent contracting). The result holds under mild assumptions and does not require any positive interaction effects among the various decisions in the revenue function, nor any cost economies of scope across transferable decisions. In fact, its underlying logic implies that it may be optimal to allocate control over all transferable decisions to the same party even when that party is less cost-efficient than the other at investing in some of those decisions and there are no interaction effects across the decisions.

Second, we consider the effects on the choice between the $\mathcal{P}$-mode and $\mathcal{A}$-mode of changing (i) the importance of the agent's moral hazard relative to the principal's, and (ii) the contractibility of the agent's and the principal's non-transferable investments. In the optimal $\mathcal{A}$-mode contract, the agent keeps a larger share of revenue than in the optimal $\mathcal{P}$-mode contract, which means the $\mathcal{A}$-mode is better at generating profit from the agent's non-transferable investments. Therefore, if increasing the importance of the agent's moral hazard means increasing the weight placed on the *profit* generated by the agent's non-transferable investments in the principal's overall profit, then this will result in a shift towards the $\mathcal{A}$-mode. However, if instead what increases is the extent to which *revenue* depends on the agent's non-transferable investments, then the tradeoff shifts in favor of the $\mathcal{A}$-mode only under special conditions. In the latter case, there is an additional effect, which operates through a change in the equilibrium choices of non-transferable investments by the agent and which can lead to counterintuitive results.

In a similar vein, making one of the agent's non-transferable investments contractible shifts the tradeoff in favor of the $\mathcal{P}$-mode when the revenue function is fully additively separable. This conforms to common intuition: eliminating one of the sources of moral hazard for the agent implies there is less need to give the agent low-powered incentives, i.e. control over transferable actions. However, this result can be over-turned if there are positive interaction effects across the agent's non-transferable decisions, leading to a counter-intuitive conclusion. This is due to a potentially countervailing effect: rendering one of the agent's non-transferable investments contractible means it can be set at a higher (and therefore more efficient) level, which in turn increases the marginal gains from raising the agent's remaining non-transferable and non-contractible investments due to the positive interaction effects. Thus, the benefit of giving the agent low-powered incentives increases.

Third and finally, we also derive results concerning the choice between $\mathcal{P}$-mode and $\mathcal{A}$-mode when there are multiple agents and spillovers across the transferable decisions of different agents. In this case, the spillover-induced distortion shifts the baseline tradeoff between $\mathcal{P}$-mode and $\mathcal{A}$-mode by either exacerbating the revenue-sharing distortion (which favors the $\mathcal{P}$-mode) or offsetting it (which favors the $\mathcal{A}$-mode). The latter scenario leads to some counter-intuitive results.

For example, consider the case when the transferable decision is a revenue-increasing, costly investment (e.g. giving kickbacks to clients) and the spillovers are negative (e.g. a sales agent of a given manufacturer steals business from the manufacturer's other sales agents by giving clients greater kickbacks). In $\mathcal{A}$-mode, individual agents invest too much by not fully internalizing the spillovers. However, these higher investments can help offset the revenue-sharing distortion, namely that the party with control rights invests too little because it keeps less than 100% of the revenue generated. The $\mathcal{A}$-mode can then be a useful way for the principal to get agents to choose higher levels of the transferable decision variable without giving them an excessively high share of revenues. This mechanism has two counterintuitive consequences. First, when negative spillovers are not too large in magnitude, an increase in their magnitude shifts the tradeoff in favor of the $\mathcal{A}$-mode—the opposite of the standard "internalize externalities" logic. Second, if the magnitude of negative spillovers is sufficiently large, then agents get a lower share of revenues in $\mathcal{A}$-mode than in $\mathcal{P}$-mode and therefore the $\mathcal{A}$-mode (respectively, the $\mathcal{P}$-mode) is more likely to be chosen when the principal's (respectively, the agents') moral hazard becomes more important. This is the opposite of the standard "give control to the party whose investments are more important" prediction, which prevails when spillovers are positive.

A similar logic also applies when the transferable decision is the price charged to customers. In this case, the principal prefers the $\mathcal{A}$-mode when spillovers are moderately negative. Indeed, negative spillovers lead agents to set prices too high in $\mathcal{A}$-mode, but this mitigates the double-sided moral hazard problem by raising the return to the agent's and the principal's non-transferable investments.

Our theory provides a natural way to conceptualize the fundamental difference between traditional firms that hire employees and platforms/marketplaces that enable independent contractors to interact with customers, based on the allocation of control rights between firm and workers over decisions that affect the revenues generated from customers. Simply put, firms that allocate more control rights to workers are closer to the platform/marketplace model. In addition to being intuitively appealing, this conceptualization is also useful because it translates into empirically testable predictions, namely that (i) one should tend to observe that control rights over various transferable actions are all allocated to the same party (firm or agents), and (ii) firms that have chosen the $\mathcal{A}$-mode (independent contracting) should leave a larger share of their revenues to agents than organizations that have chosen the $\mathcal{P}$-mode (employment). Casual empiricism suggests these predictions hold in the industry examples discussed in this paper.

The next section discusses related literature. Section 3 provides some examples of markets in which the allocation of control rights we study is relevant. Section 4 provides a general analysis of the one principal-one agent case, in which there are multiple decisions of each type (transferable, non-transferable for the agent and non-transferable for the principal). In Section 5, we study the case with one principal, multiple agents, one decision of each type and spillovers across the transferable decisions of different agents. Section 6 concludes the paper.

2 Related literature

Our paper provides a theory of decision authority. As such it relates to both the literature on the theory-of-the-firm and the literature on internal organizational design.

In our model, we use elements from the formal theories of the firm based on property rights (Grossman and Hart, 1986, Hart and Moore, 1990 - collectively GHM) and incentive systems (Holmstrom and Milgrom, 1994): there are non-contractible decisions that can be allocated to the principal or external agents, and the principal must design first-period contracts that incentivize second-period effort by both the agents and the principal. However, our approach is different in two key respects.

First, we do not tie the choice of which party holds decision authority to the ownership of assets.[1] In the Grossman-Hart-Moore (GHM) framework, the tradeoff between employment and independent contracting is driven by the split of asset ownership: this determines the ex-post payoffs earned by the

[1] Note that viewing employees as agents who have contractually agreed that control over certain transferable decisions (e.g. equipment, pricing) be given to the firm does not suffer from Alchian and Demsetz's (1972) well-known critique of Coase (1937), namely that there is no meaningful distinction between the authority that a firm has over employees and the authority it has over suppliers or other external contractors.

various parties from their respective outside options, which in turn determines their ex-ante incentives to make relationship-specific investments. By contrast, in our paper, the difference between employment and independent contracting is defined by the allocation of control rights over non-contractible decisions that are chosen ex-post and affect joint payoffs. The tradeoff between the two governance modes is then driven by double-sided moral hazard: outside options are irrelevant as there are no assets to own and no ex-post negotiation. Thus, consistent with the Holmstrom and Milgrom (1994) incentive system framework and with empirical findings (e.g. Baker and Hubbard, 2004; Simester and Wernerfelt, 2005), our paper shifts the focus from ex-ante specific investments towards moral hazard. The main difference with respect to Holmstrom and Milgrom (1994) is that we introduce double-sided moral hazard and allow some decision rights to be shifted between the principal and the agents, whereas in their framework only the agents are subject to moral hazard, decision rights are fixed and thus the tradeoff between employment and independent contracting is still determined by asset ownership and outside options.

Second, we also explore the role of spillovers created by the transferable decisions corresponding to each agent on the payoffs generated by the other agents. These spillovers are internalized when the principal controls the transferable decisions, but not when they are controlled by agents—this difference plays a key role in our model and is not present in formal theories of the firm based on property rights or incentive systems. In particular, our spillovers are quite different from Hart and Moore (1990)'s complementaries across investments. In the Hart and Moore framework, the key mechanism through which a change in the asset ownership structure affects ex-ante investment incentives is by modifying the (positive) effect of an agent's investment on the marginal investment return of other agents in some coalition of agents. By contrast, our spillovers impact the tradeoff between the two modes through their effect on the total revenues generated by the various agents: the spillover created by an agent's investment decision on another agent's revenues need not have any effect on the latter agent's marginal return on investment, as illustrated by the linear example we use in Section 5.1.

Our framework also relates to the literature on decision authority within organizations (e.g. Aghion and Tirole, 1997, Alonso et al., 2008), for example whether to give the decision authority to employees or the manager. However, our focus on distortions due to double-sided moral hazard and spillovers rather than to uncertainty, exogenously given misalignment of objectives, information asymmetry and cheap talk differs from most of this literature. While the literature does sometimes incorporate (single-sided) moral hazard issues, these papers still rely on uncertainty for their results (e.g. Zábojník, 2002, Bester and Krähmer, 2008) as well as other sources of distortion (e.g. liquidity constraints, or exogenously given misalignment of objectives). As such, the set-ups and mechanisms are very different from ours.

Our results showing that control rights over transferable decisions should all be given to the party that obtains the higher revenue share are reminiscent of the colocation of control and income results obtained by Van den Steen (2010a and b). However, the underlying mechanisms are very different: ours is driven by double-sided moral hazard considerations, whereas Van den Steen's is driven by uncertainty and differing priors between the two parties.

Our results showing how the optimal allocation of control rights is affected by the importance of the agent's moral hazard relative to the principal's are aligned with existing results going as far back as Simon (1951), where the tradeoff between giving decision authority to employer or employee is determined by whose objective function is more sensitive to departures from optimality. Again, the underlying mechanisms are very different. The effects of making revenue depend less on one party's non-transferable investments or making that party's investment contractibile in our framework are more complex and can run counter to what one might expect based on existing results such as Simon's or those regarding asset ownership, in which ownership of productive assets would optimally shift to the other party (see, for example, Wernerfelt, 2002, Baker and Hubbard, 2004).

Our results showing that negative spillovers can make the principal more likely to prefer giving control over transferable actions to the agents as the magnitude of spillovers increases or as the principal's moral hazard becomes more important relative to the agents' moral hazard, is the opposite of what one would expect based, for example, on Grossman and Hart's (1986) prediction that ownership over assets (and therefore control) should be given to the party whose investment incentives are more important, or on Wernerfelt's (2002) prediction that ownership over a productive asset should be allocated to the firm or the worker depending on whose actions have a greater or less contractible effect on the asset's depreciation. These non-standard results are reminiscent of similarly counterintuitive results obtained in the literature on internal organization and delegation. Namely, Alonso et al. (2008) show that increasing the importance of coordinating decisions can make decentralization of decisions more likely, Zábojník (2002) shows that the principal is more likely to delegate authority to the agent when the principal is better informed given there are liquidity constraints, and Bester and Krahmer (2008) show that increasing the importance of the agent's moral hazard can make the principal less likely to delegate authority to the agent. However, as pointed out above, the reasons for the counter-intuitive results in these papers are very different from ours.

Since in our model revenues must be shared between the principal and agents to incentivize both sides to make non-contractible investments, we also directly build upon principal-agent models with double-sided moral hazard (Romano, 1994, Bhattacharyya and Lafontaine, 1995). The key difference relative to these papers is that we introduce a third type of non-contractible decision, control over which can be allocated to either the principal or the agents. We also generalize these settings by allowing each type of decision variable to be multi-dimensional.

Finally, this paper relates to two of our earlier works that study how firms choose to position themselves closer to or further from a multi-sided platform business model. The focus on incentives and double-sided moral hazard in the current paper differs from the one in Hagiu and Wright (2015a), where the tradeoff between operating as a marketplace or as a reseller was driven by the importance of agents' local information relative to the firm's. Here we abstract from information asymmetries. Closer to our current model is Hagiu and Wright (2015b), which derives a tradeoff between a vertically integrated mode and a multi-sided platform mode based on agents' private information, moral hazard, spillovers generated by the choice of a transferable decision across suppliers and network effects (more suppliers joining the firm raise the demand for each supplier). A key difference relative to Hagiu and

Wright (2015b) is the introduction of double-sided moral hazard, which is fundamental to the tradeoffs that we study here. Another difference is that the current model is much more general, and applies to a wider range of organizations rather than just multi-sided platforms facing cross-group network effects.

3 Examples

There are several different categories of markets in which the choice that we study is relevant. A large category involves firms that can either employ professionals and control how they deliver services to clients, or operate as platforms enabling independent professionals to provide services directly to clients. While this choice has become particularly prominent due to the proliferation of Internet-based service platforms (e.g. Coursera, Handy, Hourly Nerd, Lyft and Uber, Rubicon Global, Task Rabbit, and Upwork), it has long been relevant in a number of "offline" industries.[2]

The hair salon industry is a good example, as it has long featured two modes of organization, that can be viewed as corresponding to our $\mathcal{P}$-mode and $\mathcal{A}$-mode respectively. Some salons employ their hairstylists and pay them fixed hourly wages plus commissions that are a percentage of sales. Such salons control how individual hair dressers are promoted, and provide most of the supplies and equipment that stylists use for hair cutting and styling ($\mathcal{P}$-mode). In contrast, other salons rent out chairs (booths) to independent hairstylists. The stylists keep all earnings minus fixed monthly booth rental fees that are paid to the salon. In such salons, individual hair stylists promote themselves, and are responsible for providing and maintaining the majority of the supplies and equipment they need ($\mathcal{A}$-mode). In both modes, the salon owners still make all necessary investments to maintain the facilities and advertise the salon to customers, while the stylists must exert effort to provide quality service to customers.

A large category of relevant markets involves firms that need salespeople, brokers, or distributors to sell their products or services. Examples include the use of salespeople by manufacturers and the use of brokers by insurance companies. Firms in these markets often use a mix of independent agents, who among other things determine the extent of kickbacks they offer to purchase managers, and employees, for whom the firm determines and provides the kickbacks that are given to purchase managers.[3] The commission rates paid out by the firms vary substantially across the two modes (Anderson, 1985).

Similarly, firms providing a wide range of products or services can do so through company-owned outlets or through independent franchisees. Most business format franchisors (e.g. hotels, fast-food outlets, and car rentals) use a combination of upfront fixed franchise fees and sales-based royalties (Blair and Lafontaine, 2005). While franchise contracts are notoriously restrictive, franchisees nevertheless control some key decisions that impact the revenues they generate (e.g. their expenditure on staff). In contrast, these decisions are made by the firm in company-owned outlets.

[2] Note our focus on the allocation of control rights between agents and the principal is consistent with legal definitions that emphasize control rights as the most important factor determining whether agents should be considered independent contractors or employees. This is reflected in lawsuits surrounding Uber and other platforms in the "sharing economy".

[3] We study the case in which some agents operate in $\mathcal{P}$-mode and others in $\mathcal{A}$-mode in an online appendix.

An example that is particularly relevant for developing countries is sharecropping, in which landowners can decide how much to share their crops and relevant decision rights with agricultural workers. At one extreme, the landowner rents the land to a lessee at a fixed rate and the lessee has full control over inputs. At the other extreme, the landowner employs agricultural laborers at fixed wages and fully controls inputs. In between these two extremes, the landowner and the sharecropper share crops[4] and decision rights over inputs. Double-sided moral hazard is key in explaining the structure of sharecropping contracts, as noted by Bhattacharyya and Lafontaine (1995).

Table 1 shows how these and other examples fit our theory. In particular, it illustrates the three different types of non-contractible decision variables featured in our model that affect the revenue generated by each agent: (i) transferable decisions that are chosen by the principal in $\mathcal{P}$-mode and by the agent in $\mathcal{A}$-mode; (ii) costly ongoing investments always chosen by the agent; and (iii) costly ongoing investments always chosen by the principal.

We have not included the price charged to customers in Table 1, which is potentially another non-contractible and transferable decision variable in some of the examples listed. This is the case when price is subject to considerable uncertainty in the contracting stage, or when the principal cannot observe price and quantity separately (e.g. some consulting service agreements). However, oftentimes price is pinned down by market constraints, in which case it can be treated as a fixed constant in our analysis. Or price may be contractible, and indeed set by the principal in its contract with the agent—that possibility is covered by Proposition 1 below for the case with one agent and discussed in Section 5.1 for the case with multiple agents.

4 One agent and multiple actions

In this section, we analyze a general setting in which there is a principal (e.g. a firm) and a single agent, but potentially many transferable and non-transferable actions.

Let the revenue generated jointly by the principal and the agent if the latter accepts the principal's contract be denoted $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$, which depends on three types of actions, all of which are non-contractible.

The vectors of actions $\mathbf{q} = \left(q^1, .., q^{M_q}\right)$ and $\mathbf{Q} = \left(Q^1, .., Q^{M_Q}\right)$ are non-transferable: the agent chooses $\mathbf{q} \in \mathbb{R}_+^{M_q}$ at cost $c(\mathbf{q}) \equiv \sum_{i=1}^{M_q} c^i\left(q^i\right)$ and the principal chooses $\mathbf{Q} \in \mathbb{R}_+^{M_Q}$ at cost $C(\mathbf{Q}) \equiv \sum_{i=1}^{M_Q} C^i\left(Q^i\right)$. This means there is double-sided moral hazard: $\mathbf{q}$ encompasses ongoing effort and investment decisions that are always made by the agent and that raise the customers' willingness to pay for the service provided (see column 3 in Table 1), while $\mathbf{Q}$ captures the ongoing investments always made by the principal (see column 4 in Table 1).[5] In contrast, the actions contained in the vector $\mathbf{a} = \left(a^1, .., a^{M_a}\right)$ are transferable, i.e. each of them can be chosen *either* by the principal *or* by

[4] While 50/50 crop sharing is the most common practice, other splits are also used, as documented by Terpstra (1998).

[5] This formulation can also encompass investments that can be made by both parties, i.e. such that one party's investment in a given action does not preclude the other party from investing in the same type of action, each carrying its own cost. For example, both the principal and the agent can invest in training. In this case, we can always define q^j as the agent's investment in training and Q^k as the principal's investment in training, with revenue affected by both.

Table 1: Examples

	Transferable decisions	*Non-transferable investment decisions made by agents*	*Non-transferable investment decisions made by the firm*
Hair salons	promotion of individual hair dressers; quality and maintenance of equipment for hair cutting and styling; spending on supplies	service quality	maintenance and advertising of the salon
Transportation (e.g. Uber vs. traditional taxi companies)	quality and appearance of the car (e.g. make, model, color); location/area of work	service quality; maintenance of the car (e.g. cleanliness)	quality of the technological infrastructure (payment, dispatch system); advertising
Consulting (e.g. Hourly Nerd vs. McKinsey) and outsourcing (e.g. Upwork vs. Infosys)	promotion of individual professionals and their skills; price	service quality	quality of the (online) system for communication, monitoring and payment; advertising
Online education (e.g. Coursera vs. University of Phoenix)	quality of the course design; advertising of individual instructors and courses	quality of course delivery	quality of the online infrastructure; advertising of the site
Waste and recycling (e.g. Rubicon Global vs. Waste Management)	condition and maintenance of equipment for waste collection and hauling	service quality	quality of the technological infrastructure (payment, scheduling routes and pick-up); advertising
Producers and sales agents	kickbacks to clients	sales effort	advertising and quality of the product or service
Franchising	expenditure on staff and their benefits	outlet manager effort	quality of the product (franchisor); national advertising
Sharecropping	quality of inputs (seeds, fertilizer and pesticides); bribes	sharecropper effort	large investments (maintenance of irrigation system)

the agent, depending on how the principal chooses to allocate control rights (see column 2 of Table 1). The party that chooses $a^i \in \mathbb{R}_+$ incurs cost $f^i\left(a^i\right)$. For convenience, we denote $f\left(\mathbf{a}\right) \equiv \sum_{i=1}^{M_a} f^i\left(a^i\right)$. We assume there is at least one action of each type, i.e. $M_q \geq 1$, $M_Q \geq 1$ and $M_a \geq 1$.

The principal chooses the set $D \subset \{1, .., M_a\}$ of transferable decisions over which it keeps control (leaving the agent to control decisions $i \in \{1, .., M\} \setminus D$) and offers a revenue-sharing contract $\Omega\left(R\right)$ to the agent, where $\Omega\left(.\right)$ can be any arbitrary function of the revenue R generated. The contract means that the agent obtains $\Omega\left(R\right)$, while the principal obtains $R - \Omega\left(R\right)$. In the analysis that follows, two extreme allocations of decision rights will play a central role: $D = \{1, .., M_a\}$, which we call the $\mathcal{P}$-mode (the principal controls all transferable decisions) and $D = \varnothing$, which we call the $\mathcal{A}$-mode (the agent controls all transferable decisions).

The key assumption in this specification is that only the realized revenue R is contractible, whereas the underlying variables $\left(\mathbf{a}, \mathbf{q}, \mathbf{Q}\right)$ are not. We later investigate what happens when some of these variables become contractible. We also assume that the principal cannot commit to "throwing away" revenue in case a target specified ex-ante is not reached (Holmstrom, 1982). Ex-ante commitments to destroy revenue seem unrealistic, as they require enforcement by an external third party, who then becomes itself subject to a moral hazard problem—this is one reason why they are seldom used in practice (Eswaran and Kotwal, 1984).

Our model can indeed be interpreted as the reduced form of a model with uncertainty, which could be one reason that the various types of actions are non-contractible in practice. With considerable uncertainty regarding the effects of the various actions on revenues, contingent decision rules could be difficult to describe. In that case, one can resort to the results of Holmstrom and Milgrom (1987) to justify our focus on linear contracts. Furthermore, linear and uniform contracts are prevalent in all of the examples listed in Table 1 (Bhattacharyya and Lafontaine, 1995 provide empirical evidence in the contexts of franchising and sharecropping).

We assume the principal holds all the bargaining power. This implies that it will set $\Omega(.)$ in both modes so that the agent is indifferent between participation and its outside option, which for convenience we normalize to zero throughout. In our model it is immaterial whether the principal or the agent collects revenues R and pays the other party their share according to contract $\Omega(R)$. For instance, if the principal is a firm that employs the agent, then the payment of $\Omega(R)$ can be interpreted as a combination of fixed wage plus bonus in an employment relationship.

The game that we study has the following timing. In stage 1, the principal chooses the allocation of control over transferable actions $D \subset \{1, .., M_a\}$ and the associated contract $\Omega(.)$; the agent decides whether or not to accept the contract. In stage 2, the principal chooses $\mathbf{Q}$ and all a^i's such that $i \in D$, while the agent simultaneously chooses $\mathbf{q}$ and all a^i's such that $i \in \{1, .., M_a\} \backslash D$. Finally, in stage 3, revenues $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ are realized; the principal receives $R - \Omega(R)$ and the agent receives $\Omega(R)$.

Throughout the paper, we use the following two notational conventions. First, for variables and parameters that apply to both the agent and the principal, we use lowercase for the agent and uppercase for the principal (e.g. $\mathbf{q}$ and $\mathbf{Q}$). Second, subscripts next to functions always indicate derivatives: for example, $f^i_{a^i}$ indicates the derivative of f^i with respect to a^i and R_{a^i} indicates the partial derivative of R with respect to the transferable action a^i.

We make the following technical assumptions:

(a1) All functions are twice continuously differentiable in all arguments.

(a2) Each of the cost functions $\left(f^1, .., f^{M_a}, c^1, .., c^{M_q}, C^1, .., C^{M_Q}\right)$ is either everywhere equal to 0 or increasing and strictly convex and there exists $i \in \{1, .., M_a\}$, $j \in \{1, .., M_q\}$ and $k \in \{1, .., M_Q\}$ such that $f^i \neq 0$, $c^j \neq 0$ and $C^k \neq 0$. For any $i \in \{1, .., M_a\}$, $j \in \{1, .., M_q\}$ and $k \in \{1, .., M_Q\}$, if $f^i \neq 0$, then $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ is strictly increasing in a^i, if $c^j \neq 0$, then $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ is strictly increasing in q^j, if $C^k \neq 0$, then $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ is strictly increasing in Q^k, and

$$f^i(0) = f^i_{a^i}(0) = c^j(0) = c^j_{q^j}(0) = C^k(0) = C^k_{Q^k}(0) = 0.$$

(a3) For all $t \in [0, 1]$, $tR(\mathbf{a}, \mathbf{q}, \mathbf{Q}) - f(\mathbf{a}) - c(\mathbf{q}) - C(\mathbf{Q})$ is strictly concave in $(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ and admits a unique finite maximizer in any subset of the $M_a + M_q + M_Q$ variables $(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ for any values of the remaining variables.

(a4) For all $\boldsymbol{\tau} \in [0,1]^{M_a+M_q+M_Q}$*, the system of equations*

$$\left\{ \begin{array}{cc} \tau^i R_{a^i}(\mathbf{a},\mathbf{q},\mathbf{Q}) = f^i_{a^i}\left(a^i\right) & \text{for } i \in \{1,..,M_a\} \\ \tau^{M_a+j} R_{q^j}(\mathbf{a},\mathbf{q},\mathbf{Q}) = c^j_{q^j}\left(q^j\right) & \text{for } j \in \{1,..,M_q\} \\ \tau^{M_a+M_q+k} R_{Q^k}(\mathbf{a},\mathbf{q},\mathbf{Q}) = C^k_{Q^k}\left(Q^k\right) & \text{for } k \in \{1,..,M_Q\} \end{array} \right. \tag{1}$$

admits a unique solution $(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau}))$.

These assumptions are standard and ensure that the optimization problems considered below are well-behaved. Assumptions (a3) and (a4) ensure that there is always a unique finite solution to the optimization problems we consider; in particular, they obviate the need for more general stability conditions, that would be quite complex here. Furthermore, the principal always finds it optimal to induce the agent to participate.

One noteworthy feature is that assumptions (a2)-(a4) allow the vectors $(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ to encompass both costly actions that always increase revenues (e.g. investments in advertising, equipment, technology infrastructure, etc.) and costless actions, in which the revenue function R is single-peaked (e.g. price, the allocation of a fixed promotional capacity between emphasizing the principal's brand vs. the agent's previous work experience). We do, however, impose that at least one action of each type is costly, for reasons discussed below.

4.1 Main results

We first establish that in our set-up, we can restrict attention to linear contracts without loss of generality (the proof, together with all other proofs, is in the appendix).

Lemma 1 *If assumptions (a1)-(a4) hold, then, given any allocation of control rights over the transferable actions, the principal can achieve the best possible outcome with a linear contract.*

This result is an extension to multiple non-contractible actions of similar results obtained in double-sided moral hazard settings by Romano (1994) and Bhattacharyya and Lafontaine (1995). It implies that we can restrict attention to contracts offered by the principal that take the form

$$\Omega(R) = (1-t)R - T,$$

where T can be interpreted as the fixed fee collected by the principal and $t \in [0,1]$ as the share of revenue kept by the principal. This means the net payoff received by the principal is $tR + T$ and the net payoff received by the agent is $(1-t)R - T$. In general, the optimal contract will have different (t, T) depending on the allocation of control over transferable actions D. Thus, it is possible for T to be negative under some allocations (i.e. the agent receives a fixed wage) and positive under other allocations (i.e. the agent pays a fixed fee).

Lemma 1 implies that, given an allocation of decision rights $D \subset \{1,..,M_a\}$ chosen by the principal, his profits can be written as[6]

$$\Pi^*(D) = \max_{t,\mathbf{a},\mathbf{q},\mathbf{Q}} \{R(\mathbf{a},\mathbf{q},\mathbf{Q}) - f(\mathbf{a}) - c(\mathbf{q}) - C(\mathbf{Q})\} \tag{2}$$

s.t.

$$\begin{cases} tR_{a^i}(\mathbf{a},\mathbf{q},\mathbf{Q}) = f^i_{a^i}(a^i) \text{ for } i \in D \\ (1-t)R_{a^i}(\mathbf{a},\mathbf{q},\mathbf{Q}) = f^i_{a^i}(a^i) \text{ for } i \in \{1,..,M_a\}\backslash D \\ (1-t)R_{q^j}(\mathbf{a},\mathbf{q},\mathbf{Q}) = c^j_{q^j}(q^j) \text{ for } j \in \{1,..,M_q\} \\ tR_{Q^k}(\mathbf{a},\mathbf{q},\mathbf{Q}) = C^k_{Q^k}(Q^k) \text{ for } k \in \{1,..,M_Q\}. \end{cases} \tag{3}$$

In general, for any D, the principal's profits are lower than the first-best profit level

$$\max_{\mathbf{a},\mathbf{q},\mathbf{Q}} \{R(\mathbf{a},\mathbf{q},\mathbf{Q}) - f(\mathbf{a}) - c(\mathbf{q}) - C(\mathbf{Q})\}.$$

The reason is that revenue $R(\mathbf{a},\mathbf{q},\mathbf{Q})$ needs to be divided between the principal and the agent to incentivize each of them to choose their respective actions. This inefficiency is the moral hazard in teams identified by Holmstrom (1982), where a team here consists of the agent and the principal.

Inspection of the program (2) makes it clear that different allocations of control rights over the transferable actions $\mathbf{a}$ lead to different profits. It is worthwhile to emphasize that, in order for the optimal choice of D to be non-trivial, at least one action for each of the three types ($\mathbf{a}$, $\mathbf{q}$ and $\mathbf{Q}$) must carry a strictly positive cost, which is assumed as part of (a2). Otherwise, if all transferable actions were costless, then, for any $i \in \{1,..,M_a\}$, the first-order condition in a^i from 3 either disappears or does not depend on t, so any allocation allocation of control rights leads to the same profits for the principal, i.e. $\Pi^*(D) = \Pi^*$ for all $D \subset \{1,..,M_a\}$. If the agent's non-transferable investments were all costless, then the principal would be able to attain the first-best outcome with $D^* = \{1,..,M_a\}$ and $t = 1$. If on the other hand the principal's non-transferable investments were all costless, then the principal would be able to attain the first-best outcome with $D^* = \varnothing$ and $t = 0$.

For future reference, denote by $t^{\mathcal{P}*}$ and $t^{\mathcal{A}*}$ the respective optimal variable fees charged by the principal in $\mathcal{P}$-mode and $\mathcal{A}$-mode, i.e. the respective solutions in t that emerge from program (2) when $D = \{1,..,M_a\}$ and $D = \varnothing$. Also let

$$\Pi^{\mathcal{P}*} \equiv \Pi^*(\{1,..,M_a\}) \text{ and } \Pi^{\mathcal{A}*} \equiv \Pi^*(\varnothing).$$

The analysis in the rest of this section will rely on the following additional assumption:

(a5) *For any* $\boldsymbol{\tau} \in [0,1)^{M_a+M_q+M_Q}$, *if* $\Pi(\boldsymbol{\tau}) \equiv R(\mathbf{a}(\boldsymbol{\tau}),\mathbf{q}(\boldsymbol{\tau}),\mathbf{Q}(\boldsymbol{\tau})) - f(\mathbf{a}(\boldsymbol{\tau})) - c(\mathbf{q}(\boldsymbol{\tau})) - C(\mathbf{Q}(\boldsymbol{\tau}))$, *where* $(\mathbf{a}(\boldsymbol{\tau}),\mathbf{q}(\boldsymbol{\tau}),\mathbf{Q}(\boldsymbol{\tau}))$ *is the unique solution to the system of equations (1), then*

[6] At the optimum, the fixed fee T of the linear contract is always set such that the participation constraint of the agent is binding, i.e.

$$(1-t)R(\mathbf{a},\mathbf{q},\mathbf{Q}) - \sum_{i\in\{1,..,M_a\}\backslash D} f^i(a^i) - c(\mathbf{q}) - T = 0.$$

$\Pi(\boldsymbol{\tau})$ *is strictly increasing in all* τ^j *such that* $j \in \{1, .., M_a\}$ *and* $f^j \neq 0$*, in all* τ^{M_a+k} *such that* $k \in \{1, .., M_q\}$ *and* $c^k \neq 0$*, and in all* $\tau^{M_a+M_q+l}$ *such that* $l \in \{1, .., M_Q\}$ *and* $C^l \neq 0$*.*

In words, this assumption requires that reducing the distortion in any second stage decision problem for a costly action (by increasing τ^i) leads to an increase in the overall net profit for the principal. Note that changing the τ^i corresponding to costless actions has no impact on the resulting net profit $\Pi(\boldsymbol{\tau})$ for the principal. In the appendix we prove that a sufficient condition for (a5) to hold is that $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ is (weakly) supermodular in all of its arguments. However, supermodularity is not necessary: (a5) can still hold even when the various non-contractible actions are strategic substitutes in the revenue function. What is required in this case is that the interaction effects are not too negative, so that they do not overwhelm the direct positive effect on net profits of increasing the incentive to invest in any given non-contractible action by raising the corresponding τ^i. Seen in this light, (a5) is a rather mild assumption.

We are now ready to derive the central results in this section.

Proposition 1 *If assumptions (a1)-(a5) hold, then:*

1. *The optimal allocation of control rights over transferable decisions is either* $D^* = \varnothing$ *($\mathcal{A}$-mode) or* $D^* = \{1, .., M_a\}$ *($\mathcal{P}$-mode).*

2. $t^{\mathcal{P}*} < 1/2$ *implies* $\Pi^{\mathcal{P}*} < \Pi^{\mathcal{A}*}$ *(i.e. the $\mathcal{A}$-mode is optimal) and* $t^{\mathcal{A}*} > 1/2$ *implies* $\Pi^{\mathcal{A}*} < \Pi^{\mathcal{P}*}$ *(i.e. the $\mathcal{P}$-mode is optimal). Furthermore,* $t^{\mathcal{P}*} = 1/2$ *implies* $\Pi^{\mathcal{P}*} \leq \Pi^{\mathcal{A}*}$ *and* $t^{\mathcal{A}*} = 1/2$ *implies* $\Pi^{\mathcal{A}*} \leq \Pi^{\mathcal{P}*}$*.*

The first result in Proposition 1 implies that, provided (a1)-(a5) hold, the principal's optimal allocation of control rights over the transferable actions can only be one of two options—the $\mathcal{P}$-mode or the $\mathcal{A}$-mode. The second result in Proposition 1 then says that the principal would never find it optimal to function in $\mathcal{P}$-mode and keep less than 50% of revenue or function in $\mathcal{A}$-mode and keep more than 50% of revenue.

The key driving force behind the proposition is that giving control rights to the party that obtains a higher share of revenues (i.e. aligning low-powered and high-powered incentives) reduces distortions and thereby raises the principal's profit. To see this, denote by t^* the optimal variable fee associated with the optimal allocation of control rights D^*. If $t^* < 1/2$ and $D^* \neq \varnothing$, then the distortions can be reduced by shifting control over all transferable actions in D^* from the principal to the agent. Indeed, this changes the first-order condition determining a^i in the second stage from

$$t^* R_{a^i}(\mathbf{a}, \mathbf{q}, \mathbf{Q}) = f^i_{a^i}\left(a^i\right)$$

to

$$(1 - t^*) R_{a^i}(\mathbf{a}, \mathbf{q}, \mathbf{Q}) = f^i_{a^i}\left(a^i\right)$$

for all $i \in D^*$. The first-order conditions in q and Q stay unchanged. Given (a5) and $1 - t^* > 1/2 > t^*$, this change leads to an outcome that is closer to the first-best and therefore higher equilibrium profits for the principal. If $t^* > 1/2$ and $D^* \neq \{1, .., M_a\}$, then, by the same logic, profits can be increased by shifting control over all transferable actions not already in D^* to the principal. Only in the special case when $t^* = 1/2$ would any split of control rights, including a strictly interior split, be optimal.

It is worthwhile to emphasize that these results are not driven by positive interaction effects between the various non-contractible actions in revenue (recall that supermodularity is not necessary), nor by any cost economies of scope across transferable actions. Indeed, we have assumed the costs of the transferable actions are independent of one another. If there were economies of scope among them, then that would provide an additional reason for giving the same party control (and therefore cost responsibility) for all of these actions.

Finally, Proposition 1 would remain valid if we added an arbitrary number of contractible decisions (e.g. price) that impact the revenue function and that the principal could set at the same time it chooses the optimal control allocation and contract for the agent.

An empirically testable implication of the first result in Proposition 1 is that one should rarely observe strictly interior splits of control rights over transferable actions (such interior splits can only be optimal if some of the non-contractible actions are strong strategic substitutes in the revenue function). For example, in the hair salon example, the promotion of individual hair dressers, the quality and maintenance of equipment for hair cutting and styling, and the spending on supplies are either all controlled by the salon (employment mode) or all controlled by the individual hairdressers (booth rental mode).

The second result in Proposition 1 can be re-stated in a more empirically useful way. Define

$$t^* \equiv \begin{cases} t^{\mathcal{P}*} & \text{if} \quad \Pi^{\mathcal{P}*} \geq \Pi^{\mathcal{A}*} \\ t^{\mathcal{A}*} & \text{if} \quad \Pi^{\mathcal{P}*} < \Pi^{\mathcal{A}*}, \end{cases}$$

which is the optimal variable fee charged by the principal in the optimal mode. The following corollary is a logical reformulation of the second result in Proposition 1.

Corollary 1 *If assumptions (a1)-(a5) hold, then* $t^* < 1/2$ *if and only if the* $\mathcal{A}$*-mode is strictly optimal (i.e.* $\Pi^{\mathcal{A}*} > \Pi^{\mathcal{P}*}$*) and* $t^* \geq 1/2$ *if and only if the* $\mathcal{P}$*-mode is weakly optimal (i.e.* $\Pi^{\mathcal{P}*} \geq \Pi^{\mathcal{A}*}$*).*

Thus, according to this prediction of our model, the agent obtains more than 50% of attributable revenues if and only if the principal has chosen the $\mathcal{A}$-mode. In other words, low-powered incentives (i.e. control over **a**) should be aligned with high-powered incentives (i.e. larger share of revenues). This provides another empirically testable implication: other things being equal, we expect that organizations that have chosen the $\mathcal{A}$-mode should leave a larger share of their revenues to agents than organizations that have chosen the $\mathcal{P}$-mode.[7] For example, hair salons that rent out chairs charge only a fixed rental fee, letting stylists keep 100% of sales, whereas traditional hair salons that employ

[7] In the case of sales agents for products, the percentage commissions for independent agents are higher than those for employees, but both are much lower than 50% (usually less than 10%) of revenues from product sales. This reflects the fact that revenue also includes production costs, which in many cases is not easily observed by the agent. For these cases, the revenue R in our model is best interpreted as revenue net of production costs.

their hairstylists offer bonuses ranging from 35% to 60% of sales.[8]

4.2 Effects of moral hazard and contractibility

We now turn to examining the impact on the optimal choice of mode of two key factors: (i) the importance of the agent's and the principal's moral hazard, and (ii) the contractibility of the agent's and the principal's non-transferable investments. Throughout this subsection, we continue to assume (a1)-(a5) hold, so that the only possible optimal allocations are $\mathcal{A}$-mode or $\mathcal{P}$-mode, following Proposition 1.

We start by precisely defining the meaning of shifting the tradeoff in favor of one mode or the other. Denote by $\widetilde{\Pi}^{\mathcal{P}*}$ and $\widetilde{\Pi}^{\mathcal{A}*}$ the respective profits in the two modes after a change in some parameters or in the contracting environment, and by $\Pi^{\mathcal{P}*}$ and $\Pi^{\mathcal{A}*}$ the respective profits before the change.

Definition *We say that the tradeoff is shifted in favor of the $\mathcal{A}$-mode if*

$$\Pi^{\mathcal{A}*} \geq \Pi^{\mathcal{P}*} \Longrightarrow \widetilde{\Pi}^{\mathcal{A}*} > \widetilde{\Pi}^{\mathcal{P}*}$$

for all initial values of the parameters. Conversely, the tradeoff is shifted in favor of the P-mode if

$$\Pi^{\mathcal{P}*} \geq \Pi^{\mathcal{A}*} \Longrightarrow \widetilde{\Pi}^{\mathcal{P}*} > \widetilde{\Pi}^{\mathcal{A}*}$$

for all initial values of the parameters.

4.2.1 Importance of moral hazard

First, consider what happens to the tradeoff between the $\mathcal{P}$-mode and the $\mathcal{A}$-mode when the agent's (respectively, the principal's) moral hazard becomes more important. To make this meaningful in our model, we further assume the revenue function is additively separable in the three types of non-contractible actions $\mathbf{a}$, $\mathbf{q}$ and $\mathbf{Q}$, i.e.

$$R(\mathbf{a}, \mathbf{q}, \mathbf{Q}) = R^a(\mathbf{a}) + \phi R^q(\mathbf{q}) + \Phi R^Q(\mathbf{Q}), \tag{4}$$

where $(\phi, \Phi) \in \mathbb{R}_+^2$ are two parameters.[9] Let also

$$c(\mathbf{q}) \equiv \sum_{j=1}^{M_q} \phi c^j\left(q^j\right) \text{ and } C(\mathbf{Q}) \equiv \sum_{k=1}^{M_Q} \Phi C^k\left(Q^k\right).$$

Thus, ϕ is the weight placed on the profit generated by the agent's non-transferable investments $\pi^q(\mathbf{q}) \equiv R^q(\mathbf{q}) - \sum_{i=1}^{M_q} c^i\left(q^i\right)$, and can be interpreted as measuring the importance of the agent's moral hazard. Similarly, Φ is the weight placed on the profit generated by the principal's non-transferable

[8] See "Hair & Nail Salons in the US," IBIS World Industry Report 81211, February 2015.

[9] If there were interaction effects across the various types of actions in the revenue function, it would be difficult to provide a clear notion of the importance of the agent's (respectively, the principal's) moral hazard.

investments $\pi^Q(\mathbf{Q}) \equiv R^Q(\mathbf{Q}) - \sum_{i=1}^{M_Q} C^i\left(Q^i\right)$, and can be interpreted as measuring the importance of the principal's moral hazard.

With these definitions, we obtain the following result.

Proposition 2 *Suppose assumptions (a1)-(a5) hold and the revenue function takes the additively separable form in* $\mathbf{a}$, $\mathbf{q}$ *and* $\mathbf{Q}$ *given by (4). If* ϕ *(respectively,* Φ*) is the weight placed on the profit generated by the agent's (respectively, the principal's) non-transferable investments, then an increase in* ϕ *(respectively, in* Φ*) shifts the tradeoff in favor of the* $\mathcal{A}$*-mode (respectively,* $\mathcal{P}$*-mode).*

Intuitively, in the optimal $\mathcal{A}$-mode contract, the agent keeps a larger share of revenue than in the optimal $\mathcal{P}$-mode contract, which means the $\mathcal{A}$-mode is better at generating profit from the agent's non-transferable investments. Therefore, increasing the weight placed on the profit generated by the agent's non-transferable investments in the principal's overall profit will result in a shift towards the $\mathcal{A}$-mode.

More formally, because ϕ multiplies both the revenue and the cost associated with the agent's non-transferable investments, it does not change the second-stage equilibrium choices of actions $(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ for any given t in either mode. Moreover, due to the envelope theorem, the effects of a change in ϕ through changes in $t^{\mathcal{P}*}$ and $t^{\mathcal{A}*}$ are second-order. As a result, the only first-order effect of ϕ on the profit differential between the $\mathcal{A}$-mode and the $\mathcal{P}$-mode is the direct effect $\pi^q\left(\mathbf{q}^{\mathcal{A}*}\right) - \pi^q\left(\mathbf{q}^{\mathcal{P}*}\right)$. Additive separability of R in $\mathbf{a}$, $\mathbf{q}$ and $\mathbf{Q}$ implies $t^{\mathcal{A}*} < t^{\mathcal{P}*}$. Along with (a5), this implies that the profit generated by the agent's non-transferable investment is larger in $\mathcal{A}$-mode, i.e. $\pi^q\left(\mathbf{q}^{\mathcal{A}*}\right) > \pi^q\left(\mathbf{q}^{\mathcal{P}*}\right)$, reflecting there is less distortion in the choice of $\mathbf{q}$ in $\mathcal{A}$-mode. Thus, if the agent's moral hazard becomes more important, i.e. ϕ increases, then the profitability of $\mathcal{A}$-mode will improve vis-à-vis $\mathcal{P}$-mode. The logic for an increase in Φ is symmetric.

However, if instead of increasing the weight on the *profit* generated by the agent's non-transferable investments, we only increased the weight on the *revenue* generated by the agent's non-transferable investments (i.e. costs remain $c(\mathbf{q}) = \sum_{j=1}^{M_q} c^j\left(q^j\right)$ regardless of ϕ), then the effect on the tradeoff between $\mathcal{A}$-mode and $\mathcal{P}$-mode would in general be ambiguous.[10] Indeed, in this case there are two first-order effects of increasing ϕ: (i) the same direct effect as above (a larger ϕ increases the relative profitability of $\mathcal{A}$-mode because the revenue generated from the agent's investments $\mathbf{q}$ is higher in $\mathcal{A}$-mode), (ii) an indirect effect which operates through a change of the optimal $\mathbf{q}$ in both modes.

To evaluate the indirect effect, note that the vector of the agent's non-transferable investments chosen in the second stage is equal to $\mathbf{q}(t\phi)$, where

$$\mathbf{q}(z) \equiv \arg\max_{\mathbf{q}} \left\{ z R^q(\mathbf{q}) - c(\mathbf{q}) \right\}. \tag{5}$$

Due to (a1)-(a4), $\mathbf{q}(z)$ is increasing, so an increase in ϕ now leads to a first-order increase in $\mathbf{q}$ in both modes, even when t is held constant. We still have $t^{\mathcal{A}*} < t^{\mathcal{P}*}$ and therefore $\mathbf{q}^{\mathcal{A}*} > \mathbf{q}^{\mathcal{P}*}$, but an

[10] If, alternatively, we only increased the weight on the agent's cost of non-transferable investments, we would obtain the reverse of the effects described in this paragraph.

increase in ϕ may lead to a larger or smaller increase in $\mathbf{q}^{\mathcal{A}*}$ relative to $\mathbf{q}^{\mathcal{P}*}$, so the indirect effect may be positive or negative. Nevertheless, we can provide a sufficient condition for the first (positive) effect to dominate. By symmetry, define

$$\mathbf{Q}(Z) \equiv \arg\max_{\mathbf{Q}} \left\{ ZR^{Q}(\mathbf{Q}) - C(\mathbf{Q}) \right\}. \tag{6}$$

We then obtain the following result.

Proposition 3 *Suppose assumptions (a1)-(a5) hold and the revenue function takes the additively separable form in* $\mathbf{a}$*,* $\mathbf{q}$ *and* $\mathbf{Q}$ *given by (4). If* ϕ *(respectively,* Φ*) is the weight placed on the revenue generated by the agent's (respectively, the principal's) non-transferable investments and if the function* $\mathbf{q}(z)$ *defined in (5) is such that* $zq_z^j(z)$ *is increasing in* z *for all* $j \in \{1,..,M_q\}$ *(respectively, the function* $\mathbf{Q}(Z)$ *defined in (6) is such that* $ZQ_Z^k(Z)$ *is increasing in* Z *for all* $k \in \{1,..,M_Q\}$*), then an increase in* ϕ *(respectively,* Φ*) shifts the tradeoff in favor of the* $\mathcal{A}$*-mode (respectively,* $\mathcal{P}$*-mode).*

This proposition emphasizes that the standard intuition—decision rights should be given to the party whose moral hazard is more "important"—only holds under special conditions when the importance of moral hazard refers to the effect on revenue rather than the effect on profit. It is easily verified that the sufficient condition provided on $\mathbf{q}(z)$ holds when R^q is linear in all of its arguments and c^i is quadratic in q^i for all $i \in \{1,..,M_q\}$, which is the functional form adopted in Section 5.1.

4.2.2 Contractibility of non-transferable investments

Next, we wish to know what happens to the tradeoff between $\mathcal{A}$-mode and $\mathcal{P}$-mode when one (or several) of the non-transferable investments—the principal's or the agent's—become contractible. For conciseness, we will investigate whether making q^1 contractible shifts the tradeoff in favor of the $\mathcal{A}$-mode or in favor of the $\mathcal{P}$-mode (the effect of making Q^1 contractible is then obtained by symmetry).

Standard intuition based on the traditional theory of the firm literature suggests that if one party's actions become more contractible, then ownership of productive assets—and with it, residual control rights—should shift to the other party. Wernerfelt (2002) proves a formal version of this idea, while Baker and Hubbard (2004) provide empirical evidence in the context of trucking. Specifically, the advent of on-board computers and GPS technology made it possible for trucking companies to condition contracts on their drivers' driving behavior. Baker and Hubbard show that this change in technology led to a significant shift among trucking companies from hiring truck drivers as independent contractors to hiring them as employees.

The corresponding intuition in our model runs as follows. If q^1 becomes contractible, then that reduces the agent's moral hazard, which means there is less need to give the agent high-powered incentives (i.e. low t). This in turn implies that giving the agent low-powered incentives (i.e. control over the a^i's) becomes relatively less attractive: in other words, the $\mathcal{P}$-mode should become relatively more attractive. This intuition is correct when q^1 is the only non-transferable investment for the agent or when the revenue function is additively separable in all of its arguments.

Proposition 4 *Suppose assumptions (a1)-(a5) hold. If $M_q = 1$ (respectively, $M_Q = 1$) or the revenue function $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ is additively separable in all of its arguments, then making q^1 contractible (respectively, making Q^1 contractible) shifts the tradeoff in favor of the $\mathcal{P}$-mode (respectively, in favor of the $\mathcal{A}$-mode).*

However, the intuition above is incomplete when there are positive interaction effects in the revenue function between the multiple non-transferable investments q^i chosen by the agent, i.e. $M_q \geq 2$ and R is strictly supermodular in $\mathbf{q}$. In that case, making q^1 contractible has another effect that can go in the opposite direction. Indeed, q^1 can now be set to a higher value, which by supermodularity implies the other q^i's will be increased as well. If $t^{\mathcal{A}*} < t^{\mathcal{P}*}$ (which holds as long as the interaction effects across $\mathbf{a}$, $\mathbf{q}$ and $\mathbf{Q}$ are not too strong), then the gain from being able to increase q^i is larger in $\mathcal{A}$-mode than in $\mathcal{P}$-mode, which means that rendering q^1 contractible makes the $\mathcal{A}$-mode relatively more attractive. The balance of the two effects can go either way, and it is straightforward to come up with functional forms, for which the prediction in Proposition 4 no longer holds.

Example *Suppose revenue and costs are*

$$\begin{aligned} R(a, q_1, q_2, Q) &= a + 0.5q_1 + 0.7q_2 + 1.6q_1q_2 + 2Q \\ f(a) &= a^2,\ c^1(q_1) = q_1^2,\ c^2(q_2) = q_2^2 \text{ and } C(Q) = Q^2. \end{aligned}$$

With these functional forms, it is easily verified that

$$\Pi^{\mathcal{P}*} - \Pi^{\mathcal{A}*} > 0 > \widetilde{\Pi}^{\mathcal{P}*} - \widetilde{\Pi}^{\mathcal{A}*},$$

which means that if all four *actions are non-contractible, then the $\mathcal{P}$-mode is preferred, whereas if q^1 is contractible, then the $\mathcal{A}$-mode is preferred.*

In summary, the principal faces the following tradeoff. On the one hand, after one of the agent's non-transferable investments becomes contractible, there are fewer sources of moral hazard for the agent. On the other hand, being able to contract on that investment can raise the marginal gains from increasing the agent's remaining non-transferable and non-contractible investments if there are positive interaction effects (note that multi-dimensionality of the agent's non-transferable investments is crucial for this effect). The net effect determines whether the benefit from giving the agent low-powered incentives increases or decreases.

4.3 Cost asymmetries

Proposition 1 assumes away differences between the principal and the agent in the costs of undertaking the transferable actions or in their impact on revenues. In some real-world examples, such differences are an important factor in determining which control rights are held by the principal and which are held by agents. For example, the principal may have economies-of-scale advantages over individual agents when incurring the cost associated with some transferable actions (e.g. volume discounts in purchasing equipment) or better information regarding the impact of those transferable actions on revenues due to access to more data (e.g. Uber and Lyft when setting prices for rides). In other contexts, the cost or information advantages lie with the agent (e.g. sharecroppers may have better knowledge than the landowners for determining expenditure on seeds, fertilizer and pesticides; the same is true for franchisees when choosing staff benefits).

Introducing cost asymmetries is straightforward in our model and provides an easy way of explaining why some control rights are held by the principal and others are held by agents. Specifically, we can simply assume that the cost of the transferable action a^i is $F^i\left(a^i\right)$ when incurred by the principal and $f^i\left(a^i\right)$ when incurred by the agent, where the functions F^i have the same properties as previously assumed for the functions f^i, with $i \in \{1,..,M_a\}$. In this formulation, actions a^i such that $F^i\left(a^i\right) < f^i\left(a^i\right)$ are more likely to be allocated to the principal and actions a^i such that $F^i\left(a^i\right) > f^i\left(a^i\right)$ are more likely to be allocated to the agent.

Nevertheless, the tendency to allocate decision rights based on relative cost advantage must still be weighed against the revenue-sharing disadvantage of splitting decision rights—the mechanism behind the result in Proposition 1. In particular, even if there are no interaction effects among the various actions and the principal has a strict cost advantage in choosing some transferable actions while the agent has a strict cost advantage in choosing others, it can still be optimal to allocate all decision rights to the same party. We illustrate this point with a linear example in the online appendix.

5 Multiple agents and spillovers

In the previous section, the fundamental source of distortions that drove all of our results and tradeoffs was revenue sharing. This is why it was necessary that at least one action of each type is costly. In this section we extend our model to $N > 1$ agents and allow for a second source of distortions: spillovers due to the transferable actions across the revenues generated by different agents.

In order to analyze the interplay between revenue-sharing and spillovers in the simplest possible way, we assume there is only one non-contractible action of each type, i.e. $M_a = M_q = M_Q = 1$. The revenue function generated jointly by the principal and agent $i \in \{1,..,N\}$ is then $R\left(a_i, \sigma_i, q_i, Q\right)$, where $\sigma_i \equiv \sigma\left(a_1,..,a_{i-1},a_{i+1},..,a_N\right)$ captures the spillovers caused by the choices of transferable actions a_j for $j \neq i$ on the revenues generated by agent i. We assume σ is a symmetric function of $N-1$ arguments with values in $\mathbb{R}_+$ and increasing in all of its arguments. In the examples in Sections 5.1 and 5.2, σ_i will be the average of the other agents' transferable actions, i.e. $\sigma_i = \left(\sum_{j \neq i} a_j\right)/\left(N-1\right)$.

The sign of the spillover is determined by the sign of the derivative of R with respect to its second argument (i.e. R_σ) and can be positive or negative. Consider the following examples from Table 1:

- Hair salons: spillovers from investments in hair styling equipment and supplies can be either positive or negative. A larger investment by an individual hairstylist may draw more traffic to the salon and benefit the other stylists working there (positive spillovers), but it may also lead to business-stealing from the other stylists (negative spillovers).

- Transportation (Uber vs. traditional taxi): spillovers from investments in car quality are likely positive. Better car quality for each individual driver improves the brand image of the entire service in the eyes of customers and therefore helps all other drivers. Business stealing is limited since users rarely, if ever, have the opportunity to choose drivers based on their cars.

- Franchising: spillovers from investments in staff are likely positive. If a given franchisee's staff are more motivated, then they provide a better quality of service to customers—this improves the brand image of the franchisor, which in turn helps all other franchisees. Moreover, business stealing among franchisees is limited, since franchisees usually have a certain degree of territorial exclusivity: consumers choose franchisees based on their location at the point of consumption.

Note that $R^i \equiv R(a_i, \sigma_i, q_i, Q)$ does not depend on the choices of non-transferable actions q_j for other agents $j \neq i$. As we discuss below, introducing this possibility would not add anything meaningful to the tradeoff between the two modes that we focus on.

The principal can choose to control all transferable actions a_i, $i \in \{1, .., N\}$ (i.e. operate in $\mathcal{P}$-mode) or allow each a_i to be chosen by agent i (i.e. operate in $\mathcal{A}$-mode).[11]

In each mode, the principal offers a revenue-sharing contract $\Omega\left(R^i\right)$ to each agent i, where, as in Section 4 above, $\Omega(.)$ can be any arbitrary function of the revenue R^i generated by agent i and the revenue-sharing contract means that agent i obtains $\Omega\left(R^i\right)$, while the principal obtains $R^i - \Omega\left(R^i\right)$.

There are two additional assumptions implicit in this extension of our model to N agents. First, in each mode, the principal offers the same contract to all agents—thus, we rule out price discrimination across agents, which is without loss of generality in our model, given all agents are symmetric. Second, the principal cannot offer an agent a contract contingent on the revenues generated by other agents.

These assumptions are reasonable in the contexts we have in mind. In the examples in Section 3, the principal offers equal terms to all agents and does not normally make payments to an agent conditional on the performance of the other agents (team payments). One reason for the absence of team payments is that agents cannot monitor the revenues generated by other agents—this is especially relevant for the $\mathcal{A}$-mode, in which agents are independent contractors. Furthermore, using team payments in $\mathcal{A}$-mode may raise antitrust concerns given that it could be perceived as a type of agreement between competitors (i.e. rival independent contractors).[12] In $\mathcal{P}$-mode, the principal chooses the a_i's, so has no need for team payments to internalize the spillovers across them. Instead, the principal could try to use team payments to overcome the double-sided moral hazard problem by using one agent as a budget breaker for her contract with other agents. However, in $\mathcal{P}$-mode agents are employees within the same organization and should be able to undo any such attempt by cooperating. In the online appendix we prove formally that such cooperation renders team payments irrelevant.

It is straightforward to show that, after adjusting assumptions (a1)-(a4) accordingly, Lemma 1 continues to apply here, so the principal can achieve the best possible outcome with a linear contract in both modes. The formal proof is in the online appendix.

[11] In the online appendix we explore the hybrid possibility that some but not all agents control their a_i.

[12] Similarly, the fact Uber sets the price of rides has led to an antitrust case against it (see https://www.competitionpolicyinternational.com/us-judge-rules-uber-ceo-must-face-antitrust-lawsuit/).

The interaction between revenue sharing and spillovers creates the possibility of interesting new results. The revenue-sharing distortion implies that we are in a second-best world in both modes. In this context, positive spillovers lead to the a_i's being set too low in $\mathcal{A}$-mode, which exacerbates the revenue-sharing distortion. On the other hand, negative spillovers lead to the a_i's being set too high in $\mathcal{A}$-mode, which can offset the distortion due to revenue-sharing. As we discuss in greater detail in the following two subsections, this possibility has counterintuitive implications for the tradeoff between the two modes. For example, unlike the case with positive spillovers, an increase in the magnitude of negative spillovers can shift the tradeoff in favor of the $\mathcal{A}$-mode.

The two simplest scenarios in which these results appear are:

1. Costly transferable actions a_i and an additively separable revenue function $R(a, \sigma, q, Q)$.

2. Costless transferable actions a_i (namely, prices) and a non-additively separable revenue function $R(a, \sigma, q, Q)$.

The two cases exhibit different mechanisms—we analyze them in the next sections through specific examples. These two cases correspond to realistic scenarios. In many contexts prices are either pinned down by market constraint, or easily observable and contracted upon, which means that they do not have an impact on the $\mathcal{P}$-mode versus $\mathcal{A}$-mode choice. Alternatively, in other cases the optimal price is subject to considerable uncertainty in the contracting stage; or it may be that the principal cannot observe price and quantity separately, so can only offer contracts based on revenue. Price then becomes a relevant transferable and non-contractible variable.

5.1 Linear example

In this section we assume the revenue generated by agent i is

$$R(a_i, \overline{a}_{-i}, q_i, Q) = \beta a_i + x(\overline{a}_{-i} - a_i) + \phi q_i + \Phi Q,$$

where $\overline{a}_{-i}$ is the average of the transferable actions chosen for $j \neq i$, i.e.

$$\overline{a}_{-i} \equiv \frac{\sum_{j \neq i} a_j}{N-1}.$$

Costs are assumed to be quadratic:

$$f(a) = \frac{\beta}{2}a^2, \ \ c(q) = \frac{\phi}{2}q^2 \quad \text{and} \quad C(Q) = \frac{\Phi}{2}Q^2.$$

Consistent with Section 4.2.1, $\phi > 0$ is the weight on the profit generated by the agent's non-transferable investments (and measures the importance of the agent's moral hazard) and $\Phi > 0$ is the weight on the profit generated by the principal's non-transferable investments (and measures the importance of the principal's moral hazard). All of the qualitative results we derive below also hold if the "importance of moral hazard" is measured by the effect of investment on revenue rather than on profit (i.e. ϕ and Φ do not multiply costs). This is because the linear specification shuts down the indirect effect discussed in Section 4.2.1. The results are almost identical: the only change in the expressions for optimal tariffs and profits below is that ϕ and Φ are everywhere replaced with ϕ^2 and Φ^2 respectively.

When spillovers are negative ($x < 0$), revenue R is decreasing in $\overline{a}_{-i}$, which means that in $\mathcal{A}$-mode the transferable actions a_i are set too high. Conversely, when spillovers are positive ($x > 0$), revenue R is increasing in $\overline{a}_{-i}$, so that in $\mathcal{A}$-mode the a_i's are set too low. The particular specification we use in which x multiplies $\overline{a}_{-i} - a_i$ rather than just $\overline{a}_{-i}$ reflects that it is usually not the absolute level of other agents' actions that creates a spillover, but whether an agent's level of action (investment) is above or below the average level of other agents. This normalization also turns out to simplify the analysis.

We assume

$$x < \beta \quad \text{and} \quad x\left(1 - \frac{x}{\beta}\right) < N\Phi, \tag{7}$$

which ensures that (i) $R\left(a_i, \overline{a}_{-i}, q_i, Q\right)$ is increasing in a_i, (ii) all optimization problems are well defined, and (iii) the optimal variable fees in both modes ($t^{\mathcal{P}*}$ and $t^{\mathcal{A}*}$) are strictly between 0 and 1. Note that all $x < 0$ are permissible under (7).

We obtain (all calculations are available in the online appendix)

$$t^{\mathcal{P}*} = \frac{\beta + N\Phi}{\beta + \phi + N\Phi} \text{ and } t^{\mathcal{A}*} = \frac{N\Phi - \frac{x}{\beta}\left(\beta - x\right)}{\frac{1}{\beta}\left(\beta - x\right)^2 + \phi + N\Phi} \tag{8}$$

and the following proposition.[13]

Proposition 5 *The principal prefers the $\mathcal{A}$-mode to the $\mathcal{P}$-mode if and only if*

$$\left|x\frac{\phi}{\beta} + \beta + N\Phi\right| < \sqrt{\beta\left(\beta + \phi + N\Phi\right) + \phi^2}. \tag{9}$$

Consider first the baseline case with no spillovers, i.e. $x = 0$. Then the principal prefers the $\mathcal{A}$-mode to the $\mathcal{P}$-mode if and only if

$$\phi > N\Phi. \tag{10}$$

In other words, the principal prefers the $\mathcal{A}$-mode if the agents' moral hazard is more important than the principal's moral hazard, consistent with the intuition developed in Section 4.2.1. Moreover, the tradeoff does not depend on β, the impact of the transferable action on revenues. The reason is that in both modes the share of revenues retained by the party that chooses the transferable action (t^{E*} in $\mathcal{P}$-mode and $\left(1 - t^{A*}\right)$ in $\mathcal{A}$-mode) is increasing in β. Since t^{E*} and $\left(1 - t^{A*}\right)$ increase at the same rate in this particular example (due to the symmetry of $\mathcal{P}$-mode and $\mathcal{A}$-mode profits in $N\Phi$ and ϕ), the resulting tradeoff does not depend on β.

[13] It is straightforward to verify that the entire range of x defined by (9) is permissible by assumptions (7) if $\beta < 4N\Phi$.

Consider now the tradeoff for general x. If $\beta + N\Phi < \sqrt{\beta(\beta + \phi + N\Phi) + \phi^2}$ (which is equivalent to $\phi > N\Phi$), so that moral hazard considerations favor the $\mathcal{A}$-mode, then the $\mathcal{A}$-mode is preferred if and only if the magnitude of spillovers $|x|$ is not too large. Indeed, for large spillovers, the coordination benefits of the $\mathcal{P}$-mode dominate. On the other hand, if $\phi < N\Phi$, so that moral hazard considerations favor the $\mathcal{P}$-mode, then the $\mathcal{A}$-mode is still preferred for an intermediate, bounded range of negative spillovers. To understand why, note that in $\mathcal{A}$-mode, negative spillovers cause the agents to set their a_i's too high relative to what the principal would like them to choose all else equal. But this implies that in $\mathcal{A}$-mode, negative spillovers help offset to a certain extent the primary revenue distortion, i.e. a_i's being set too low because the party choosing a_i does not receive the full marginal return when $0 < t < 1$. When this offsetting effect is moderately strong (i.e. the magnitude of negative spillovers is not too large), the resulting levels of a_i's are closer to first-best in $\mathcal{A}$-mode than in $\mathcal{P}$-mode, so the $\mathcal{A}$-mode can dominate (this advantage of $\mathcal{A}$-mode must still be traded-off against the moral hazard advantage of the $\mathcal{P}$-mode when $\phi < N\Phi$). When the offsetting effect becomes too strong, the resulting levels of a_i's in $\mathcal{A}$-mode are too far above the first-best levels, so the $\mathcal{P}$-mode dominates again.

Inspection of (9) reveals that the range of spillover values x for which the principal prefers the $\mathcal{A}$-mode is skewed towards negative values, consistent with the explanation in the previous paragraph. Positive spillovers cause the a_i's to be set too low in $\mathcal{A}$-mode, which exacerbates the primary revenue distortion. This makes the $\mathcal{A}$-mode relatively less likely to dominate. There still exists a range of positive spillovers for which the $\mathcal{A}$-mode is preferred provided the agents' moral hazard is more important than that of the principal, but that range is smaller than the corresponding range of negative spillovers.

The skew towards negative values of x in condition (9) also implies that, if spillovers are moderately negative, then an increase in their magnitude (i.e. a *decrease* in x) shifts the trade-off in favor of the $\mathcal{A}$-mode.[14] This result runs counter to the common intuition, according to which spillovers should always make centralized control (i.e. $\mathcal{P}$-mode in our model) more desirable due to the ability to coordinate decisions. The reason behind this counterintuitive result is that, when spillovers are moderately negative and their magnitude increases, the $\mathcal{A}$-mode levels of a_i's get closer to the first-best level through the offsetting effect described above, so the $\mathcal{A}$-mode becomes relatively more attractive (the $\mathcal{P}$-mode levels of a_i's are unchanged). If spillovers are positive or very negative, then an increase in their magnitude moves the $\mathcal{A}$-mode levels of a_i's away from the first-best level, so the standard effect is restored.

We can interpret this result in the context of one of the examples noted in Section 3, namely consultancies. If promoting an individual consultant steals business from the other consultants in the same consulting firm (negative spillovers), then consultants do too much self-promotion when they are independent contractors ($\mathcal{A}$-mode), relative to what the firm would choose, other things equal. But this effect can help compensate for sub-optimal incentives to invest in promotion whenever the commission paid to consultants is less than 100%. In this context, if the business-stealing effect of self-promotion across consultants is moderate, then an increase in its magnitude can make the $\mathcal{A}$-mode

[14] Specifically, if $-\beta - N\Phi < x\frac{\phi}{\beta} < 0$, then condition (9) is more likely to hold when x decreases.

relatively more desirable, by allowing the firm to pay lower commissions while keeping consultants' incentives constant.

Next, we investigate the impact of ϕ and $N\Phi$ on the tradeoff between $\mathcal{A}$-mode and $\mathcal{P}$-mode, by considering their effect on the profit differential $\Pi^{\mathcal{A}*} - \Pi^{\mathcal{P}*}$. From (9), this impact seems difficult to ascertain. Fortunately, one can use first-order conditions and the envelope theorem, which lead to simple conditions (see the online appendix for calculations).

Proposition 6 *A larger ϕ shifts the tradeoff in favor of $\mathcal{A}$-mode (i.e. $\frac{d\left(\Pi^{\mathcal{A}*}-\Pi^{\mathcal{P}*}\right)}{d\phi} > 0$) if and only if $t^{\mathcal{A}*} < t^{\mathcal{P}*}$. A larger Φ shifts the tradeoff in favor of $\mathcal{P}$-mode (i.e. $\frac{d\left(\Pi^{\mathcal{A}*}-\Pi^{\mathcal{P}*}\right)}{d(N\Phi)} < 0$) if and only if $t^{\mathcal{A}*} < t^{\mathcal{P}*}$.*

In other words, the effects of both types of moral hazard on the tradeoff conform to common intuition whenever the share of revenues retained by the principal is larger in $\mathcal{P}$-mode, i.e. $t^{\mathcal{P}*} > t^{\mathcal{A}*}$. Additive separability implies that the total net profit as a function of t is the sum of the profit from each of the actions. Due to the envelope theorem, a change in the parameter ϕ only has a first-order effect on the profit that comes from the agents' non-transferable investments, which is a decreasing function of t (a higher t lowers the agents' incentives to invest). Since agents control their investments in both modes, the expression of the profit that comes from their investments is the same in both modes and therefore ϕ has the expected effect when $t^{\mathcal{P}*} > t^{\mathcal{A}*}$. Similarly for the effect of Φ. From expressions (8), this is always the case in the absence of spillovers ($x = 0$). The problem is that negative spillovers provide a reason to increase $t^{\mathcal{A}*}$, so we can end up with $t^{\mathcal{A}*} > t^{\mathcal{P}*}$. This arises if and only if

$$\frac{x}{\beta} + \frac{\beta}{\beta - x} < -\frac{\beta + N\Phi}{\phi}, \tag{11}$$

i.e. if the spillover x is sufficiently negative.[15] Thus, when the inequality in (11) holds, an increase in the importance of the agents' (respectively, the principal's) moral hazard shifts the trade-off in favor of the $\mathcal{P}$-mode (respectively, $\mathcal{A}$-mode), the opposite of the baseline tradeoff given by (10) for the case without spillovers.

This result runs counter to the common intuition that control should be given to the party whose investment incentives are more important. The interpretation runs as follows. Negative spillovers partially offset the revenue-sharing distortion in $\mathcal{A}$-mode. As a result, a higher t induces less distortion of the transferable actions a_i in $\mathcal{A}$-mode, so the principal can charge a higher t in $\mathcal{A}$-mode, to the point that $t^{\mathcal{A}^*} > t^{\mathcal{P}^*}$ if spillovers are sufficiently negative. However, when this occurs, agents retain a lower share of revenues in $\mathcal{A}$-mode than in $\mathcal{P}$-mode, so their choice of non-transferable effort q_i is *lower* in $\mathcal{A}$-mode. Consequently, when agents' effort (moral hazard) becomes more important in this parameter region, the $\mathcal{P}$-mode becomes relatively more attractive. Similarly, when the principal's investment (moral hazard) becomes more important in the same parameter region, the $\mathcal{A}$-mode becomes relatively more attractive.

[15] Recall that all $x < 0$ are permissible by assumptions (7). Furthermore, it is easily verified that the respective ranges in x defined by (9) and (11) have a non-empty intersection.

Finally, note that the linear example used in this section implicitly assumes that the price to customers is fixed, so is held the same across the two modes, and that there are no production costs. These are not critical assumptions. In the online appendix, we show that Proposition 5 remains unchanged even if the principal chooses price along with the fees (t, T) in its contract, and there are production costs. In other words, the trade-off between the two modes remains the same, even though the profit-maximizing price will differ across the two modes (it is higher for the mode generating higher profits).

5.2 Price as the transferable action

We now turn to the other case of interest identified above—the transferable action is price and therefore does not carry any fixed costs. Specifically, the revenue generated by agent i is now

$$R\left(p_i, \overline{p}_{-i}, q_i, Q\right) = p_i \left(d + \beta p_i + x\left(\overline{p}_{-i} - p_i\right) + \phi q_i + \Phi Q\right), \tag{12}$$

where $d > 0$ is the demand intercept and $\overline{p}_{-i}$ is the average of the prices chosen for $j \neq i$. Thus, the revenue function is not additively separable, although the underlying demand function is. The costs of the non-transferable actions remain the same as in Section 5.1.[16]

To ensure that $R\left(p_i, \overline{p}_{-i}, q_i, Q\right)$ is increasing in p_i and that all optimization problems are well defined, we assume

$$\begin{aligned} \beta &< 0 \\ -2\beta + \min\{0, 2x\} &> \max\{N\Phi, \phi\}. \end{aligned} \tag{13}$$

Note that (13) implies all $x > 0$ are permissible.

From (12), positive spillovers ($x > 0$) correspond to the usual case with prices: when other agents increase their prices, this increases the demand faced by agent i. Also, one could reinterpret p_i as quantity instead of price, but then the usual case would be captured by negative spillovers ($x < 0$).

Define

$$k \equiv \frac{N\Phi\phi}{N\Phi + \phi} \in (0, |\beta|),$$

which can be viewed as a measure of the combined importance of the agent's and principal's moral hazards (k is symmetric in $N\Phi$ and ϕ and increasing in both).

We then obtain the following proposition (calculations are in the online appendix).[17]

Proposition 7 *The principal prefers the $\mathcal{A}$-mode if and only if*

$$-\frac{4k\left(k + \beta\right)}{k + 2\beta} < x < 0. \tag{14}$$

[16] If costs do not depend on ϕ and Φ, then the results below are identical except that the parameters ϕ and Φ are everywhere squared.

[17] Recall $0 < k < -\beta$ so $-\frac{4k(k+\beta)}{k+2\beta} < 0$.

First, note that the proposition identifies a meaningful tradeoff since any positive x and any x satisfying (14) also satisfy (13) provided β is sufficiently negative.

Second, the $\mathcal{P}$-mode is always preferred if spillovers are positive or very negative. The logic here is different from the case with costly transferable actions. Given that the transferable action here (price) does not carry any costs, there is no distortion of price in either mode due to revenue-sharing between the principal and each agent. As a result, the variable fee t can be used in both modes to balance double-sided moral hazard (q_i versus Q) equally well. Furthermore, due to the strategic complementarity between p_i and (q_i, Q), the choice of p_i can either offset or compound the effects of double-sided moral hazard. However, the $\mathcal{P}$-mode has an advantage in internalizing spillovers across the agents' services. This explains why there is a larger region over which the $\mathcal{P}$-mode dominates.

The fact that agents do not internalize spillovers in $\mathcal{A}$-mode can work in favor of the $\mathcal{A}$-mode when spillovers are negative ($x < 0$). Namely, when $x < 0$, the $\mathcal{A}$-mode leads to excessively high choices of p_i, which can help offset the effects of double-sided moral hazard. If this offsetting effect is moderately strong, then the resulting levels of q_i's and Q are closer to first-best in $\mathcal{A}$-mode than in $\mathcal{P}$-mode, so the $\mathcal{A}$-mode dominates. On the other hand, if the offsetting effect is too strong, then negative spillovers over-compensate and the resulting levels of q_i's and Q in $\mathcal{A}$-mode are too far above the first-best levels, so the $\mathcal{P}$-mode is preferred. In contrast, when $x > 0$, the $\mathcal{A}$-mode leads to p_i being set too low, which compounds the effects of double-sided moral hazard. As a result, the $\mathcal{P}$-mode always dominates in that case.

We can interpret the result that negative spillovers across agents' prices (i.e. *positive* demand externalities) can favor the $\mathcal{A}$-mode in the context of the franchising example. As discussed above, due to customer mobility and the over-arching brand of the franchisor, its various franchisees create positive demand externalities on one another. In turn, this implies that independent franchisees tend to set their prices too high relative to what the franchisor would find optimal, so the latter would prefer to control prices (Blair and Lafontaine, 2005, chapter 7). However, our analysis suggests that the excessive prices charged by independent franchisees can help offset the insufficient on-going investments by both the franchisees and the franchisor due to revenue sharing, so giving franchisees discretion over prices will sometimes be preferred.

Third, the parameters measuring the importance of moral hazard for the principal and agent, $N\Phi$ and ϕ, have the same effect on the tradeoff between the two modes (through k). This result stands in contrast to the linear example with costly transferable actions, where $N\Phi$ and ϕ always had opposite effects on the tradeoff between the two modes. The explanation is as follows. Since the transferable action (price) is not distorted by the variable fee t in either mode, both modes do just as well in terms of balancing the double-sided moral hazard problem. Thus, the extent to which the $\mathcal{A}$-mode is preferred over the $\mathcal{P}$-mode when moral hazard becomes more important does not depend on the source of the moral hazard, but only on its aggregate magnitude, measured by k.

6 Conclusion

By substantially reducing the costs of remote monitoring, communication, and collaboration, Internet and mobile technologies have made it possible to build marketplaces and platforms for a rapidly increasing variety of services, ranging from house cleaning to programming, consulting, and legal advice. Consequently, the choice facing firms of whether to control the provision of services to customers by employing workers or whether to enable independent contractors to take control of service provision, and the associated tradeoffs that we have examined in this paper are becoming increasingly relevant in a growing number of sectors throughout the economy.

At the most fundamental level, we have shown that the tradeoffs arise from the need to balance double-sided moral hazard, while at the same time minimizing distortions in the choice of transferable actions due to revenue sharing. This implies that low-powered incentives (control over the transferable actions) should be aligned with high-powered incentives (higher share of revenues), a mechanism which underlies most of our key results. Spillovers across the transferable decisions of different agents introduce an additional source of distortions. Depending on whether the spillover-induced distortion exacerbates the revenue-sharing distortion or offsets it, spillovers may shift the baseline tradeoff in favor of the $\mathcal{P}$-mode (as standard intuition would suggest) or in favor of the $\mathcal{A}$-mode (with counter-intuitive consequences).

Our analysis is also relevant to current legal and regulatory debates about whether "sharing economy" service marketplaces (e.g. Handy, Lyft, Postmates, Uber) should be forced to treat professionals that work with them as employees rather than as independent contractors.[18] All existing legal definitions emphasize control rights as the most important factor in determining this issue, which is consistent with our analysis. However, the assignment of fixed costs to one party or the other would seem to matter much less for the distinction. Indeed, a simple observation based on our model is that any fixed cost that would be borne by the firm when it employs workers and by independent contractors otherwise (e.g. health insurance and worker tax filings) makes no difference to the firm's choice of which party to give decision rights to. Even if independent contractors incur the fixed cost, it must be internalized through a lower fixed fee charged to these agents because the agents' outside option remains unchanged. However, in practice, some of the fixed costs incurred by the firm when it employs workers are not pure transfers to workers (e.g. compliance costs). This drives a fixed cost wedge between the two modes, shifting the tradeoff in favor of the use of independent contractors that control relevant decisions, but other considerations in our model continue to hold.

Needless to say, there are other considerations that are relevant to the policy debate regarding the proper boundary between employees and independent contractors, but are not captured in our model: the impact of the work being done and the control rights associated with it on agents' outside options, the intensity of competition faced by the firms, and the efficiency effects of different regulatory and tax regimes on firms' choices between the two options. Incorporating some of these aspects into the analysis provides a promising avenue for future research.

[18] See for example Justin Fox "Uber and the Not-Quite-Independent Contractor" Bloomberg, June 30, 2015.

There are several other directions in which our work can be extended. One would be to introduce uncertainty and allow that agents are risk averse or wealth-constrained, so that they cannot pay large fixed fees upfront. This would increase the firm's optimal share of revenues regardless of the allocation of control rights, and so should shift the tradeoff in favor of the firm controlling transferable decisions ($\mathcal{P}$-mode). Similarly, one could allow agents to have positive bargaining power and assume that fixed fees are constrained (due to risk-aversion, budget constraints, or other reasons). In this setting, the optimal revenue sharing would be determined by the interaction between double-sided moral hazard considerations and the relative bargaining power of the firm and the agents. Finally, it would be interesting to study competition among firms that can each choose their optimal allocation of control rights over transferable decisions, potentially leading to equilibria in which firms compete with different allocations of control rights.

References

[1] Aghion, P. and J. Tirole (1997) "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1), 1-29.

[2] Alchian, A.A. and H. Demsetz (1972) "Production, Information Costs, and Economic Organization," *American Economic Review*, 62, 777-795.

[3] Alonso, R., W. Dessein and N. Matouschek (2008) "When does coordination require centralization?" *American Economic Review*, 98(1), 145–179.

[4] Anderson, E. (1985) "The Salesperson as Outside Agent or Employee: A Transaction Cost Analysis." *Marketing Science*, 4(3), 234-54.

[5] Baker, G.P. and T.N. Hubbard (2004) "Contractibility and Asset Ownership: On-Board Computers and Governance in U. S. Trucking," *Quarterly Journal of Economics*, 119(4), 1443-1479.

[6] Bester, H. and D. Krahmer (2008) "Delegation and incentives," *RAND Journal of Economics*, 39(3), 664-682.

[7] Bhattacharyya, S. and F. Lafontaine (1995) "Double-Sided Moral Hazard and the Nature of Share Contracts," *RAND Journal of Economics*, 26(4), 761-781.

[8] Blair, R. D. and F. Lafontaine (2005) *The Economics of Franchising*, New York, NY: Cambridge University Press.

[9] Coase, R. (1937) "The Nature of the Firm," *Economica*, 4, 386-405.

[10] Eswaran, M. and A. Kotwal (1984) "The Moral Hazard of Budget-Breaking," *RAND Journal of Economics*, 15(4), 578-581.

[11] Grossman, S. and O. Hart (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Ownership," *Journal of Political Economy*, 94(4), 691–719.

[12] Hagiu, A. and J. Wright (2015a) "Marketplace or Reseller?" *Management Science*, 61(1), 184–203.

[13] Hagiu, A. and J. Wright (2015b) "Multi-sided Platforms" *International Journal of Industrial Organization*, 43, 162–174.

[14] Hart, O. and J. Moore (1990) "Property rights and the nature of the firm," *Journal of Political Economy,* 98, 1119–1158.

[15] Holmstrom, B. (1982) "Moral Hazard in Teams," *Bell Journal of Economics*, 13(2), 324-340.

[16] Holmstrom, B. and P. Milgrom (1994) "The Firm as an Incentive System," *American Economic Review*, 84(4), 972-991.

[17] Romano, R. E. (1994) "Double Moral Hazard and Resale Price Maintenance," *RAND Journal of Economics*, 25(3), 455–466.

[18] Terpstra, A. (1998) *Tenancy and irrigation water management in South-Eastern Punjab, Pakistan*, IIMI Pakistan Report, R-046. Pakistan: International Irrigation Management Institute (IIMI).

[19] Van den Steen, E. (2010a) "Disagreement and the Allocation of Control," *Journal of Law, Economics, & Organization*, 26(2), 385-426.

[20] Van den Steen, E. (2010b) "Interpersonal Authority in a Theory of the Firm," *American Economic Review*, 100(1), 466-490.

[21] Vives, X. (1999) *Oligopoly Pricing: Old Ideas and New Tools*, Cambridge, Mass.: MIT Press.

[22] Waterhouse, W. C. (1983) "Do Symmetric Problems Have Symmetric Solutions?" *American Mathematical Monthly*, 90(6), 378-387.

[23] Wernerfelt, B. (2002) "Why Should the Boss Own the Assets," *Journal of Economics & Management Strategy*, 11(3), 473-485.

7 Appendix

Proof of Lemma 1

Consider first the $\mathcal{P}$-mode. The principal's optimal contract $\Omega^*(.)$ solves

$$\begin{aligned}
\Pi^{\mathcal{P}*} \quad = \quad & \max_{\Omega(.),\mathbf{Q},\mathbf{a},\mathbf{q}} \left\{ R(\mathbf{a},\mathbf{q},\mathbf{Q}) - \Omega(R(\mathbf{a},\mathbf{q},\mathbf{Q})) - f(\mathbf{a}) - C(\mathbf{Q}) \right\} \qquad (15) \\
& \text{s.t.} \\
& \quad \mathbf{a} = \arg\max\nolimits_{\mathbf{a}'} \left\{ R(\mathbf{a}',\mathbf{q},\mathbf{Q}) - \Omega(R(\mathbf{a}',\mathbf{q},\mathbf{Q})) - f(\mathbf{a}') \right\} \\
& \quad \mathbf{q} = \arg\max\nolimits_{\mathbf{q}'} \left\{ \Omega(R(\mathbf{a},\mathbf{q}',\mathbf{Q})) - c(\mathbf{q}') \right\} \\
& \quad \mathbf{Q} = \arg\max\nolimits_{\mathbf{Q}'} \left\{ R(\mathbf{a},\mathbf{q},\mathbf{Q}') - \Omega(R(\mathbf{a},\mathbf{q},\mathbf{Q}')) - C(\mathbf{Q}') \right\} \\
& \quad 0 \leq \Omega(R(\mathbf{a},\mathbf{q},\mathbf{Q})) - c(\mathbf{q}).
\end{aligned}$$

Let $(\mathbf{a}^*, \mathbf{q}^*, \mathbf{Q}^*)$ denote the outcome of this optimization problem and define $R^* \equiv R(\mathbf{a}^*, \mathbf{q}^*, \mathbf{Q}^*)$.

In the online appendix we prove the following lemma.

Lemma 2 *$\Omega^*(.)$ is continuous and differentiable at R^*.*

The lemma and program (15) imply that $(\mathbf{a}^*, \mathbf{q}^*, \mathbf{Q}^*)$ solve

$$\begin{cases} (1 - \Omega_R^*(R^*)) R_{a^i}(\mathbf{a}^*, \mathbf{q}^*, \mathbf{Q}^*) = f_{a^i}^i\left(a^{*i}\right) \text{ for } i \in \{1, .., M_a\} \\ \Omega_R^*(R^*) R_{q^j}(\mathbf{a}^*, \mathbf{q}^*, \mathbf{Q}^*) = c_{q^j}^j\left(q^{*j}\right) \text{ for } j \in \{1, .., M_q\} \\ (1 - \Omega_R^*(R^*)) R_{Q^k}(\mathbf{a}^*, \mathbf{q}^*, \mathbf{Q}^*) = C_{Q^k}\left(Q^{k*}\right) \text{ for } k \in \{1, .., M_Q\}. \end{cases}$$

Let then $t^* \equiv 1 - \Omega_R^*(R^*)$ and $T^* \equiv (1 - t^*) R^* - \Omega^*(R^*)$. Clearly, the linear contract $\widehat{\Omega}(R) = (1 - t^*) R - T^*$ can generate the same stage-2 symmetric Nash equilibrium $(\mathbf{a}^*, \mathbf{q}^*, \mathbf{Q}^*)$ as the initial contract $\Omega^*(R)$. Furthermore, both $\Omega^*(R)$ and $\widehat{\Omega}(R)$ cause the agents' participation constraint to bind and therefore result in the same profits for the principal.

A similar proof applies to the case when the principal chooses the $\mathcal{A}$-mode.

7.1 Supermodularity implies (a5)

Recall the definition of $\Pi(\boldsymbol{\tau})$ for any vector $\boldsymbol{\tau} \in [0, 1]^{M_a + M_q + M_Q}$:

$$\Pi(\boldsymbol{\tau}) \equiv R(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau})) - f(\mathbf{a}(\boldsymbol{\tau})) - c(\mathbf{q}(\boldsymbol{\tau})) - C(\mathbf{Q}(\boldsymbol{\tau})), \tag{16}$$

where $(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau}))$ is the unique solution to

$$\begin{cases} \tau^j R_{a^j}(\mathbf{a}, \mathbf{q}, \mathbf{Q}) = f_{a^j}^j\left(a^j\right) & \text{for } j \in \{1, .., M_a\} \\ \tau^{M_a + k} R_{q^k}(\mathbf{a}, \mathbf{q}, \mathbf{Q}) = c_{q^k}^k\left(q^k\right) & \text{for } k \in \{1, .., M_q\} \\ \tau^{M_a + M_q + l} R_{Q^l}(\mathbf{a}, \mathbf{q}, \mathbf{Q}) = C_{Q^l}^l\left(Q^l\right) & \text{for } l \in \{1, .., M_Q\} \end{cases} \tag{17}$$

We wish to prove that if $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ is (weakly) supermodular in all of its arguments, then $\Pi(\boldsymbol{\tau})$ is increasing in τ^i for all i corresponding to a costly action, i.e. in all τ^j such that $j \in \{1, .., M_a\}$ and $f^j \neq 0$, in all $\tau^{M_a + k}$ such that $k \in \{1, .., M_q\}$ and $c^k \neq 0$ and in all $\tau^{M_a + M_q + l}$ such that $l \in \{1, .., M_Q\}$ and $C^l \neq 0$.

We begin by showing that supermodularity implies $(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau}))$ is strictly increasing in τ^i for all i corresponding to a costly action. To do so, note that the solution $(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau}))$ to (17) corresponds to a game in which there are $M_a + M_q + M_Q$ players, and where each player $j \in \{1, .., M_a\}$ sets

$$a^j = \arg\max_a \left\{\tau^j R\left(a^1, .., a^{j-1}, a, a^{j+1}, .., a^{M_a}, \mathbf{q}, \mathbf{Q}\right) - f^j(a)\right\};$$

each player $M_a + k$ for $k \in \{1, .., M_q\}$ sets

$$q^k = \arg\max_q \left\{\tau^{M_a + k} R\left(\mathbf{a}, q^1, .., q^{k-1}, q, q^{k+1}, .., q^{M_q}, \mathbf{Q}\right) - c^k(q)\right\};$$

and each player $M_a + M_q + l$ for $l \in \{1, .., M_Q\}$ sets

$$Q^l = \arg\max_{Q} \left\{ \tau^{M_a+M_q+l} R\left(\mathbf{a}, \mathbf{q}, Q^1, .., Q^{l-1}, Q, Q^{l+1}, .., Q^{M_Q}\right) - C^l(Q) \right\}.$$

Since $R(\mathbf{a}, \mathbf{q}, \mathbf{Q})$ is supermodular in all of its arguments, this game is supermodular, with payoffs having weakly increasing differences in the actions and the parameters $\left(\tau^1, .., \tau^{M_a+M_q+M_Q}\right)$. From standard supermodularity results (Vives, 1999), we know that an increase in any of the parameters $\left(\tau^1, .., \tau^{M_a+M_q+M_Q}\right)$ will increase each of the solutions $a^j(\boldsymbol{\tau})$ for $j \in \{1, .., M_a\}$, $q^k(\boldsymbol{\tau})$ for $k \in \{1, .., M_q\}$ and $Q^l(\boldsymbol{\tau})$ for $l \in \{1, .., M_Q\}$ in a weak sense. To obtain the strict increase, note that if τ^i increases for some i corresponding to a costly action and no $a^j(\boldsymbol{\tau})$, $q^k(\boldsymbol{\tau})$ or $Q^l(\boldsymbol{\tau})$ increases, then no $a^j(\boldsymbol{\tau})$, $q^k(\boldsymbol{\tau})$ or $Q^l(\boldsymbol{\tau})$ can change since none can decrease. But if $(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau}))$ remain unchanged, then, since τ^i is higher and the corresponding cost function is non-zero, the first-order conditions (17) can no longer hold. Thus, at least one $a^j(\boldsymbol{\tau})$ or one $q^k(\boldsymbol{\tau})$ or one $Q^l(\boldsymbol{\tau})$ must strictly increase. Furthermore, at least one costly action $a^j(\boldsymbol{\tau})$ or $q^k(\boldsymbol{\tau})$ or $Q^l(\boldsymbol{\tau})$ must increase, otherwise assumption (a3) would be violated.

Next, for any $i \in \{1, .., M_a + M_q + M_Q\}$ corresponding to a costly action, we have

$$\begin{aligned} \frac{d\Pi(\boldsymbol{\tau})}{d\tau^i} &= \sum_{j=1}^{M_a} \left(R_{a^j}(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau})) - f^j_{a^j}\left(a^j(\boldsymbol{\tau})\right)\right) \frac{da^j}{d\tau^i} + \sum_{k=1}^{M_q} \left(R_{q^k}(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau})) - c^k_{q^k}\left(q^k(\boldsymbol{\tau})\right)\right) \frac{dq^k}{d\tau^i} \\ &\quad + \sum_{l=1}^{M_Q} \left(R_{Q^l}(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau})) - C^l_{Q^l}\left(Q^l(\boldsymbol{\tau})\right)\right) \frac{dQ^l}{d\tau^i} \\ &= \sum_{j=1}^{M_a} \left(1-\tau^j\right) R_{a^j}(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau})) \frac{da^j}{d\tau^i} + \sum_{k=1}^{M_q} \left(1-\tau^{M_a+k}\right) R_{q^k}(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau})) \frac{dq^k}{d\tau^i} \\ &\quad + \sum_{l=1}^{M_Q} \left(1-\tau^{M_a+M_q+l}\right) R_{Q^l}(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau})) \frac{dQ^l}{d\tau^i}, \end{aligned}$$

where we have used (17) to replace $f^j_{a^j}\left(a^j(\boldsymbol{\tau})\right)$, $c^k_{q^k}\left(q^k(\boldsymbol{\tau})\right)$ and $C^l_{Q^l}\left(Q^l(\boldsymbol{\tau})\right)$. Assumption (a2) implies $R_{a^j} \geq 0$, $R_{q^k} \geq 0$ and $R_{Q^l} \geq 0$ for all $j \in \{1, .., M_a\}$, $k \in \{1, .., M_q\}$ and $l \in \{1, .., M_Q\}$ (with strict inequalities for costly actions). Furthermore, since $(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau}))$ is strictly increasing in τ^i, we have $\frac{da^j(\boldsymbol{\tau})}{d\tau^i} \geq 0$, $\frac{dq^k(\boldsymbol{\tau})}{d\tau^i} \geq 0$ and $\frac{dQ^l(\boldsymbol{\tau})}{d\tau^l} \geq 0$ for all $j \in \{1, .., M_a\}$, $k \in \{1, .., M_q\}$ and $l \in \{1, .., M_Q\}$, with at least one strict inequality. Finally, one of these strict inequalities occurs for a costly action—suppose it's a^j. This implies $\left(1-\tau^j\right) R_{a^j}(\mathbf{a}(\boldsymbol{\tau}), \mathbf{q}(\boldsymbol{\tau}), \mathbf{Q}(\boldsymbol{\tau})) \frac{da^j}{d\tau^i} > 0$, therefore we can conclude that $\frac{d\Pi(\boldsymbol{\tau})}{d\tau^i} > 0$ for any $\tau \in [0,1)^{M_a+M_q+M_Q}$.

Proof of Proposition 1

Suppose the principal chooses t and $D \subset \{1, .., M_a\}$ as the subset of transferable decisions that it controls (the agent is therefore given control over decisions $j \in \{1, .., M_a\} \backslash D$). Then the profit obtained by the principal is equal to $\Pi(\boldsymbol{\tau}(D,t))$, where $\Pi(\boldsymbol{\tau})$ is defined by (16) and $\boldsymbol{\tau}(D,t)$ is the vector of $M_a + M_q + M_Q$ coordinates defined as follows

$$\tau (D,t)^i = \begin{cases} t & \text{if} \quad i \in D \cup \{M_a + M_q + 1, .., M_a + M_q + M_Q\} \\ 1-t & \text{if} \quad i \in (\{1, .., M_a\} \backslash D) \cup \{M_a + 1, .., M_a + M_q\}. \end{cases}$$

Consider part (1) of the Proposition first. Denote by t^* the optimal variable fee and by $(D^*, \{1, .., M_a\} \backslash D^*)$ the optimal allocation of control rights over the transferable actions. Suppose $D^* \neq \emptyset$ and $D^* \neq \{1, .., M_a\}$. If $t^* < 1 - t^*$ (i.e. $t^* < 1/2$), then the principal could weakly increase profits by giving up control over all actions a_j for $j \in D^*$ to the agent and keeping t^* unchanged—strictly if at least one of the actions a_j for $j \in D^*$ is costly. To see this, note that the change in profits is $\Pi(\boldsymbol{\tau}(\emptyset, t^*)) - \Pi(\boldsymbol{\tau}(D^*, t^*))$. If $t^* > 0$ and one of the actions a_j for $j \in D^*$ is costly, then (a5) implies this difference is positive, because $0 < t^* < 1 - t^* < 1$ and $D^* \neq \emptyset$ imply $\boldsymbol{\tau}(D^*, t^*) \in [0,1)^{M_a + M_q + M_Q}$, $\boldsymbol{\tau}(\emptyset, t^*) \in [0,1)^{M_a + M_q + M_Q}$ and $\boldsymbol{\tau}(\emptyset, t^*) > \boldsymbol{\tau}(D^*, t^*)$. If $t^* > 0$ and all actions a_j for $j \in D^*$ are costless, then switching control over them to the agent has no impact, so the resulting profit is unchanged. If $t^* = 0$, then the change in profits can be written

$$\begin{aligned} \Pi(\boldsymbol{\tau}(\emptyset, 0)) - \Pi(\boldsymbol{\tau}(D^*, 0)) &= \max_{\mathbf{a}, \mathbf{q}} \{R(\mathbf{a}, \mathbf{q}, \mathbf{Q} = \mathbf{0}) - f(\mathbf{a}) - c(\mathbf{q})\} \\ &\quad - \begin{pmatrix} \max_{\mathbf{a}, \mathbf{q}} \{R(\mathbf{a}, \mathbf{q}, \mathbf{Q} = \mathbf{0}) - f(\mathbf{a}) - c(\mathbf{q})\} \\ \text{s.t. } a^j = 0 \text{ if } j \in D^* \end{pmatrix}. \end{aligned}$$

In this case, $D^* \neq \emptyset$ and assumptions (a1)-(a4) imply $\Pi(\boldsymbol{\tau}(\emptyset, 0)) - \Pi(\boldsymbol{\tau}(D^*, 0)) > 0$.

By a symmetric argument, if $t^* > 1 - t^*$, then the principal could weakly increase profits by taking control over all actions $j \in \{1, .., M_a\} \backslash D^*$ and keeping t^* unchanged—strictly if at least one of the actions $j \in \{1, .., M_a\} \backslash D^*$ is costly.

Finally, if $t^* = 1/2$, then any allocation of control rights yields the same payoffs, so the pure modes remain weakly optimal.

Let us now turn to part (2) of the proposition. Consider the $\mathcal{P}$-mode. If $t^{\mathcal{P}*} < 1/2$, then the principal could strictly increase profits by giving up control over the transferable actions to the agent and keeping the variable fee unchanged, equal to $t^{\mathcal{P}*}$. To see this, note that the change in profits is $\Pi\left(\boldsymbol{\tau}\left(\emptyset, t^{\mathcal{P}*}\right)\right) - \Pi\left(\boldsymbol{\tau}\left(\{1, .., M_a\}, t^{\mathcal{P}*}\right)\right)$. If $t^{\mathcal{P}*} > 0$, then (a5) implies this difference is positive (recall at least one of the transferable actions is costly), because $0 < t^{\mathcal{P}*} < 1/2$ implies $\boldsymbol{\tau}\left(\emptyset, t^{\mathcal{P}*}\right) \in (0,1)^{M_a + M_q + M_Q}$, $\boldsymbol{\tau}\left(\{1, .., M_a\}, t^{\mathcal{P}*}\right) \in (0,1)^{M_a + M_q + M_Q}$ and $\boldsymbol{\tau}\left(\emptyset, t^{\mathcal{P}*}\right) > \boldsymbol{\tau}\left(\{1, .., M_a\}, t^{\mathcal{P}*}\right)$. If $t^{\mathcal{P}*} = 0$, then the change in profits is

$$\begin{aligned} & \Pi(\boldsymbol{\tau}(\emptyset, 0)) - \Pi(\boldsymbol{\tau}(\{1, .., M_a\}, 0)) \\ = \; & \max_{\mathbf{a}, \mathbf{q}} \{R(\mathbf{a}, \mathbf{q}, \mathbf{Q} = \mathbf{0}) - f(\mathbf{a}) - c(\mathbf{q})\} - \max_{\mathbf{q}} \{R(\mathbf{a} = \mathbf{0}, \mathbf{q}, \mathbf{Q} = \mathbf{0}) - c(\mathbf{q})\}, \end{aligned}$$

which is positive due to assumptions (a1)-(a4). Thus, if $t^{\mathcal{P}*} < 1/2$, then

$$\Pi^{\mathcal{P}*} = \Pi\left(\boldsymbol{\tau}\left(\{1, .., M_a\}, t^{\mathcal{P}*}\right)\right) < \Pi\left(\boldsymbol{\tau}\left(\emptyset, t^{\mathcal{P}*}\right)\right) \leq \Pi^{\mathcal{A}*}.$$

If $t^{\mathcal{P}*} = 1/2$, then

$$\Pi^{\mathcal{P}*} = \Pi\left(\boldsymbol{\tau}\left(\{1,..,M_a\},t^{\mathcal{P}*}\right)\right) = \Pi\left(\boldsymbol{\tau}\left(\emptyset,t^{\mathcal{P}*}\right)\right) \leq \Pi^{\mathcal{A}*}.$$

By a symmetric argument, $t^{\mathcal{A}*} > 1/2$ implies that $\Pi^{\mathcal{A}*} < \Pi^{\mathcal{P}*}$ and $t^{\mathcal{A}*} = 1/2$ implies $\Pi^{\mathcal{A}*} \leq \Pi^{\mathcal{P}*}$.

7.2 Proofs of Propositions 2 and 3

We use some common notation for both proofs. For all $(w,z) \in [0,1]^2$, define

$$\begin{aligned}
\pi^a(\mathbf{a}) &\equiv R^a(\mathbf{a}) - f(\mathbf{a}) \text{ and } \mathbf{a}(z) \equiv \arg\max_{\mathbf{a}}\{zR^a(\mathbf{a}) - f(\mathbf{a})\} \\
\pi^q(\mathbf{q},w) &\equiv wR^q(\mathbf{q}) - c(\mathbf{q}) \text{ and } \mathbf{q}(z) \equiv \arg\max_{\mathbf{q}}\{zR^q(\mathbf{q}) - c(\mathbf{q})\} \\
\pi^Q(\mathbf{Q},w) &\equiv wR^Q(\mathbf{Q}) - C(\mathbf{Q}) \text{ and } \mathbf{Q}(z) \equiv \arg\max_{\mathbf{Q}}\left\{zR^Q(\mathbf{Q}) - C(\mathbf{Q})\right\}.
\end{aligned}$$

Assumption (a5) implies $\pi^a(\mathbf{a}(z))$ is increasing in z, while $\pi^q(\mathbf{q}(z),w)$ and $\pi^Q(\mathbf{Q}(z),w)$ are increasing in z for all $w \geq z$.

7.2.1 Proof of Proposition 2

The principal's profits in each mode as functions of (ϕ,Φ) can be written

$$\begin{aligned}
\Pi^{\mathcal{P}*}(\phi,\Phi) &\equiv \max_t\left\{\pi^a(\mathbf{a}(t)) + \phi\pi^q(\mathbf{q}(1-t),1) + \Phi\pi^Q(\mathbf{Q}(t),1)\right\} \\
\Pi^{\mathcal{A}*}(\phi,\Phi) &\equiv \max_t\left\{\pi^a(\mathbf{a}(1-t)) + \phi\pi^q(\mathbf{q}(1-t),1) + \Phi\pi^Q(\mathbf{Q}(t),1)\right\}.
\end{aligned}$$

Also, denote by $t^{\mathcal{P}*}(\phi,\Phi)$ and $t^{\mathcal{A}*}(\phi,\Phi)$ the corresponding optimal variable fees

$$\begin{aligned}
t^{\mathcal{P}*}(\phi,\Phi) &\equiv \arg\max_t\left\{\pi^a(\mathbf{a}(t)) + \phi\pi^q(\mathbf{q}(1-t),1) + \Phi\pi^Q(\mathbf{Q}(t),1)\right\} \\
t^{\mathcal{A}*}(\phi,\Phi) &\equiv \arg\max_t\left\{\pi^a(\mathbf{a}(1-t)) + \phi\pi^q(\mathbf{q}(1-t),1) + \Phi\pi^Q(\mathbf{Q}(t),1)\right\}.
\end{aligned}$$

Because $\pi^a(\mathbf{a}(t))$, $\pi^q(\mathbf{q}(t),1)$ and $\pi^Q(\mathbf{Q}(t),1)$ are increasing in t, we have $t^{\mathcal{A}*}(\phi,\Phi) \leq t^{\mathcal{P}*}(\phi,\Phi)$, which implies

$$\pi^Q\left(\mathbf{Q}\left(t^{\mathcal{A}*}(\phi,\Phi)\right),1\right) \leq \pi^Q\left(\mathbf{Q}\left(t^{\mathcal{P}*}(\phi,\Phi)\right),1\right) \text{ and } \pi^q\left(\mathbf{q}\left(1-t^{\mathcal{A}*}(\phi,\Phi)\right),1\right) \geq \pi^q\left(\mathbf{q}\left(1-t^{\mathcal{P}*}(\phi,\Phi)\right),1\right).$$

Thus, using the envelope theorem, we have

$$\frac{d\left(\Pi^{\mathcal{A}*}(\phi,\Phi) - \Pi^{\mathcal{P}*}(\phi,\Phi)\right)}{d\phi} = \pi^q\left(\mathbf{q}\left(1-t^{\mathcal{A}*}(\phi,\Phi)\right),1\right) - \pi^q\left(\mathbf{q}\left(1-t^{\mathcal{P}*}(\phi,\Phi)\right),1\right) \geq 0 \quad (18)$$

$$\frac{d\left(\Pi^{\mathcal{A}*}(\phi,\Phi) - \Pi^{\mathcal{P}*}(\phi,\Phi)\right)}{d\Phi} = \pi^Q\left(\mathbf{Q}\left(t^{\mathcal{A}*}(\phi,\Phi)\right),1\right) - \pi^Q\left(\mathbf{Q}\left(t^{\mathcal{P}*}(\phi,\Phi)\right),1\right) \leq 0. \quad (19)$$

If $t^{\mathcal{A}*}(\phi,\Phi) < t^{\mathcal{P}*}(\phi,\Phi)$, then the last two inequalities are strict, so any increase in ϕ (respectively, in Φ) shifts the tradeoff in favor of the $\mathcal{A}$-mode (respectively, $\mathcal{P}$-mode).

If $t^{\mathcal{A}*}(\phi,\Phi) = t^{\mathcal{P}*}(\phi,\Phi)$, then there are only two possibilities:

- $t^{\mathcal{A}*}(\phi,\Phi) = t^{\mathcal{P}*}(\phi,\Phi) = 0$ (which occurs when ϕ is very large relative to Φ), so $\Pi^{\mathcal{P}*}(\phi,\Phi) < \Pi^{\mathcal{A}*}(\phi,\Phi)$. Combined with (18), this implies that any increase in ϕ shifts the tradeoff in favor of the $\mathcal{A}$-mode.
- $t^{\mathcal{A}*}(\phi,\Phi) = t^{\mathcal{P}*}(\phi,\Phi) = 1$ (which occurs when Φ is very large relative to ϕ), so $\Pi^{\mathcal{P}*}(\phi,\Phi) > \Pi^{\mathcal{A}*}(\phi,\Phi)$. Combined with (19), this implies that any increase in Φ shifts the tradeoff in favor of the $\mathcal{P}$-mode.

7.2.2 Proof of Proposition 3

The principal's profits in each mode as functions of (ϕ,Φ) can now be written

$$\begin{aligned}
\widetilde{\Pi}^{\mathcal{P}*}(\phi,\Phi) &\equiv \max_t \left\{\pi^a(\mathbf{a}(t)) + \pi^q(\mathbf{q}((1-t)\phi),\phi) + \pi^Q(\mathbf{Q}(t\Phi),\Phi)\right\} \\
\widetilde{\Pi}^{\mathcal{A}*}(\phi,\Phi) &\equiv \max_t \left\{\pi^a(\mathbf{a}(1-t)) + \pi^q(\mathbf{q}((1-t)\phi),\phi) + \pi^Q(\mathbf{Q}(t\Phi),\Phi)\right\}.
\end{aligned}$$

Meanwhile, the corresponding optimal variable fees are

$$\begin{aligned}
\widetilde{t}^{\mathcal{P}*}(\phi,\Phi) &\equiv \arg\max_t \left\{\pi^a(\mathbf{a}(t)) + \pi^q(\mathbf{q}((1-t)\phi),\phi) + \pi^Q(\mathbf{Q}(t\Phi),\Phi)\right\} \\
\widetilde{t}^{\mathcal{A}*}(\phi,\Phi) &\equiv \arg\max_t \left\{\pi^a(\mathbf{a}(1-t)) + \pi^q(\mathbf{q}((1-t)\phi),\phi) + \pi^Q(\mathbf{Q}(t\Phi),\Phi)\right\}.
\end{aligned}$$

Because $\pi^a(\mathbf{a}(t))$**,** $\pi^q(\mathbf{q}((1-t)\phi),\phi)$ and $\pi^Q(\mathbf{Q}(t\Phi),\Phi)$ are increasing in t (note that $(1-t)\phi \leq \phi$ and $t\Phi \leq \Phi$**)**, we have $\widetilde{t}^{\mathcal{A}*}(\phi,\Phi) \leq \widetilde{t}^{\mathcal{P}*}(\phi,\Phi)$, with equality only if $\widetilde{t}^{\mathcal{A}*}(\phi,\Phi) = \widetilde{t}^{\mathcal{P}*}(\phi,\Phi) = 0$ or $\widetilde{t}^{\mathcal{A}*}(\phi,\Phi) = \widetilde{t}^{\mathcal{P}*}(\phi,\Phi) = 1$. Using the envelope theorem, we have

$$\frac{d\left(\widetilde{\Pi}^{\mathcal{A}*}(\phi,\Phi) - \widetilde{\Pi}^{\mathcal{P}*}(\phi,\Phi)\right)}{d\phi} = \frac{d\pi^q(\mathbf{q}((1-t)\phi),\phi)}{d\phi}_{t=\widetilde{t}^{\mathcal{A}*}(\phi,\Phi)} - \frac{d\pi^q(\mathbf{q}((1-t)\phi),\phi)}{d\phi}_{t=\widetilde{t}^{\mathcal{P}*}(\phi,\Phi)}$$

Thus, by the same reasoning as above, a sufficient condition for an increase in ϕ to shift the tradeoff in favor of the $\mathcal{A}$-mode is that $\frac{d\pi^q(\mathbf{q}(t\phi),\phi)}{d\phi}$ increases with t. We have

$$\begin{aligned}
\frac{d\pi^q(\mathbf{q}(t\phi),\phi)}{d\phi} &= R^q(\mathbf{q}(t\phi)) + \sum_{j=1}^{M_q}\left(\phi R^q_{q^j}(\mathbf{q}(t\phi)) - c^j_{q^j}\left(q^j(t\phi)\right)\right) t q^j_z(t\phi) \\
&= R^q(\mathbf{q}(t\phi)) + t(1-t)\phi\sum_{j=1}^{M_q} R^q_{q^j}(\mathbf{q}(t\phi))\, q^j_z(t\phi).
\end{aligned}$$

Taking the derivative in t, we obtain

$$\frac{d^2\pi^q(\mathbf{q}(t\phi),\phi)}{d\phi dt} = \sum_{j=1}^{M_q}\left(\begin{array}{c} 2\phi(1-t) R^q_{q^j}(\mathbf{q}(t\phi))\, q^j_z(t\phi) + t(1-t)\phi^2 R^q_{q^j}(\mathbf{q}(t\phi))\, q^j_{zz}(t\phi) \\ +t(1-t)\phi^2 R^q_{q^j q^j}(\mathbf{q}(t\phi))\left(q^j_z(t\phi)\right)^2 + t(1-t)\phi^2\sum_{k\neq j} R^q_{q^j q^k}(\mathbf{q}(t\phi))\, q^j_z(t\phi)\, q^k_z(t\phi) \end{array}\right)$$

For all $j \in \{1,..,M_q\}$, the first order condition defining $q^j(t\phi)$ is

$$t\phi R^q_{q^j}(\mathbf{q}(t\phi)) - c^j_{q^j}\left(q^j(t\phi)\right) = 0.$$

Differentiating with respect to t, we obtain

$$\phi R^q_{q^j}(\mathbf{q}(t\phi)) + t\phi^2 R^q_{q^j q^j}(\mathbf{q}(t\phi))\, q^j_z(t\phi) + t\phi^2 \sum_{k\neq j} R^q_{q^j q^k}(\mathbf{q}(t\phi))\, q^k_z(t\phi) = \phi c^j_{q^j q^j}\left(q^j(t\phi)\right) q^j_z(t\phi).$$

Multiplying by $(1-t)\, q^j_z(t\phi)$, this is equivalent to

$$\begin{aligned} & \phi(1-t) R^q_{q^j}(\mathbf{q}(t\phi))\, q^j_z(t\phi) + t(1-t)\phi^2 R^q_{q^j q^j}(\mathbf{q}(t\phi))\left(q^j_z(t\phi)\right)^2 + t(1-t)\phi^2 \sum_{k\neq j} R^q_{q^j q^k}(\mathbf{q}(t\phi))\, q^j_z(t\phi)\, q^k_z(t\phi) \\ = \; & \phi c^j_{q^j q^j}\left(q^j(t\phi)\right)\left(q^j_z(t\phi)\right)^2. \end{aligned}$$

We can now plug this back into the last expression of $\frac{d^2\pi^q(\mathbf{q}(t\phi),\phi)}{d\phi dt}$ to obtain

$$\begin{aligned} \frac{d^2\pi^q(\mathbf{q}(t\phi),\phi)}{d\phi dt} &= \sum_{j=1}^{M_q}\left(\phi(1-t) R^q_{q^j}(\mathbf{q}(t\phi))\, q^j_z(t\phi) + t(1-t)\phi^2 R^q_{q^j}(\mathbf{q}(t\phi))\, q^j_{zz}(t\phi) + \phi c^j_{q^j q^j}\left(q^j(t\phi)\right)\left(q^j_z(t\phi)\right)^2\right) \\ &= \sum_{j=1}^{M_q}\left(\phi(1-t) R^q_{q^j}(\mathbf{q}(t\phi))\left(q^j_z(t\phi) + t\phi q^j_{zz}(t\phi)\right) + \phi c^j_{q^j q^j}\left(q^j(t\phi)\right)\left(q^j_z(t\phi)\right)^2\right) \geq 0, \end{aligned}$$

where the last inequality follows from $c^j_{q^j q^j} \geq 0$ (assumption (a2)), $R^q_{q^j}(\mathbf{q}(t\phi)) \geq 0$ (assumption (a2) along with definition of $\mathbf{q}(t\phi)$) and $q^j_z(t\phi) + t\phi q^j_{zz}(t\phi) \geq 0$ (assumption in the text of Proposition 3).

By a symmetric argument, if $Q^k_z(z) + zQ^k_{zz}(z) \geq 0$ for all $k \in \{1,..,M_Q\}$ and $z \geq 0$, then $\frac{d^2\pi^Q(\mathbf{Q}(t\Phi),\Phi)}{d\Phi dt} \geq 0$, which implies that an increase in Φ shifts the tradeoff in favor of the $\mathcal{P}$-mode.

7.3 Proof of Proposition 4

Suppose first $M_q = 1$. If q^1 is contractible, then

$$\begin{aligned} \widetilde{\Pi}^{\mathcal{P}*} = \; & \max_{t,\mathbf{a},q^1,\mathbf{Q}} \left\{R\left(\mathbf{a}, q^1, \mathbf{Q}\right) - f(\mathbf{a}) - c^1\left(q^1\right) - C(\mathbf{Q})\right\} \\ & \text{s.t.} \\ & \begin{cases} tR_{a^i}\left(\mathbf{a}, q^1, \mathbf{Q}\right) = f^i_{a^i}\left(a^i\right) & \text{for } i \in \{1,..,M_a\} \\ tR_{Q^k}\left(\mathbf{a}, q^1, \mathbf{Q}\right) = C_{Q^k}\left(Q^k\right) & \text{for } k \in \{1,..,M_Q\}, \end{cases} \end{aligned}$$

so the principal can achieve the first best level of profits in $\mathcal{P}$-mode by setting $t = 1$, i.e.

$$\widetilde{\Pi}^{\mathcal{P}*} = \max_{\mathbf{a},q^1,\mathbf{Q}} \left\{R\left(\mathbf{a}, q^1, \mathbf{Q}\right) - f(\mathbf{a}) - c^1\left(q^1\right) - C(\mathbf{Q})\right\}.$$

This is clearly better than the profits that can be achieved in $\mathcal{A}$-mode,

$$\begin{aligned}
\widetilde{\Pi}^{\mathcal{A}*} &= \max_{t,\mathbf{a},q^1,\mathbf{Q}} \left\{ R\left(\mathbf{a}, q^1, \mathbf{Q}\right) - f\left(\mathbf{a}\right) - c^1\left(q^1\right) - C\left(\mathbf{Q}\right) \right\} \\
& \text{s.t.} \\
& \begin{cases} (1-t) R_{a^i}\left(\mathbf{a}, q^1, \mathbf{Q}\right) = f^i_{a^i}\left(a^i\right) \text{ for } i \in \{1, .., M_a\} \\ t R_{Q^k}\left(\mathbf{a}, q^1, \mathbf{Q}\right) = C_{Q^k}\left(Q^k\right) \text{ for } k \in \{1, .., M_Q\}. \end{cases}
\end{aligned}$$

Thus, making q^1 is contractible trivially shifts the tradeoff in favor of the $\mathcal{P}$-mode. By a symmetric argument, if $M_Q = 1$, then making Q^1 is contractible shifts the tradeoff in favor of the $\mathcal{A}$-mode.

Assume now $M_q > 1$, $M_Q > 1$ and R is additively separable in all its arguments, i.e.

$$R\left(\mathbf{a}, \mathbf{q}, \mathbf{Q}\right) = \sum_{i=1}^{M_a} R^{a^i}\left(a^i\right) + \sum_{j=1}^{M_q} R^{q^j}\left(q^j\right) + \sum_{j=1}^{M_Q} R^{Q^k}\left(Q^k\right).$$

For any $t \in [0,1]$, define

$$\begin{aligned}
\pi^{a^i}(a) &\equiv R^{a^i}(a) - f^i(a) \text{ and } a^i(t) \equiv \arg\max_a \left\{ t R^{a^i}(a) - f^i(a) \right\} \text{ for all } i \in \{1, .., M_a\} \\
\pi^{q^j}(q) &\equiv R^{q^j}(q) - c^j(q) \text{ and } q^j(t) \equiv \arg\max_q \left\{ t R^{q^j}(q) - c^j(q) \right\} \text{ for all } j \in \{1, .., M_q\} \\
\pi^{Q^k}(Q) &\equiv R^{Q^k}(Q) - C^k(Q) \text{ and } Q^k(t) \equiv \arg\max_Q \left\{ t R^{Q^k}(Q) - C^k(Q) \right\} \text{ for all } k \in \{1, .., M_Q\}.
\end{aligned}$$

Assumptions (a2)-(a3) imply that $\pi^{a^i}\left(a^i(t)\right)$, $\pi^{q^j}\left(q^j(t)\right)$ and $\pi^{Q^k}\left(Q^k(t)\right)$ are all weakly increasing in t.

Then, when all q^j's are non-contractible, we have

$$\begin{aligned}
\Pi^{\mathcal{P}*} &= \max_t \left\{ \sum_{i=1}^{M_a} \pi^{a^i}\left(a^i(t)\right) + \sum_{j=1}^{M_q} \pi^{q^j}\left(q^j(1-t)\right) + \sum_{k=1}^{M_Q} \pi^{Q^k}\left(Q^k(t)\right) \right\} \\
\Pi^{\mathcal{A}*} &= \max_t \left\{ \sum_{i=1}^{M_a} \pi^{a^i}\left(a^i(1-t)\right) + \sum_{j=1}^{M_q} \pi^{q^j}\left(q^j(1-t)\right) + \sum_{k=1}^{M_Q} \pi^{Q^k}\left(Q^k(t)\right) \right\}
\end{aligned}$$

After q^1 becomes contractible, we have

$$\begin{aligned}
\widetilde{\Pi}^{\mathcal{P}*} &= \max_t \left\{ \sum_{i=1}^{M_a} \pi^{a^i}\left(a^i(t)\right) + \sum_{j=2}^{M_q} \pi^{q^j}\left(q^j(1-t)\right) + \sum_{k=1}^{M_Q} \pi^{Q^k}\left(Q^k(t)\right) \right\} + \max_q \left\{ \pi^{q^1}(q) \right\} \\
\widetilde{\Pi}^{\mathcal{A}*} &= \max_t \left\{ \sum_{i=1}^{M_a} \pi^{a^i}\left(a^i(1-t)\right) + \sum_{j=2}^{M_q} \pi^{q^j}\left(q^j(1-t)\right) + \sum_{k=1}^{M_Q} \pi^{Q^k}\left(Q^k(t)\right) \right\} + \max_q \left\{ \pi^{q^1}(q) \right\}
\end{aligned}$$

Define then for all $x \in [0,1]$

$$\begin{aligned}
\widetilde{\Pi}^{\mathcal{P}*}(x) &\equiv \max_t \left\{ \sum_{i=1}^{M_a} \pi^{a^i}\left(a^i(t)\right) + \left(x\pi^{q^1}\left(q^1(1-t)\right) + \sum_{j=2}^{M_q} \pi^{q^j}\left(q^j(1-t)\right) \right) + \sum_{k=1}^{M_Q} \pi^{Q^k}\left(Q^k(t)\right) \right\} \\
&\quad + (1-x)\max_q \left\{ \pi^{q^1}(q) \right\} \\
\widetilde{\Pi}^{\mathcal{A}*}(x) &\equiv \max_t \left\{ \sum_{i=1}^{M_a} \pi^{a^i}\left(a^i(1-t)\right) + \left(x\pi^{q^1}\left(q^1(1-t)\right) + \sum_{j=2}^{M_q} \pi^{q^j}\left(q^j(1-t)\right) \right) + \sum_{k=1}^{M_Q} \pi^{Q^k}\left(Q^k(t)\right) \right\} \\
&\quad + (1-x)\max_q \left\{ \pi^{q^1}(q) \right\}.
\end{aligned}$$

Clearly, $\Pi^{\mathcal{P}*} = \widetilde{\Pi}^{\mathcal{P}*}(1)$, $\widetilde{\Pi}^{\mathcal{P}*} = \widetilde{\Pi}^{\mathcal{P}*}(0)$, $\Pi^{\mathcal{A}*} = \widetilde{\Pi}^{\mathcal{A}*}(1)$ and $\widetilde{\Pi}^{\mathcal{A}*} = \widetilde{\Pi}^{\mathcal{A}*}(0)$. Let also

$$\begin{aligned}
\widetilde{t}^{\mathcal{P}*}(x) &\equiv \arg\max_t \left\{ \sum_{i=1}^{M_a} \pi^{a^i}\left(a^i(t)\right) + \left(x\pi^{q^1}\left(q^1(1-t)\right) + \sum_{j=2}^{M_q} \pi^{q^j}\left(q^j(1-t)\right) \right) + \sum_{k=1}^{M_Q} \pi^{Q^k}\left(Q^k(t)\right) \right\} \\
\widetilde{t}^{\mathcal{A}*}(x) &\equiv \arg\max_t \left\{ \sum_{i=1}^{M_a} \pi^{a^i}\left(a^i(1-t)\right) + \left(x\pi^{q^1}\left(q^1(1-t)\right) + \sum_{j=2}^{M_q} \pi^{q^j}\left(q^j(1-t)\right) \right) + \sum_{k=1}^{M_Q} \pi^{Q^k}\left(Q^k(t)\right) \right\}.
\end{aligned}$$

We have $\widetilde{t}^{\mathcal{P}*}(x) \geq \widetilde{t}^{\mathcal{A}*}(x)$ for all $x \in [0,1]$. This implies

$$\frac{d\left(\widetilde{\Pi}^{\mathcal{P}*}(x) - \widetilde{\Pi}^{\mathcal{A}*}(x)\right)}{dx} = \pi^{q^1}\left(q^1\left(1-\widetilde{t}^{\mathcal{P}*}(x)\right)\right) - \pi^{q^1}\left(q^1\left(1-\widetilde{t}^{\mathcal{A}*}(x)\right)\right) \leq 0.$$

If $\widetilde{t}^{\mathcal{P}*}(x) > \widetilde{t}^{\mathcal{A}*}(x)$ for some $x \in [0,1]$, then $\widetilde{\Pi}^{\mathcal{P}*}(x) - \widetilde{\Pi}^{\mathcal{A}*}(x)$ is decreasing in x on a positive-measure interval of $[0,1]$, so we can directly conclude

$$\widetilde{\Pi}^{\mathcal{P}*} - \widetilde{\Pi}^{\mathcal{A}*} = \widetilde{\Pi}^{\mathcal{P}*}(0) - \widetilde{\Pi}^{\mathcal{A}*}(0) > \widetilde{\Pi}^{\mathcal{P}*}(1) - \widetilde{\Pi}^{\mathcal{A}*}(1) = \Pi^{\mathcal{P}*} - \Pi^{\mathcal{A}*},$$

i.e. making q^1 contractible shifts the tradeoff in favor of the $\mathcal{P}$-mode.

If $\widetilde{t}^{\mathcal{P}*}(x) = \widetilde{t}^{\mathcal{A}*}(x)$ for all $x \in [0,1]$, then there are only two possibilities:

- $\widetilde{t}^{\mathcal{P}*}(x) = \widetilde{t}^{\mathcal{A}*}(x) = 0$ for all $x \in [0,1]$, which implies $\widetilde{\Pi}^{\mathcal{P}*}(x) < \widetilde{\Pi}^{\mathcal{A}*}(x)$ for all $x \in [0,1]$.
- $\widetilde{t}^{\mathcal{P}*}(x) = \widetilde{t}^{\mathcal{A}*}(x) = 1$ for all $x \in [0,1]$, which implies $\widetilde{\Pi}^{\mathcal{P}*}(x) > \widetilde{\Pi}^{\mathcal{A}*}(x)$ for all $x \in [0,1]$.

In both cases, making q^1 contractible trivially shifts the tradeoff in favor of the $\mathcal{P}$-mode.

By a symmetric argument, making Q^1 contractible shifts the tradeoff in favor of the $\mathcal{A}$-mode.