Controlling versus enabling∗

Andrei Hagiu† and Julian Wright‡

January 31, 2016

Abstract

We study the choice that a firm makes between an employment mode, in which the firm controls service provision by employing professionals, sales representatives, or other types of agents, and an agency mode, in which the firm enables agents to provide their services on terms that they control. The choice of mode is determined by the need to balance double-sided moral hazard problems arising from investments that only agents can make and investments that only the firm can make, while at the same time minimizing distortions in decisions that either party can make. Distortions arise due to the need to share revenues and because of spillovers. Surprisingly, increasing the magnitude of negative spillovers across agents can shift the tradeoff in favor of the agency mode, and provided negative spillovers are not too strong, increasing the agents’ (respectively, the firm’s) moral hazard can shift the tradeoff in favor of the employment (respectively, the agency) mode.

JEL classification: D4, L1, L5

Keywords: control rights, employment, independent contractors, platforms, theory of the firm.

1 Introduction

A key decision for many firms is whether to control the provision of services to customers by employing workers or whether to let independent contractors take control of service provision. This decision 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, and Northwestern University for their helpful comments. 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
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 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, equipment, training, and promotion. This contrast motivates our theoretical study of a firm’s choice between two modes of organization—an employment mode versus an agency mode—where the key difference is that the firm gives agents (e.g. professionals) more control rights in the agency mode than in the employment mode.

In these settings, the revenues generated by the firm typically depend on both its ongoing investments as well as those made by the agents. When neither the firm’s nor the agents’ investments are contractible, joint production calls for some sharing of revenues between the firm and the agents to help balance the resulting double-sided moral hazard problem. At the same time, there are other non-contractible decisions, such as pricing, expenditure on equipment, training, and promotion, which also affect revenues, but can be controlled by either the firm or the agents. 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 decision variables: two costly and non-transferable investment decisions—one for the firm and one for the agents—and a transferable decision variable that can be controlled by either the firm or the agents. The allocation of control over the transferable decision variable is what determines the mode of organization in our model. If control is given to the agents, then the firm operates in the agency mode. If instead control is kept by the firm, then it operates in the employment mode. Given double-sided moral hazard, a meaningful tradeoff between the agency mode and the employment mode only exists if the transferable decision variable is costly or exhibits spillovers across multiple agents.

In our model, there is no uncertainty and we assume the firm cannot credibly commit to throw away revenues in either mode and cannot pay independent contractors based on revenues generated by other independent contractors in the agency mode. These assumptions imply that linear contracts are optimal in our setting and that, since revenues must be shared between the firm and the agents to incentivize their respective non-transferable investments, neither the employment nor the agency mode will achieve the first-best. Thus, the firm is choosing between two second-best arrangements, with a primary distortion due to revenue sharing that affects both modes, and a secondary distortion that arises from spillovers across the transferable decisions and is only present in the agency mode. Broadly speaking, the revenue-sharing distortion creates a baseline tradeoff between the two modes.

\footnote{Such “team payments” are not necessary in the employment mode, where the firm has direct control over the transferable decisions that induce spillovers and can therefore internalize them.}
driven by the importance of the firm’s moral hazard relative to that of the agents’. The spillover-induced distortion shifts this baseline tradeoff by either exacerbating the revenue-sharing distortion (which favors the employment mode) or offsetting it (which favors the agency mode).

More specifically, consider first the case when the transferable decision is a revenue-increasing costly investment (e.g. marketing or equipment). If a larger investment by one agent also increases the revenues obtained by other agents providing services through the same firm (i.e. the spillover is positive), then an increase in the magnitude of the spillover always shifts the tradeoff between the two modes in favor of the employment mode, as standard intuition would suggest. This is because in the employment mode, the firm coordinates investment decisions to fully internalize the spillover. However, in the agency mode, the firm can only induce individual agents to partially internalize spillovers by sharing some revenues with them, implying that agents invest too little. Furthermore, the employment (respectively, agency) mode is more likely to be chosen when the firm’s (respectively, the agents’) moral hazard becomes more important—consistent with the predictions of traditional theories of the firm. In our model, the logic is as follows: the party whose moral hazard problem is more important should receive a greater share of the revenue, which implies that the same party should also be given control over the transferable decision variable to lessen the distortion from the costly transferable decision variable being set too low. In other words, low-powered incentives (i.e. control over the transferable decision) should be aligned with high-powered incentives (i.e. higher share of revenues).

This logic changes dramatically with negative spillovers (e.g. more self-promotion by a sales agent of a given manufacturer steals business from the manufacturer’s other sales agents). In the agency mode, individual agents now 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 agency mode can then be a useful way for the firm 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 agency mode. Second, if the magnitude of negative spillovers is sufficiently large, then agents get a lower share of revenues in the agency mode than in the employment mode. This leads to a reversal of the standard logic, which prevailed in the case with positive spillovers.

When the transferable decision variable is the price charged to customers, the tradeoff between the two modes is determined by different considerations. Absent production costs, revenue sharing does not distort price-setting in either mode, so revenue-sharing works equally well in both modes to balance the double-sided moral hazard problem. However, a higher price raises the return to each party from costly investments that increase demand, thereby mitigating each party’s moral hazard problem. When services are substitutes (positive spillovers in our formulation), independent agents set prices too low in the agency mode, which therefore exacerbates moral hazard. As a result, the employment mode dominates. On the other hand, if agents’ services are complements (negative spillovers), then
independent agents set prices too high in the agency mode, thus mitigating each of the moral hazard problems. Consequently, we find that the agency mode can be preferred.

The next section discusses related literature. Section 3 provides some examples of markets in which the choice between employment mode and agency mode is relevant. Section 4 sets up our general model: the case without spillovers is analyzed in Section 5, while the case with spillovers is analyzed in Section 6. We then fully solve two specific examples of the general model in Sections 7 and 8 to obtain further results. Section 9 extends our benchmark model to consider private benefits, different timing, cost asymmetries, and the possibility of hybrid modes. Section 10 concludes the paper.

2 Related literature

Our theory 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) and incentive systems (Holmstrom and Milgrom, 1994): there are non-contractible decisions that can be allocated to the firm or external parties (independent contractors), and the firm must design first-period contracts that incentivize second-period investments by both the agents and the firm. However, our approach is different in two key respects. First, the allocation of decision rights between the firm and agents is not tied to the ownership of any asset, so ours is not a theory of ownership structure. Rather, the employee vs. independent contractor distinction in our model is solely based on whether the firm or the agents are given control over certain non-contractible and transferable decisions that impact payoffs. This is entirely consistent with legal definitions that emphasize control rights as the most important factor determining whether agents should be considered independent contractors or employees, as highlighted by the current debates and lawsuits surrounding Uber and other firms in the “sharing economy”. Second, our setting features spillovers created by the transferable decisions corresponding to each agent on the payoffs generated by the other agents. These spillovers are internalized when the firm controls the transferable decisions in the employment mode, but not when agents control the transferable decisions in the agency mode—this difference plays a key role in the tradeoffs that we study and is not present in formal theories of the firm based on property rights or incentive systems.

Given that our theory is not based on any change in ownership structure and features coordination issues due to spillovers, one may wonder whether it is better thought of as a theory of delegation within organizations (e.g. Aghion and Tirole, 1997, Alonso et al., 2008). For example, in our setting, one could ask: how much control should a firm give its employees and how much of its revenues should it share with them? However, we think that our theory is less suited for analyzing delegation choices

---

Note also that viewing employees as agents who have contractually agreed that control of certain transferable actions (e.g. 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.
within organizations, as in our agency mode, independent contractors are external to the firm, in contrast to employees in the employment mode. This distinction is important when there are multiple agents and spillovers arise between them. If agents are internal to the firm, then the firm should be able to use team-based payments and other mechanisms to induce them to internalize spillovers. This is less plausible for independent contractors because, among other reasons, making payments to each independent contractor conditional on the performance of other independent contractors is likely to be deemed illegal under antitrust law. The inability to use such mechanisms in the agency mode is fundamental to our main results. Furthermore, a key difference relative to the delegation literature is that in our model there is no uncertainty and therefore no information asymmetry. Instead, the allocation of control rights is entirely driven by double-sided moral hazard considerations and spillovers.

In terms of insights, a major contribution of our paper is showing how revenue-sharing between the firm and its agents interacts with spillovers created by transferable decisions across agents to produce results that are sometimes counterintuitive. For instance, negative spillovers can make the firm more likely to prefer the agency mode over the employment mode as the magnitude of spillovers increases or as the firm’s moral hazard becomes more important relative to the agents’ moral hazard. This is the opposite of what one might 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. Our 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 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, the set-ups and mechanisms in these papers are very different from ours: they rely on private information and cheap talk (Alonso et al.) or uncertainty and exogenously given misalignment of objectives (Bester and Krahmer).

Since in our model revenues must be shared between agents and the firm 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 firm or the agents.

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 third-party suppliers’ 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 suppliers’ 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 firms 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 marketplaces (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.

The hair salon industry is a good example, as it has long featured both modes of organization. Some salons employ their hairstylists and pay them fixed hourly wages plus commissions that are a percentage of sales. Such salons control schedules, provide the hair products, do all the marketing to customers, and provide stylists with some training and guidance. 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, stylists decide their own schedules, provide their own hair products, and advertise themselves to customers. The salon owners still make all necessary investments to maintain the facilities, as well as to advertise the salon 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 have to train and promote themselves, and employees, whom the firm trains and promotes. 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 directly impact the revenues they generate (e.g. the quality and maintenance of their particular outlets, and the benefits and training offered to their staff). In contrast, these decisions are made by the firm in company-owned outlets.

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\(^3\) 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 where firms may choose between the two modes 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) a transferable decision that is chosen by the firm in the employment mode and by the agent in the agency mode; (ii) a costly ongoing effort always chosen by the agent; and (iii) a costly ongoing investment always chosen by the firm. We have not included the price charged to customers in Table 1, which is potentially another transferable decision variable in each of the examples listed. This is because the price is sometimes pinned down by market constraints, in which case it can be treated as a fixed constant in our analysis.\(^4\)

<table>
<thead>
<tr>
<th>Table 1: Examples</th>
<th>Transferable decisions</th>
<th>Non-transferable investment decisions made by agents</th>
<th>Non-transferable investment decisions made by the firm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hair salons</td>
<td>hair products; training and promotion of individual hairdressers</td>
<td>service quality</td>
<td>maintenance and advertising of salon</td>
</tr>
<tr>
<td>Transportation (e.g. Uber vs. traditional taxi companies)</td>
<td>car quality and maintenance</td>
<td>service quality</td>
<td>quality of the technological infrastructure (payment, dispatch system); advertising</td>
</tr>
<tr>
<td>Consulting (e.g. Hourly Nerd vs. McKinsey) and outsourcing (e.g. Upwork vs. Infosys)</td>
<td>training</td>
<td>service quality</td>
<td>quality of the (online) system for communication, monitoring and payment; advertising</td>
</tr>
<tr>
<td>Online education (e.g. Coursera vs. University of Phoenix)</td>
<td>curriculum design; advertising of individual instructors and courses</td>
<td>quality of content and its delivery</td>
<td>quality of the online infrastructure; advertising of the site</td>
</tr>
<tr>
<td>Waste and recycling (e.g. Rubicon Global vs. Waste Management)</td>
<td>equipment for waste collection and hauling</td>
<td>service quality</td>
<td>quality of the technological infrastructure (payment, scheduling routes and pickup); advertising</td>
</tr>
<tr>
<td>Producers and sales agents</td>
<td>training and promotion of individual sales agents</td>
<td>sales effort</td>
<td>advertising and quality of the product or service</td>
</tr>
<tr>
<td>Franchising vs. company owned outlets</td>
<td>quality and maintenance of outlets; staff benefits and training</td>
<td>outlet manager effort</td>
<td>quality of the product (franchisor); national advertising</td>
</tr>
<tr>
<td>Sharecropping</td>
<td>quality of inputs (seeds, fertilizer and pesticides); bribes</td>
<td>sharecropper effort</td>
<td>large investment (maintenance of irrigation system)</td>
</tr>
</tbody>
</table>

---

\(^3\)While 50/50 crop sharing is the most common practice, other splits are also used, as documented by Terpstra (1998).

\(^4\)The price may also be set by the firm in its contract with the agent, a case that we discuss in Section 7.
4 The model and preliminaries

There is a firm (the principal) and $N$ symmetric agents. The revenue function generated jointly by
the firm and agent $i$ who joins the firm is identical for all agents and is given by $R(a_i, \sigma_i, q_i, Q)$, which
depends on three types of actions, all of which are non-contractible.

Actions $q_i$ and $Q$ are non-transferable: agent $i$ always chooses $q_i \in \mathbb{R}_+$ at cost $c(q_i)$ and the firm
always chooses $Q \in \mathbb{R}_+$ at cost $C(Q)$. This means there is double-sided moral hazard. One can think
of $q_i$ as the effort or investment made by the agent to raise the quality of the service she provides
and of $Q$ as capturing the firm’s ongoing investments in advertising or a common infrastructure. In
particular, the firm chooses a single $Q$ that impacts the revenues attributable to all $N$ agents.

Action $a_i$ is transferable, i.e. it can be chosen either by the firm or by the agent, depending on the
mode in which the firm chooses to operate. The party that chooses $a_i \in \mathbb{R}_+$ incurs cost $f(a_i)$. Our
analysis encompasses two possibilities:

- Costly actions that always increase revenues, i.e. $f(a_i) > 0$ for $a_i > 0$ and $R$ increasing in $a_i$.
  Examples include investments in equipment, training or promotion of agents (see Table 1).

- Costless actions ($f = 0$), such that $R$ is single-peaked in $a_i$. Price is the most natural example,
  but such actions also include “horizontal choices” (Hagiu and Wright, 2015a), such as the allo-
  cation of a fixed promotional capacity between emphasizing an agent’s previous education and
  work experience and emphasizing her/his performance on recent projects through the firm.

The remaining argument in the revenue function,

$$\sigma_i \equiv \sigma(a_{-i}),$$

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 $\sigma$ is a symmetric function with values in $\mathbb{R}_+$ and increasing in all
of its arguments. Thus, the sign of the spillover is determined by the sign of the derivative of $R$ with
respect to its second argument. For instance, in some circumstances, higher spending by an individual
hairstylist on self-promotion and fancy hair products may draw more traffic to the salon and benefit
the other stylists working there (positive spillovers), whereas in other circumstances it may lead to
business-stealing from the other hairstylists (negative spillovers). In the examples in Sections 7 and
8, $\sigma$ will be the average of the other agents’ transferable actions, i.e.

$$\sigma(a_{-i}) = \frac{\sum_{j \neq i} a_j}{N-1},$$

For convenience, denote

$$s(a) \equiv \sigma(a, ..., a)$$

the spillover imposed on agent $i$’s revenues by all other agents $j \neq i$ choosing the same transferable
action $a_j = a$ (by symmetry, this function does not depend on $i$). Note that $s(a)$ is increasing because $\sigma(a_{-i})$ is increasing in all of its arguments.

Finally, note that $R_i^e \equiv R(a_i, \sigma(a_{-i}), q_i, Q)$ 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 firm can choose to operate in one of two modes: $E$-mode (employment) and $A$-mode (agency). The difference between the two modes is that in $E$-mode, the firm controls all transferable actions $a_i$, $i \in \{1, ..., N\}$, whereas in $A$-mode each $a_i$ is chosen by agent $i$.

In each mode, the firm offers a revenue-sharing contract $\Omega(R_i)$ to each agent $i$, where $\Omega(.)$ can be any arbitrary function of the revenue $R_i$ generated by agent $i$ and the revenue-sharing contract means that agent $i$ obtains $\Omega(R_i)$, while the firm obtains $R_i - \Omega(R_i)$. There are four implicit assumptions in this specification. First, in each mode, the firm 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, only the realized revenue $R_i$ is contractible, whereas the underlying variables $(a_i, q_i, Q)$ are not. Third, the firm cannot commit to “throwing away” revenue in case a target specified ex-ante is not reached (Holmstrom, 1982). And fourth, the firm 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 particular, 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). Moreover, making payments to each independent contractor conditional on the performance of the other independent contractors may raise antitrust concerns.\footnote{This assumption is only relevant for the $A$-mode, since the firm internalizes spillovers in $E$-mode and therefore does not need to make contracts conditional on team performance.}

Given that there is no uncertainty in our model, these assumptions imply that we can restrict attention to linear contracts without loss of generality (this is formally proven below). Bhattacharyya and Lafontaine (1995) provide empirical support for the prevalence of revenue sharing contracts that are linear and uniform across agents, in markets exhibiting double-sided moral hazard, such as franchising and sharecropping.

We assume the firm holds all the bargaining power. This implies that it will set $\Omega(.)$ in both modes so that each agent $i$ is indifferent between participation and its outside option, which for convenience we normalize to zero throughout. Note that in our model it is immaterial whether the firm or agent $i$ collects revenues $R_i$ and pays the other party their share according to contract $\Omega(R_i)$. If in $E$-mode the firm collects revenues and pays $\Omega(R_i)$ to agent $i$, then this can be interpreted as a bonus in an employment relationship.

The game that we study has the following timing. In stage 0, the firm chooses whether to operate in $E$-mode or $A$-mode. In stage 1, the firm sets $\Omega(.)$ and the agents simultaneously decide whether to accept this. In stage 2, there are two possibilities depending on the firm’s choice in stage 0. In $E$-mode, the firm chooses $Q$ and $a_i$ for each participating agent, and each agent simultaneously chooses $q_i$. In $A$-mode, the firm chooses $Q$ and each agent simultaneously chooses $q_i$ and $a_i$. Finally, in stage...
is concave and admits a unique finite maximizer in $(a, R)$. Throughout the paper, we use the following two notational conventions. First, for variables and parameters that apply to both the agents and the firm, we use lowercase for agents and uppercase for the firm (e.g. $q$ and $Q$). Second, subscripts next to functions always indicate derivatives: for example, $f_a$ indicates the derivative of $f$ with respect to $a$, $R_a$ indicates the partial derivative of $R$ with respect to the transferable action $a$ (the first argument) and $R_\sigma$ indicates the partial derivative of $R$ with respect to the spillover $\sigma$ (the second argument). However, subscripts next to variables indicate the choice of variables associated with the $i^{th}$ agent, such as in $a_i$ and $q_i$.

We make the following technical assumptions:

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

(a2) The cost functions $c$ and $C$ are increasing and strictly convex in their arguments. If $f \neq 0$, then $f$ is also increasing and strictly convex. Furthermore,

$$f(0) = f_a(0) = c(0) = c_q(0) = C(0) = C_Q(0) = 0.$$ 

(a3) The revenue function $R(a_i, \sigma_i, q_i, Q)$ is non-negative and strictly increasing in $(q_i, Q)$. If $f \neq 0$, then $R(a_i, \sigma_i, q_i, Q)$ is also strictly increasing in $a_i$ and $\sum_{i=1}^{N} R(a_i, \sigma(a_{-i}), q_i, Q)$ is strictly increasing in each $a_i$, for $i \in \{1, \ldots, N\}$.

(a4) For all $(\sigma_i, Q) \in \mathbb{R}^2$, and $t \in (0, 1]$, $tR(a_i, \sigma_i, q_i, Q) - f(a_i) - c(q_i)$ is concave and admits a unique finite maximizer in $(a_i, q_i)$. For all $t \in (0, 1]$ and $(q_1, \ldots, q_N) \in \mathbb{R}_+^N$,

$$\sum_{i=1}^{N} (tR(a_i, \sigma(a_{-i}), q_i, Q) - f(a_i) - c(q_i)) - C(Q)$$

is concave and admits a unique finite maximizer in $(a_1, \ldots, a_N, Q)$, which is symmetric in $(a_1, \ldots, a_N)$.

(a5) For any $(t_1, t_2, t_3) \in [0, 1]^2$ and $I \in \{0, 1\}$, the following system of equations admits a unique solution $(a, q, Q)$:

$$\begin{cases} 
    t_1(R_a(a, s(a), q, Q) + Is_a(a) R_\sigma(a, s(a), q, Q)) = f_a(a) \\
    t_2 R_q(a, s(a), q, Q) = c_q(q) \\
    t_3 NR_Q(a, s(a), q, Q) = C_Q(Q). 
\end{cases}$$

These assumptions are quite standard and are made to ensure that the optimization problems considered below are well-behaved. Assumption (a3) ensures that the spillover is not so large that it overcomes the direct effect of $a_i$. Assumptions (a4) and (a5) ensure that there is always a unique finite solution to the optimization problems we consider; in particular, they allow us to avoid having to impose stability conditions, which would be quite complex in this set-up (see, for instance, Vives, 1999).

Furthermore, the firm always finds it optimal to induce all $N$ agents to participate and given our completely symmetric setup, it is natural to focus on the symmetric solutions to the optimization
problems that we consider.  

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

**Proposition 1** If assumptions (a1)-(a5) hold, then in both modes the firm can achieve the best possible symmetric outcome with a linear contract.

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

\[ \Omega(R) = (1 - t)R - T, \]

where \( T \) can be interpreted as the fixed fee and \( t \in [0, 1] \) as the share of revenue generated by each agent that the firm keeps. This means the net payoff received by agent \( i \) is \((1 - t)R^i - T\). In general, the respective optimal contracts in the two modes will have different \((t, T)\). Thus, it is possible for \( T \) to be negative under \( \mathcal{E} \)-mode (i.e. the agents receive a fixed wage) and positive under \( \mathcal{A} \)-mode (i.e. the agents pay a fixed fee). Nevertheless, in practice, if the agents’ outside option is high enough, then they would receive a net payment in both modes.

## 5 Baseline case with no spillovers

To build some intuition, it is instructive to first focus on the case when there are no spillovers across agents, i.e. when \( R_{\sigma} = 0 \). In this case, the revenue function depends on only three arguments, so we write it as \( R(a, q, Q) \) for convenience. Proposition 1 and symmetry across \( i \in \{1, ..., N\} \) imply that the firm’s profits in \( \mathcal{E} \)-mode can be written as\(^7\)

\[
\Pi^{\mathcal{E}^*} = \max_{t,a,q,Q} \left\{ N (R(a, q, Q) - f(a) - c(q)) - C_Q(Q) \right\} \tag{1}
\]

\[
\text{s.t.} \begin{cases} 
  tR_a(a, q, Q) = f_a(a) \\
  (1 - t) R_q(a, q, Q) = c_q(q) \\
  tNR_Q(a, q, Q) = C_Q(Q). 
\end{cases} \tag{2}
\]

\(^6\)It is straightforward to show directly that the solutions must be symmetric when there are no spillovers (Section 5) or when the revenue function is additively separable in its arguments (this is the case with the linear example we analyze in Section 7). Waterhouse (1983) provides general conditions for a symmetric constrained optimization problem to have symmetric solutions.

\(^7\)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. \((1 - t) R(a, q, Q) - c(q) - T = 0\).
Similarly, the firm’s A-mode profits are

\[ \Pi^{A*} = \max_{t,a,q,Q} \left\{ N(R(a,q,Q) - f(a) - c(q)) - C(Q) \right\} \]  

s.t. \begin{align*}
(1-t)R_a(a,q,Q) &= f_a(a) \\
(1-t)R_q(a,q,Q) &= c_q(q) \\
tNR_Q(a,q,Q) &= C_Q(Q).
\end{align*} \quad (4)

In general, the respective profits yielded by both modes are lower than the first-best profit level

\[ \max_{a,q,Q} \left\{ N(R(a,q,Q) - f(a) - c(q)) - C(Q) \right\}. \]

The reason is that revenue \( R(a_i,q_i,Q) \) for each \( i \in \{1, \ldots, N\} \) needs to be divided between the firm and agent \( i \) 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 an agent and the firm.

Comparing programs (1) and (3) makes it clear that the difference between the two modes comes from the choice of the non-transferable action \( a \). The tradeoff between the E-mode and the A-mode boils down to whether it is better to align the choice of \( a \) with the firm’s choice of investment \( Q \) (E-mode) or with the agents’ choice of effort \( q \) (A-mode).

**Proposition 2** Compare the firm’s profits under the two modes.

(a) If the transferable action \( a \) is contractible or costless (i.e. \( f = 0 \)), then the two modes are equivalent and lead to the same firm profits \((\Pi^{E*} = \Pi^{A*})\).

(b) Suppose the transferable action \( a \) is non-contractible and costly. If the non-transferable action \( q \) is contractible or if it has no impact on revenue \((R_q = 0)\), then \( \Pi^{E*} > \Pi^{A*} \). If the non-transferable action \( Q \) is contractible or if it has no impact on revenue \((R_Q = 0)\), then \( \Pi^{A*} > \Pi^{E*} \).

Thus, in the absence of spillovers, for there to exist a meaningful tradeoff between the two modes, (i) all three actions must be non-contractible and have a strictly positive impact on revenues \( R \), and (ii) the non-transferable action \( a \) must carry a strictly increasing cost \( f(a) \). Part (a) of the proposition implies that if the transferable action \( a \) is price, then, even if it cannot be contracted on, the two modes are equivalent. As we will see in Section 6, this no longer holds when there are spillovers from the choice of price by each agent on the revenues generated by the other agents.

In the general case of interest, when all three actions are non-contractible, have a positive impact on revenues and carry strictly increasing costs, each mode distorts the choice of \( a \) in a different way, which leads to different profits. Heuristically, if the firm’s moral hazard is more important (in the sense that it has a larger impact on \( R \)), then the optimal \( t \) is higher in both modes, but then the E-mode induces relatively less distortion in \( a \) and is therefore more likely to be preferred. Conversely, if the agents’ moral hazard is more important, then the optimal \( t \) is lower in both modes, so the A-mode induces less distortion in \( a \) and is therefore more likely to be preferred. In other words, low-powered
incentives (i.e. control over a) should be aligned with high-powered incentives (i.e. higher share of revenues). This intuition is confirmed by the linear example analyzed in Section 7 below.

We can derive a useful result for the case in which \( R(a,q,Q) \) is supermodular in its arguments, i.e. the actions \( a, q \) and \( Q \) are (weak) strategic complements. Denote by \( t^{E*} \) and \( t^{A*} \) the respective optimal variable fees charged by the firm in the two modes, i.e. the respective solutions in \( t \) that emerge from programs (1) and (3).

**Proposition 3** Suppose \( a \) is costly (i.e. \( f \neq 0 \)) and \( R(a,q,Q) \) is supermodular in its arguments. Then, \( t^{E*} < 1/2 \) implies \( \Pi^{E*} < \Pi^{A*} \) and \( t^{A*} > 1/2 \) implies \( \Pi^{A*} < \Pi^{E*} \). Furthermore, \( t^{E*} = 1/2 \) implies \( \Pi^{E*} \leq \Pi^{A*} \) and \( t^{A*} = 1/2 \) implies \( \Pi^{A*} \leq \Pi^{E*} \).

In other words, when the three non-contractible actions are (weak) strategic complements, the firm would never find it optimal to function in \( E \)-mode and keep less than 50% of revenue or function in \( A \)-mode and keep more than 50% of revenue. The key driving force behind this result is that reducing the distortions in the firm’s and the agents’ second stage objective functions relative to the firm’s first-stage objective function raises the firm’s profit. For example, if \( t^{E*} < 1/2 \), then the distortions can be reduced by shifting control over the transferable actions from the firm to the agents. Indeed, this changes the first-order condition determining \( a \) in the second stage from

\[
t^{E*}R_a (a,q,Q) = f_a (a)
\]

to

\[
\left(1 - t^{E*}\right)R_a (a,q,Q) = f_a (a).
\]

The other two first-order conditions stay unchanged. Since \( 1 - t^{E*} > t^{E*} \) and the three actions are strategic complements, this change results in higher second-stage equilibrium levels of \((a,q,Q)\). This in turn means the outcome is closer to the first-best and therefore equilibrium profits are higher.

Proposition 3 can be re-stated in a more empirically useful way. To do so, define

\[
t^* \equiv \begin{cases} 
  t^{E*} & \text{if } \Pi^{E*} \geq \Pi^{A*} \\
  t^{A*} & \text{if } \Pi^{E*} < \Pi^{A*},
\end{cases}
\]

which is the optimal variable fee charged by the firm in the optimal mode. The following corollary is a logical reformulation of Proposition 3.

**Corollary 1** Suppose \( a \) is costly (i.e. \( f \neq 0 \)) and \( R(a,q,Q) \) is supermodular in its arguments. Then \( t^* < 1/2 \) if and only if the \( A \)-mode is strictly optimal (i.e. \( \Pi^{A*} > \Pi^{E*} \)) and \( t^* \geq 1/2 \) if and only if the \( E \)-mode is weakly optimal (i.e. \( \Pi^{E*} \geq \Pi^{A*} \)).

Thus, according to this prediction of our model, each agent obtains more than 50% of attributable revenues if and only if the firm is functioning in \( A \)-mode. This means that, other things being equal,
we would expect that organizations that have chosen the \( A \)-mode should leave a larger share of their revenues to agents than organizations that have chosen the \( E \)-mode. For example, salons that rent out chairs charge only a fixed rental fee, letting stylists keep 100% of sales, whereas traditional hair salons that employ their hairstylists offer bonuses ranging from 35% to 60% of sales.\(^8\)

6 General case with spillovers

Having established the baseline results for the case without spillovers in the previous section, we now turn to the analysis of the general case, in which spillovers are present.

Proposition 1 implies that the firm’s profits in \( E \)-mode can be written as

\[
\Pi^E = \max_{t,a,q,Q} \{ N (R(a,s(a),q,Q) - f(a) - c(q)) - C(Q) \}
\]

\[\text{s.t.}
\begin{align*}
& t(R_a(a,s(a),q,Q) + s_a(a)R_\sigma(a,s(a),q,Q)) = f_a(a) \\
& (1-t)R_q(a,s(a),q,Q) = c_q(q) \\
& tNR_Q(a,s(a),q,Q) = C_Q(Q).
\end{align*}
\]

Similarly, the firm’s profits in \( A \)-mode can be written as

\[
\Pi^A = \max_{t,a,q,Q} \{ N (R(a,s(a),q,Q) - f(a) - c(q)) - C(Q) \}
\]

\[\text{s.t.}
\begin{align*}
& (1-t)R_a(a,s(a),q,Q) = f_a(a) \\
& (1-t)R_q(a,s(a),q,Q) = c_q(q) \\
& tNR_Q(a,s(a),q,Q) = C_Q(Q).
\end{align*}
\]

Comparing the two programs above, there are now two differences between the two modes, both originating in the choice of the non-transferable actions. The first difference is the same as in the case without spillovers: the first-order condition in \( a \) has a factor \( t \) in \( E \)-mode and a factor \( (1-t) \) in \( A \)-mode. The second difference is new and stems from the presence of spillovers across the \( N \) agents: in \( E \)-mode the firm internalizes the spillover when setting \( a_i \) for \( i = 1,..,N \), whereas the spillovers are left uninternalized in \( A \)-mode when each \( a_i \) is chosen by the individual agent \( i \).

We can now derive the corresponding proposition to Proposition 2.

**Proposition 4** Compare the firm’s profits under the two modes.

(a) If the transferable action \( a \) is contractible, then the two modes are equivalent and lead to the same firm profits \( (\Pi^E = \Pi^A) \). If the transferable action is costless and non-contractible (i.e. \( f = 0 \)), then the two modes lead to different profits when there are spillovers \( (R_\sigma \neq 0) \). If in addition the revenue function is additively separable in \( (a,\sigma) \) and \( (q,Q) \) (i.e. if it can be written

\[\text{See “Hair & Nail Salons in the US,” IBIS World Industry Report 81211, February 2015.}\]
\[ R1(a, \sigma) + R2(q, Q), \text{ then } \Pi^{E*} > \Pi^{A*}. \]

(b) Suppose the transferable action \( a \) is non-contractible. If the non-transferable action \( q \) is contractible or if it has no impact on revenue \( (R_q = 0) \), then \( \Pi^{E*} > \Pi^{A*} \). If \( f = 0 \) and the non-transferable action \( Q \) is contractible or has no impact on revenue \( (R_Q = 0) \), then \( \Pi^{E*} > \Pi^{A*} \).

There are two key differences in Proposition 4 relative to Proposition 2. First, due to spillovers, \( f = 0 \) no longer leads to equivalence. This reflects that in \( E \)-mode, spillovers across the choices of \( a_i \)'s are internalized, whereas in \( A \)-mode they are not. One may think that this always leads to the \( E \)-mode to dominate the \( A \)-mode; however, this is only true when the revenue function is additively separable in \( (a_i, \sigma_i) \) and \( (q_i, Q) \), \( Q \) is contractible, or \( Q \) has no impact on revenue. If instead all three types of actions are non-contractible and impact revenues, and there are interaction effects between the transferable action and the two types of non-transferable investments, then either mode may dominate. In particular, interaction effects between \( a_i \) and \( q_i \) or between \( a_i \) and \( Q \) may either exacerbate or dampen the disadvantage of the \( A \)-mode in terms of not internalizing spillovers.

The second difference is that in case (b), contractibility of \( Q \) or \( R_Q = 0 \) no longer necessarily implies that the \( A \)-mode dominates. In this case, the advantage of the \( A \)-mode in achieving the constrained first-best level of the \( q_i \)'s \((\text{by setting } t^{A*} = 0)\) must still be traded-off against the advantage of the \( E \)-mode in internalizing spillovers across the \( a_i \)'s. If, in addition, the transferable action is costless, then the \( E \)-mode can also achieve the constrained first-best level of \( q_i \) \((\text{the choice of the transferable action does not depend on } t^E, \text{ so } t^{E*} = 0)\), which implies that the \( E \)-mode is strictly better.

Finally, note that all the results in Proposition 4 would continue to hold even if we allowed for spillovers of effort \( q_i \) across revenues attributable to other agents \( j \neq i \) \((\text{accompanied by the appropriate changes in assumptions (a3)-(a5)})\). Indeed, the respective first-order conditions corresponding to \( q \) in programs (5) and (7) would stay the same: the spillover from agents’ non-transferable effort remains uninternalized in both \( E \)-mode and \( A \)-mode because in both modes agents choose \( q_i \)'s individually. Thus, the tradeoff between the two modes would not be impacted materially by spillovers generated by the non-contractible, non-transferable efforts \( q_i \). This is why we have abstracted away from such spillovers.

In the presence of spillovers, we can no longer derive results similar to the ones in Proposition 3, even when the revenue function \( R(a, \sigma, q, Q) \) is supermodular in all of its arguments. The only exception is the case when spillovers are positive, i.e. \( R_\sigma(a, \sigma, q, Q) > 0 \) for all \((a, \sigma, q, Q)\), and \( t^{A*} \geq 1/2 \): in this case we can conclude that \( \Pi^{A*} \leq \Pi^{E*} \). The logic is very similar to the one in Proposition 3: if \( t^{A*} \geq 1/2 \), then the distortions in the choices of non-contractible actions can be reduced by shifting control over the transferable actions from the agents to the firm. Indeed, this changes the first-order condition determining the equilibrium \( a \) in the second stage from

\[ (1 - t^{A*}) R_a(a, s(a), q, Q) = f_a(a) \]

to

\[ t^{A*} (R_a(a, s(a), q, Q) + s_a(a) R_\sigma(a, s(a), q, Q)) = f_a(a). \]
The other two first-order conditions stay unchanged. Since \( t^{A^*} \geq 1 - t^{A^*} \), \( s_a R_\sigma > 0 \) and \( (a, s(a), q, Q) \) are strategic complements in the revenue function, this change results in second-stage equilibrium levels of \((a, q, Q)\) that are closer to the first-best levels, and therefore equilibrium profits that are higher.

However, this argument breaks down if spillovers are not everywhere positive, because then the second-stage equilibrium levels of \((a, q, Q)\) may be lower. Furthermore, regardless of the sign of spillovers, \( t^{E^*} \leq 1/2 \) does not necessarily imply \( \Pi^{E^*} \leq \Pi^{A^*} \): shifting control over the transferable actions from the firm to the agents reduces the revenue-sharing distortion \( (1 - t^{E^*} \geq t^{E^*}) \), but creates the distortion due to spillovers being left uninternalized by the agents, so it is unclear which of these two effects dominates.\(^9\)

The interaction between revenue sharing and spillovers lies at the heart of the tradeoff that we study. 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 \( A \)-mode, which exacerbates the revenue-sharing distortion. Thus, as the magnitude of positive spillovers increases, the \( E \)-mode becomes relatively more attractive. However, this does not necessarily mean the \( E \)-mode dominates whenever spillovers are positive: the tradeoff also hinges on the relative magnitudes of double-sided moral hazard. On the other hand, negative spillovers lead the \( a_i \)'s to be set too high in \( A \)-mode, which can offset the distortion due to revenue-sharing. As we discuss in greater detail in the next two sections, this possibility has important and 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 \( A \)-mode.

Based on Proposition 4, the two simplest scenarios in which the tradeoff between the two modes is meaningful 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 and illustrate the interactions between the distortions due to revenue-sharing and spillovers—we analyze them in the next sections through specific examples. These two cases correspond to realistic scenarios. In many contexts prices are easily observable and contracted on, which means that they do not have an impact on the \( E \)-mode versus \( A \)-mode distinction. Alternatively, in other cases parties cannot observe price or quantity separately, so can only contract on revenue. Price then becomes a relevant transferable and non-contractible variable.

\(^9\)Nevertheless, in the online appendix we show numerically using a linear example, that knowing whether agents receive more or less than 50% of variable revenues still allows us to correctly predict the choice of mode most of the time. This suggests that Corollary 1 still has some predictive power, even in the presence of spillovers.
7 Linear example

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

$$R(a_i, \pi_i, q_i, Q) = \beta a_i + x (\pi_i - a_i) + \phi q_i + \Phi Q,$$

where $\pi_i = \sigma(a_i)$ is the average of the transferable actions chosen for $j \neq i$, i.e.

$$\pi_i = \frac{\sum_{j \neq i} a_j}{N - 1}.$$

Costs are assumed to be quadratic:

$$f(a) = \frac{1}{2} a^2, \quad c(q) = \frac{1}{2} q^2 \quad \text{and} \quad C(Q) = \frac{1}{2} Q^2.$$

Thus, when spillovers are negative ($x < 0$), revenue $R$ is decreasing in $\pi_i$, which means that in $A$-mode the transferable actions $a_i$ are set too high. Conversely, when spillovers are positive ($x > 0$), revenue $R$ is increasing in $\pi_i$, so that in $A$-mode the $a_i$'s are set too low. For example, if $a_i$ represents advertising, then negative (respectively, positive) spillovers occur when other agents' advertising decreases (respectively, increases) the demand realized by agent $i$.

We assume

$$x < \beta \quad \text{and} \quad x (\beta - x) < N \Phi^2,$$

which ensures that (i) assumptions (a3)-(a5) are satisfied for this example, and (ii) the optimal variable fee will be strictly between 0 and 1 in both modes. Note that all $x < 0$ are permissible under (9).

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

$$t_{E^*} = \frac{\beta^2 + N \Phi^2}{\beta^2 + \phi^2 + N \Phi^2}$$

$$t_{A^*} = \frac{N \Phi^2 - x (\beta - x)}{(\beta - x)^2 + \phi^2 + N \Phi^2}$$

and the following proposition.$^{10}$

**Proposition 5** The firm prefers the $A$-mode to the $E$-mode if and only if

$$\left| x \frac{\phi^2}{\beta} + \beta^2 + N \Phi^2 \right| \leq \sqrt{\beta^2 (\beta^2 + \phi^2 + N \Phi^2) + \phi^4}. \quad (11)$$

Consider first the baseline case with no spillovers, i.e. $x = 0$. Then the firm prefers the $A$-mode to the $E$-mode if and only if

$$\phi^2 > N \Phi^2. \quad (12)$$

$^{10}$It is straightforward to verify that the entire range of $x$ defined by (11) is permissible by assumptions (9) for $\beta$ sufficiently large.
In other words, the firm prefers the $A$-mode if the agents’ moral hazard is more important than the firm’s moral hazard, consistent with the intuition developed in Section 5. In particular, the tradeoff does not depend on $\beta$, 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 $E$-mode and $(1 - t^A_*)$ in $A$-mode) is increasing in $\beta$. Since $t^E_*$ and $(1 - t^A_*)$ increase at the same rate in this particular example (due to the symmetry of $E$-mode and $A$-mode profits in $N\Phi^2$ and $\phi^2$), the resulting tradeoff does not depend on $\beta$.

Consider now the tradeoff for general $x$. If $\beta^2 + N\Phi^2 < \sqrt{\beta^2(\beta^2 + \phi^2 + N\Phi^2)} + \phi^4$ (which is equivalent to $\phi^2 > N\Phi^2$), so that moral hazard considerations favor the $A$-mode, then the $A$-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 $E$-mode dominate. On the other hand, if $\phi^2 < N\Phi^2$, so that moral hazard considerations favor the $E$-mode, then the $A$-mode is still preferred for an intermediate, bounded range of negative spillovers. To understand why, note that in $A$-mode, negative spillovers cause the agents to set their $a_i$’s too high relative to what the firm would like them to choose all else equal. But this implies that in $A$-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 $A$-mode than in $E$-mode, so the $A$-mode can dominate (this advantage of $A$-mode must still be traded-off against the moral hazard advantage of the $E$-mode when $\phi^2 < N\Phi^2$). When the offsetting effect becomes too strong, the resulting levels of $a_i$’s in $A$-mode are too far above the first-best levels, so the $E$-mode dominates again.

Inspection of (11) reveals that the range of spillover values $x$ for which the firm prefers the $A$-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 $A$-mode, which exacerbates the primary revenue distortion. This makes the $A$-mode relatively less likely to dominate. There still exists a range of positive spillovers for which the $A$-mode is preferred provided the agents’ moral hazard is more important than that of the firm, but that range is smaller than the corresponding range of negative spillovers.

The skew towards a negative value of $x$ in condition (11) 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 $A$-mode.\footnote{Specifically, if $-\beta^2 - N\Phi^2 < x^2 < 0$, then condition (11) is more likely to hold when $x$ decreases.} This result runs counter to the common intuition, according to which spillovers should always make centralized control (i.e. $E$-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 $A$-mode levels of $a_i$’s get closer to the first-best level through the offsetting effect described above, so the $A$-mode becomes relatively more attractive (the $E$-mode levels of $a_i$’s are unchanged). If spillovers are positive or very negative, then an increase in their magnitude moves the $A$-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 discussed 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 (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 A-mode relatively more desirable, by allowing the firm to pay lower commissions while keeping consultants’ incentives constant.

We now investigate the impact of $\phi^2$ and $N\Phi^2$ on the tradeoff between A-mode and E-mode, i.e. on the profit differential $\Pi^{A*} - \Pi^{E*}$. From (11), 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 $\phi$ shifts the tradeoff in favor of A-mode (i.e. $\frac{d(\Pi^{A*} - \Pi^{E*})}{d(\phi^2)} > 0$) if and only if $t^{A*} < t^{E*}$. A larger $\Phi$ shifts the tradeoff in favor of E-mode (i.e. $\frac{d(\Pi^{A*} - \Pi^{E*})}{d(N\Phi^2)} < 0$) if and only if $t^{A*} < t^{E*}$.

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 firm is larger in E-mode, i.e. $t^{E*} > t^{A*}$. From expressions (10), this is always the case in the absence of spillovers ($x = 0$). However, with spillovers, $t^{A*} > t^{E*}$ if and only if

$$\frac{x}{\beta} + \frac{\beta}{\beta - x} < -\frac{\beta^2 + N\Phi^2}{\phi^2},$$

i.e. if the spillover $x$ is sufficiently negative.\(^{12}\) Thus, when the inequality in (13) holds, an increase in the importance of the agents’ (respectively, the firm’s) moral hazard shifts the trade-off in favor of the E-mode (respectively, A-mode), the opposite of what one might expect and of the baseline tradeoff given by (12) for the case without spillovers.

This second counterintuitive result represents a novel contribution of our paper relative to traditional theory of the firm models, which predict that ownership and/or control over assets should be given to the party whose investment incentives are more important (see Grossman and Hart, 1986). The result can be interpreted as follows. Negative spillovers partially offset the revenue-sharing distortion in A-mode. As a result, a higher $t$ induces less distortion of the transferable actions $a_i$ in A-mode, so the firm can charge a higher $t$ in A-mode, to the point that $t^{A*} > t^{E*}$ if spillovers are sufficiently negative. However, when this occurs, agents retain a lower share of revenues in A-mode than in E-mode, so their choice of non-transferable effort $q_i$ is more distorted in A-mode. Consequently,

\(^{12}\)Recall that all $x < 0$ are permissible by assumptions (9). Furthermore, it is easily verified that the respective ranges in $x$ defined by (11) and (13) have a non-empty intersection.
when agents’ effort (moral hazard) becomes more important in this parameter region, the $E$-mode becomes relatively more attractive. Similarly, when the firm’s investment (moral hazard) becomes more important in the same parameter region, the $A$-mode becomes relatively more attractive.

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 firm 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).

8 Price as the transferable action

We now turn to the other case of interest identified in Section 6—the transferable action is the price that is either set by agent $i$ or the firm, and therefore does not carry any fixed costs. Specifically, the revenue generated by agent $i$ is now

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

where $d > 0$ is the demand intercept and $\bar{p}_{-i}$ 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 7.

To ensure assumptions (a1)-(a5) are satisfied for this example, we assume

$$\beta < 0, \phi > 0, \Phi > 0$$

$$-2\beta + \min \{0, 2x\} > \max \{ N\Phi^2, \phi^2 \}.$$ (15)

Note that (15) implies all $x > 0$ are permissible and $x > \beta$, so demand $d + \beta p_i + x (\bar{p}_{-i} - p_i) + \phi q_i + \Phi Q$ is decreasing in $p_i$.

From (14), 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 note that one could reinterpret $p_i$ as quantity instead of price, but then the usual case would be captured by negative spillovers ($x < 0$).

Define

$$k \equiv \frac{N\Phi^2 \phi^2}{N\Phi^2 + \phi^2} \in (0,|\beta|),$$

which can be viewed as a measure of the combined importance of both types of moral hazard ($k$ is symmetric in $N\Phi^2$ and $\phi^2$ and increasing in both).

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

\textsuperscript{13}Recall $0 < k < -\beta$ so $-\frac{4k(k+\beta)}{k+2\beta} < 0.$
Proposition 7 The firm prefers the $A$-mode if and only if

$$-\frac{4k(k + \beta)}{k + 2\beta} < x < 0.$$  \hspace{1cm} (16)

First, note that the proposition identifies a meaningful tradeoff since any $x$ satisfying (16) also satisfies (15) provided $\beta$ is sufficiently negative. This is true for all positive $x$.

Second, the $E$-mode is always preferred if spillovers are positive or if spillovers are 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 firm and each agent. As a result, the variable fee $t$ can be used in both modes to balance the double-sided moral hazard problem ($q_i$ versus $Q$) equally well.\textsuperscript{14} Furthermore, due to the strategic complementarity between $p_i$ and $(q_i, Q)$, the choice of $p_i$ can either offset or compound the double-sided moral hazard problem. However, the $E$-mode has an advantage in internalizing spillovers across the agents’ services. This explains why there is a larger region over which the $E$-mode dominates.

The fact that agents do not internalize spillovers in $A$-mode can work in favor of the $A$-mode when spillovers are negative ($x < 0$). Namely, when $x < 0$, the $A$-mode leads to excessively high choices of $p_i$, which can help offset the double-sided moral hazard problem. This is because a higher $p_i$ leads to higher $q_i$ and $Q$ due to strategic complementarity, which offsets the problem of $q_i$ and $Q$ being too low that arises from revenue-sharing. If this offsetting effect is moderately strong, then the resulting levels of $q_i$’s and $Q$ are closer to first-best in $A$-mode than in $E$-mode, so the $A$-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 $A$-mode are too far above the first-best levels, so the $E$-mode is preferred. In contrast, when $x > 0$, the $A$-mode leads to $p_i$ being set too low, which compounds the double-sided moral hazard problem. As a result, the $E$-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 $A$-mode in the context of one of the examples from Section 3, namely franchising. In general, due to customer mobility and the over-arching brand of the franchisor, its various franchisees create positive demand externalities on one another, i.e. a lower price charged by one franchisee likely increases demand for the other franchisees as well. 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 own the franchised outlets or at least 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.

\textsuperscript{14}For this reason, there is no underlying tendency to have a high $t$ under $E$-mode and a low $t$ under $A$-mode. Thus, unlike the case in which the transferable action was costly, knowing whether agents receive more or less than 50% of variable revenues does not, in general, help predict the choice of mode when the transferable action is costless.
Third, the parameters measuring the strength of the two moral hazard problems, \( N\Phi^2 \) and \( \phi^2 \), 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 the two moral hazards 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. As noted above, when spillovers are negative, raising prices reduces the moral hazard problems due to the strategic complementarity between prices and investments, and this works equally well for both \( q_i \) and \( Q \). Thus, the extent to which the \( A \)-mode is preferred over the \( E \)-mode when moral hazard problems become more important does not depend on the source of the moral hazard, but only on its aggregate magnitude, measured by \( k \).

Finally, it is easily verified that \( -\frac{4k(k+|\beta|)}{k+2|\beta|} \) is decreasing for \( k \in (0, (2 - \sqrt{2}) |\beta|) \) and increasing for \( k \in [(2 - \sqrt{2}) |\beta|, |\beta|) \). Thus, assuming negative spillovers, when \( \phi^2 \) and \( N\Phi^2 \) are small (i.e. \( k \) is small), the range of \( x \) over which the \( A \)-mode is preferred increases as \( \phi^2 \) and \( N\Phi^2 \) increase; and vice versa when \( \phi^2 \) and \( N\Phi^2 \) are large. In other words, when the double-sided moral hazard problem is of small importance, the effectiveness of the \( A \)-mode in compensating for moral hazard with excessive prices increases as moral hazard becomes more important, so the tradeoff shifts in favor of the \( A \)-mode. And vice versa when double-sided moral hazard is already very important.

9 Extensions

This section explores several extensions to our model: changing the timing of the infrastructure investment \( Q \); allowing for cost asymmetries between the firm and the agent(s); allowing the firm to choose a hybrid mode that lies between the pure \( E \)-mode and pure \( A \)-mode; and allowing for private benefits.

9.1 Timing

When \( Q \) represents a basic infrastructure investment that is fundamental to the firm’s operations, in some cases this investment is made before the choice of business model (\( E \)-mode versus \( A \)-mode) rather than afterwards. This reflects that it may be easier for a firm to change its business model than its basic infrastructure. In this case, the model becomes very similar to that in Hagiu and Wright (2015b), but without private information.

The net result of this change in timing is to shift the firm’s business model tradeoff in favor of the \( A \)-mode. Indeed, if the firm is able to commit to its choice of \( Q \) prior to the choice of business model, then the need to keep a larger share of variable revenues to motivate investments in \( Q \) disappears. This observation implies that one of the factors that determine the choice of mode is the extent to which the firm’s investments are determined upfront. Thus, when the firm’s \textit{ongoing} investments in infrastructure (or other forms of common investment) increase in importance relative to its \textit{ex-ante} investments, the tradeoff shifts in the same way as predicted by our analysis above when the firm’s moral hazard problem becomes more important.
9.2 Cost asymmetries

Throughout the analysis above we have assumed there are no asymmetries between the firm and the agents in the costs of undertaking the transferable action or in its impact on revenues. In some real-world examples, such asymmetries are an important factor in determining which control rights are held by the firm and which are held by agents. For example, the firm may have economies-of-scale advantages over individual agents when incurring the cost associated with the transferable action \( a \) (e.g. economies of scale in purchasing equipment) or better information regarding the impact of the transferable action on revenues due to access to more data (e.g. Uber and Lyft when setting prices for rides).

Introducing cost asymmetries between the firm and the agents is straightforward in our model without spillovers. We simply assume that the cost of the transferable action \( a \) is \( F(a) \) when incurred by the firm in \( E \)-mode and \( f(a) \) when incurred by the agents in \( A \)-mode, where the functions \( F \) and \( f \) have the same properties as previously assumed for the function \( f \) used throughout the paper. Proposition 1 continues to hold for both modes, such that the optimal contracts are linear. In the online appendix, we confirm that, not surprisingly, a higher relative cost advantage for the firm (respectively, the agents) shifts the trade-off in favor of the \( E \)-mode (respectively, \( A \)-mode). Note that such cost asymmetries are not equivalent to the asymmetric private benefits studied in Section 9.5. With cost asymmetries, the total payoff maximized by the firm changes from one mode to the other, which was not the case with private benefits.

9.3 Hybrid mode across agents

Hybrid modes, with some agents offering their services in \( E \)-mode and others in \( A \)-mode, are found quite often in the markets where our theory is relevant (e.g. consultancies, hair salons, and sales representatives for industrial companies). In this subsection we show that a strictly hybrid mode can be optimal even without spillovers (i.e. we assume \( R_s = 0 \)) and despite the fact that all \( N \) agents are identical. This is because \( Q \) is a common investment across all the agents’ services (e.g. a common infrastructure) and because of the concavity of the profit function with respect to \( Q \).

We use the linear example from Section 7 with no spillovers (\( x = 0 \)) and quadratic costs. At first glance, this seems like the least likely scenario for a hybrid mode to be optimal, because there are no interaction effects and no asymmetries between firm and agents. In the online appendix, we establish the following result.

**Proposition 8** The optimal number of employees (as opposed to independent contractors) is

\[
n^* = \begin{cases} 
N & \text{if } \Phi^2 > \beta^2 + \phi^2 \\
N \left(1 - \frac{\phi^2 (\beta^2 + \phi^2 - N\Phi^2)}{2N\Phi^2\phi^2}\right) & \text{if } \beta^2 + \phi^2 > N\Phi^2 > \phi^2 - \frac{\beta^2\phi^2}{2\beta^2 + \phi^2} \\
0 & \text{if } N\Phi^2 < \phi^2 - \frac{\beta^2\phi^2}{2\beta^2 + \phi^2}.
\end{cases}
\]

Note that \( n^* \) is increasing in \( N\Phi^2 \) (the importance of the firm’s moral hazard) and decreasing in
\( \phi^2 \) (the importance of agents' moral hazard), consistent with the intuition built in Section 7 for the case without spillovers.

The reason why the optimal choice of \( n \) can be interior (i.e. strictly between 0 and \( N \)) is that the firm can only choose a single \( Q \) (e.g. infrastructure investment), which affects all agents. If the firm could choose a different \( Q \) for agents in each mode, then it would choose a higher \( Q \) for agents in \( E \)-mode than for agents in \( A \)-mode. The optimal level of \( Q \) in each mode (and the firm’s corresponding profit per agent) would be independent of the number of agents in each mode. The optimal solution would then be \( n = N \) or \( n = 0 \), depending on which mode yields higher profit per agent.

In contrast, when the firm must choose the same \( Q \) for agents in both modes, the firm’s choice of \( Q \) depends on how many agents are in each mode, since it depends on the weighted average of the variable fee in each mode, i.e. \( t(n) \equiv \frac{N}{N} t^E + \frac{N-n}{N} t^A \). Since the firm’s profit function is concave in \( Q \), it becomes concave in \( n \) (instead of linear): substituting an agent in \( A \)-mode for an agent in \( E \)-mode has a lower impact on profits when the number of agents in \( E \)-mode is larger. This explains why a mix of employees and independent contractors sometimes allows the firm to do better.

### 9.4 Hybrid modes across actions

Until now, we have always restricted attention to a single transferable action for concision. In many real-world examples, however, there are multiple relevant transferable actions (see Table 1 in Section 3). This provides another dimension along which firms can (and oftentimes do) operate in hybrid modes, with some transferable decisions controlled by agents and others by the firm.

Our model can be extended to encompass this dimension as well. For simplicity, consider the case with no spillovers, but multiple transferable and costly actions, \( a^j \), with \( j \in \{1, \ldots, M\} \), so the revenue function for agent \( i \) is \( R(a^1_i, \ldots, a^M_i, q_i, Q) \). The fixed cost associated with transferable action \( a^j_i \) is \( f^j(a^j_i) \). The costs associated with non-transferable actions are \( c(q_i) \) and \( C(Q) \) as before.

The firm optimizes its profits over the contract \((t, T)\) and the set \( D \subset \{1, \ldots, M\} \) of decisions over which it keeps control (agent \( i \) controls decisions \( j \in \{1, \ldots, M\} \setminus D \)). The following proposition establishes that the hybrid interior mode can never be optimal when \( R(a^1_i, \ldots, a^M_i, q_i, Q) \) is supermodular in its arguments.

**Proposition 9** If the revenue function \( R \) is supermodular in \((a^1_i, \ldots, a^M_i, q_i, Q)\), then the optimal mode is either \( D = \emptyset \) (\( A \)-mode) or \( D = \{1, \ldots, M\} \) (\( E \)-mode).

The key driving force behind the result in Proposition 9 is that reducing the distortions in the firm’s and the agents’ second stage objective functions relative to the firm’s first-stage objective function raises the firm’s profit. If \( t^* < 1/2 \), the distortions can be reduced by giving control over all transferable actions to agents. If \( t^* > 1/2 \), the distortions can be reduced by giving control over all transferable actions to the firm. Only in the special case when \( t^* = \frac{1}{2} \), would any split of control rights, including a strictly interior split, be optimal. However, the proposition still applies given that pure modes remain weakly optimal in this case. The proposition suggests that, in the absence of spillovers, a strictly
interior split of control rights can only be optimal if there are negative interaction effects across the actions \((a_1^1, ..., a_i^M, q_i, Q)\).

Spillovers provide a natural explanation for the adoption of a hybrid mode, especially if they are positive for some actions and negative for others, or are present for some actions and not others. Alternatively, allowing for asymmetries here (as in Section 9.2 above) would provide a natural way of explaining which control rights are held by the firm and which are held by agents. For instance, Uber and Lyft have a clear advantage over their drivers in setting prices for rides (better information), whereas drivers are in a better position to choose their work schedules and the amount of maintenance their individual cars need. Hair salons may take control over the choice of “uniforms” because they can obtain lower costs due to scale effects (relative to individual hairstylists), but the marketing of hairstylists may be more efficiently done by each individual.

### 9.5 Private benefits

Consider the benchmark model without spillovers from Section 5. The transferable action \(a\) can drive an additional wedge between the two modes when one or both parties derive private benefits from the choice of \(a\). Examples of private benefits include the enhancement of individual agents’ reputation and outside opportunities by the marketing of their services (e.g. hairdressers, consultants, sales representatives) and the improved reputation of the firm by the choice of better equipment (e.g. hair salons, hospitals and clinics, taxi companies).

Formally, suppose that each \(a_i\) influences some non-contractible outside payoffs, \(B(a_i)\) for the firm and \(b(a_i)\) for agent \(i\), where the functions \(B\) and \(b\) are non-negative, twice-continuously differentiable and increasing. It is easily seen that, after adjusting assumptions (a1)-(a5) accordingly, the proof of Proposition 1 continues to apply, so linear contracts remain optimal in both modes.

Private benefits change the programs (1) and (3) that determine \(E\)-mode and \(A\)-mode profits in two ways. First, since the agents’ private benefits affect their willingness to participate, the function maximized by the firm in both modes is now

\[
N(R(a, q, Q) + B(a) + b(a) - f(a) - c(q)) - C(Q).
\]

Second, the first-order condition in \(a\) for the \(E\)-mode is now\(^{15}\)

\[
tR_a(a, q, Q) + B_a(a) = f_a(a),
\]  

(17)

while the new first-order condition in \(a\) for the \(A\)-mode is

\[
(1 - t)R_a(a, q, Q) + b_a(a) = f_a(a).
\]

(18)

Comparing the new first-order conditions (17) and (18) with the ones in the programs (2) and (4),

\(^{15}\)The reason we focus on private benefits influenced by \(a\) only is that any private benefit influenced by \(q\) or \(Q\) would not create any difference between the sets of first-order conditions (2) and (4) corresponding to the two modes.
it is clear that Proposition 2 no longer holds. In particular:

- Even if \( a \) is costless, as long as the private benefit functions \( b \) and \( B \) are different, the resulting profits in \( E \)-mode and \( A \)-mode are different.

- Even if \( q \) (respectively, \( Q \)) is contractible or does not impact revenues, the \( A \)-mode (respectively, \( E \)-mode) might still dominate if \( b \) (respectively, \( B \)) is sufficiently large.

Heuristically, holding everything else constant, if the firm’s private benefit \( B \) is more (respectively, less) important than the agents’ private benefit \( b \), then the \( E \)-mode is more (respectively, less) likely to be preferred to the \( A \)-mode.

As in Section 5, we can reach more precise results for the case when \( a \) is costly by imposing supermodularity on \( R \). In the online appendix we extend Proposition 3 and Corollary 1 to the case with linear private benefits, i.e. \( B(a) = Ba \) and \( b(a) = ba \). This provides a way to continue to predict which mode a firm is operating in based on the observed share of variable revenues retained by the agent. We also obtain a precise comparison of the two modes by resorting to the linear example of Section 7, to which we add linear private benefits \( B(a) = Ba \) and \( b(a) = ba \).\(^{16}\) Extending the result in (12) to this case, we obtain that the firm prefers the \( A \)-mode to the \( E \)-mode if and only if

\[
\left( (\beta + b)^2 - B^2 \right) \phi^2 > \left( (\beta + B)^2 - b^2 \right) N \Phi^2.
\]

Thus, the tradeoff in (12) is robust to the introduction of private benefits: the tradeoff shifts in favor of the \( A \)-mode when the agents’ moral hazard and private benefit become more important and in favor of the \( E \)-mode when the firm’s moral hazard and private benefit become more important. In particular, note that \( b = B \) implies \( \Pi^A > \Pi^E \) if and only if \( \phi^2 > N \Phi^2 \); i.e. (12) is restored when the private benefits of the two parties are equally important. Furthermore, \( \phi^2 = N \Phi^2 \) implies \( \Pi^A > \Pi^E \) if and only if \( b > B \). In other words, if the agents’ and the firm’s moral hazard are equally important, then the choice of mode only depends on which party’s private benefits are more important.

10 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 between employment mode and agency mode, 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 balancing double-sided moral hazard, while at the same time minimizing distortions in the choice of transferable actions due

\(^{16}\)If instead of additive private benefits, the private benefits are proportional to the demand underlying the linear revenue function, then it turns out they are irrelevant for the tradeoff between the two modes. See online appendix.
to revenue sharing and spillovers. Our modeling approach and some of our key results are reminiscent of the theory of the firm based on property rights and incentive systems. In particular, in the baseline model without spillovers, a key prediction is that control rights over transferable actions should be given to whichever party’s (the firm or the agents) moral hazard problem is more important. However, we have shown that this prediction can be overturned when spillovers are introduced, both with costly and costless transferable actions. The counterintuitive scenario occurs when the spillover distortion offsets the revenue-sharing distortion. Moreover, when the transferable action is price (and therefore there are interaction effects with the non-transferable actions), the tradeoff between the two modes no longer depends on the source of moral hazard, but only on its magnitude.

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.\(^\text{17}\) 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 in the employment mode and by independent contractors in the agency mode (e.g. health insurance and worker tax filings) makes no difference to the tradeoff between the two modes. Although such a cost is not incurred directly by the firm in the agency mode, it must be internalized through a lower fixed fee charged to agents because the agents’ outside option remains unchanged. However, in practice, some of the fixed costs incurred by the firm in the employment mode 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 agency mode, 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. 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 in both modes, and so should shift the tradeoff in favor of the employment 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. One could further pursue the analysis of hybrid modes across actions to take into account cost or information asymmetries between the firm and the agents, or negative interaction effects across the transferable

\(^{17}\)See for example Justin Fox “Uber and the Not-Quite-Independent Contractor” Bloomberg, June 30, 2015.
actions. Finally, it would be interesting to study competition among firms that can each choose between the employment mode and the agency mode, potentially leading to equilibria in which firms compete with different models.

References


11 Appendix

Proof of Proposition 1

Consider first the $E$-mode. Denote by

$$\bar{a}_n \equiv (a, ..., a)$$

the vector of $n \in \{1, .., N\}$ coordinates all equal to $a$ (thus, $\sigma(\bar{a}_{N-1}) = s(a)$). Since we have assumed that it is optimal for the firm to induce all $N$ agents to join and we are focusing on symmetric solutions, the optimal contract $\Omega^*(a)$ (i.e. payment to each agent) solves

$$\Pi^{E*} = \max_{\Omega(a,q),a,q} \{ N[R(a, \sigma(\bar{a}_{N-1}), q, Q) - \Omega(R(a, \sigma(\bar{a}_{N-1}), q, Q)) - f(a)] - C(Q) \}$$

s.t.

$$a = \arg \max_{a'} \left\{ \frac{R(a', \sigma(\bar{a}_{N-1}), q, Q) - \Omega(R(a', \sigma(\bar{a}_{N-1}), q, Q))}{(N-1)} + (N-1) (R(a, \sigma(a', \bar{a}_{N-2}), q, Q) - \Omega(R(a, \sigma(a', \bar{a}_{N-2}), q, Q))) - f(a') \right\}$$

$$q = \arg \max_{q'} \{ \Omega(R(a, \sigma(\bar{a}_{N-1}), q', Q)) - c(q') \}$$

$$Q = \arg \max_{Q'} \{ N[R(a, \sigma(\bar{a}_{N-1}), q, Q')] - \Omega(R(a, \sigma(\bar{a}_{N-1}), q, Q')) \} - C(Q') \}$$

$$0 \leq \Omega(R(a, \sigma(\bar{a}_{N-1}), q, Q)) - c(q).$$

Let then $(a^*, q^*, Q^*)$ denote the symmetric outcome of this optimization problem. Also define

$$R^* \equiv R(a^*, \sigma(\bar{a}_{N-1}^*), q^*, Q^*) = R(a^*, s(a^*), q^*, Q^*).$$

In the online appendix we prove the following lemma.
Lemma 1 \( \Omega^* (.) \) is continuous and differentiable at \( R^* \).

The lemma and program (19) imply that \((a^*, q^*, Q^*)\) solve

\[
\begin{align*}
(1 - \Omega^*_R (R^*)) (R_a (a^*, s (a^*), q^*, Q^*) + s_a (a^*) R_{s} (a^*, s (a^*), q^*, Q^*)) &= f_a (a^*) \\
\Omega^*_R (R^*) R_q (a^*, s (a^*), q^*, Q^*) &= c_q (q^*) \\
N (1 - \Omega^*_R (R^*)) R_Q (a^*, s (a^*), q^*, Q^*) &= C_Q (Q^*) \, .
\end{align*}
\]

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

A similar proof applies to the case when the firm chooses the \( \mathcal{A} \)-mode.

Proof of Proposition 2

For (a), if \( a \) is contractible, then the constraint in \( a \) disappears in both modes, so the programs (1) and (3) become identical. If \( f = 0 \), then the constraint in \( a \) is the same in both modes and is defined by \( R_a (a, q, Q) = 0 \), so the two modes are equivalent once again.

For (b), if the agents’ effort has no impact on revenues \( (R_q = 0) \) then each agent sets \( q_i = 0 \) in both modes. In \( \mathcal{E} \)-mode it is then optimal for the firm to retain the entire revenue \( (t = 1) \), so profits are

\[
\Pi^{\mathcal{E}} = \max_{a, Q} \{ N (R (a, 0, Q) - f(a)) - C (Q) \} \, .
\]

This is clearly higher than profits under \( \mathcal{A} \)-mode:

\[
\Pi^{\mathcal{A}} = \max_{t, a, Q} \{ N (R (a, 0, Q) - f(a)) - C (Q) \}
\]

s.t.

\[
\begin{align*}
(1 - t) R_a (a, 0, Q) &= f_a (a) \\
tNR_Q (a, 0, Q) &= C_Q (Q) \, .
\end{align*}
\]

If the agents’ effort is contractible, then in \( \mathcal{E} \)-mode the firm optimally sets \( t = 1 \) and profits are

\[
\Pi^{\mathcal{E}} = \max_{a, q, Q} \{ N (R (a, q, Q) - f(a) - c(q)) - C (Q) \} \, .
\]

This is the first-best level of profits, which strictly dominate the profits that can be achieved in \( \mathcal{A} \)-mode.

By a symmetric argument, we obtain the result for the case when the firm’s investment has no impact on revenues \( (R_Q = 0) \) or \( Q \) is contractible.
Proof of Propositions 3 and 9

Since both Proposition 3 and Proposition 9 rely on supermodularity of the revenue function and their proofs are similar, we prove them together. To do so, we start directly with the set-up with multiple actions \( a^j \), where \( j \in \{1, \ldots, M\} \), so that the revenue function is \( R(a^1, \ldots, a^M, q, Q) \). The proof of Proposition 3 follows for the special case \( M = 1 \).

Suppose the fixed cost associated with transferable action \( a^j \) is \( f^j(a^j) \). To simplify notation, let 
\[
\begin{align*}
a^{M+1} & \equiv q \quad \text{and} \quad f^{M+1}(.) \equiv c(.) \\
a^{M+2} & \equiv Q \quad \text{and} \quad f^{M+2}(.) \equiv \frac{1}{N}C(.)
\end{align*}
\]

In other words, rename the agents’ non-transferable effort \( q \) as \( a^{M+1} \) and the firm’s non-transferable investment \( Q \) as \( a^{M+2} \). Supermodularity of the revenue function implies that \( R(a^j, a^k) \geq 0 \) for \( j \neq k \in \{1, \ldots, M\} \).

For any vector \( \vec{\tau} = (\tau_1, \ldots, \tau_{M+2}) \in [0,1]^{M+2} \), let 
\[
\Pi(\vec{\tau}) \equiv N\left(R(a^1, \ldots, a^{M+2}) - \sum_{j=1}^{M+2} f^j(a^j)\right),
\]

where \( (a^1, \ldots, a^{M+2}) \) is the unique solution to the \((M + 2)\) equations
\[
\begin{align*}
\tau_1 R_{a^1}(a^1, \ldots, a^{M+2}) &= f^1_{a^1}(a^1) \\
\vdots \\
\tau_{M+2} R_{a^{M+2}}(a^1, \ldots, a^{M+2}) &= f^{M+2}_{a^{M+2}}(a^{M+2}).
\end{align*}
\]

We first prove two preliminary lemmas, which will help in proving both Propositions 3 and 9.

**Lemma 2** For all \( k \in \{1, \ldots, M + 2\} \), the solution \( (a^1, \ldots, a^{M+2}) \) to (20) is strictly increasing in \( \tau_k \), i.e. \( a^j \) is (weakly) increasing in \( \tau_k \) for all \( j \in \{1, \ldots, M + 2\} \), and is strictly increasing in \( \tau_k \) for at least one \( j \).

**Proof.** Note that the solution \( (a^1, \ldots, a^{M+2}) \) corresponds to a game in which there are \( M + 2 \) players, and each player \( j \) sets \( a \) to maximize \( r^j \equiv \tau_j R(a^1, \ldots, a^{j-1}, a, a^{j+1}, \ldots, a^{M+2}) - f^j(a) \). Given \( R_{a^j a^k} \geq 0 \) for \( j \neq k \), we have \( r^j_{a^j a^k} \geq 0 \) for \( j \neq k \). Furthermore, \( r^j_{a^j a^j} > 0 \) and \( r^j_{a^j a^k} = 0 \) for \( k \neq j \). The game is therefore supermodular with payoffs having strictly increasing differences in the actions and the parameters \( (\tau_1, \ldots, \tau_{M+2}) \). From standard supermodularity results (Vives, 1999), we know that an increase in any of the parameters \( (\tau_1, \ldots, \tau_{M+2}) \) will increase each of the solutions \( a^j \) for \( j \in \{1, \ldots, M + 2\} \) in a weak sense. To obtain the strict comparative static result, note that the solution is defined by the set of equations defined in (20). This means that if \( \tau_k \) increases for some \( k \in \{1, \ldots, M + 2\} \) and \( a^j \) does not strictly increase for at least one \( j \in \{1, \ldots, M + 2\} \), then no \( a^j \) can change since none can
decrease. But if all \(a^j\)'s remain unchanged, then, since \(\tau_k\) is higher, the first-order conditions (20) can no longer hold. Thus, at least one \(a^j\) must strictly increase.

**Lemma 3** When \((\tau_1, .., \tau_{M+2}) \in [0, 1)^{M+2}\), the payoff function \(\Pi(\tau_1, .., \tau_{M+2})\) is strictly increasing in \(\tau_j\) for all \(j \in \{1, .., M + 2\}\).

**Proof.** For any \(j \in \{1, .., M + 2\}\), we have

\[
\frac{d\Pi}{d\tau_j} = N \sum_{k=1}^{M+2} \left( R_{a^k}(a^1, .., a^{M+2}) - f^k_{a^k}(a^k) \right) \frac{da^k}{d\tau_j}
\]

where we have used (20) to replace \(f^k_{a^k}(a^k)\). By assumption, \(R_{a^k} > 0\), and from Lemma 2, we know that \(\frac{da^k}{d\tau_j} \geq 0\) for all \(k \in \{1, .., M + 2\}\), with strict inequality for some \(k\). Thus, if \((\tau_1, .., \tau_{M+2}) \in [0, 1)^{M+2}\), then we can conclude that \(\frac{d\Pi}{d\tau_j} > 0\). ■

Suppose the firm chooses \(t\) and \(D \subset \{1, .., M\}\) as the subset of transferable decisions that it controls (the agents are therefore given control over decisions \(j \in \{1, .., M\} \setminus D\)). Then the profit obtained by the firm is equal to \(\Pi(\vec{\tau}(D, t))\), where \(\vec{\tau}(D, t)\) is the vector of \((M + 2)\) coordinates defined so that the \(j^{th}\) element is

\[
\tau(D, t)_j = \begin{cases} t & \text{if } j \in D \cup \{M + 2\} \\ 1 - t & \text{if } j \in \{1, .., M\} \setminus D \cup \{M + 1\} \end{cases}
\]

We can now use the lemmas to prove Propositions 3 and 9.

**Proof of Proposition 3**

Let \(M = 1\) and consider first the \(E\)-mode. If \(t^{E*} < 1/2\), then the firm could strictly increase profits by giving up control over the transferable actions to the agents and keeping the variable fee unchanged, equal to \(t^{E*}\). To see this, note that the change in profits is

\[
\Pi \left( \left( 1 - t^{E*}, 1 - t^{E*}, t^{E*} \right) \right) - \Pi \left( \left( t^{E*}, 1 - t^{E*}, t^{E*} \right) \right).
\]

If \(t^{E*} > 0\), then this difference is positive by Lemma 3, because \(0 < t^{E*} < 1/2\) implies \((t^{E*}, 1 - t^{E*}, t^{E*}) \in (0, 1)^3\) and \((1 - t^{E*}, 1 - t^{E*}, t^{E*}) > (t^{E*}, 1 - t^{E*}, t^{E*})\). If \(t^{E*} = 0\), then the change in profits is

\[
\Pi \left( \left( 1, 1, 0 \right) \right) - \Pi \left( \left( 0, 1, 0 \right) \right) = N \max_{a, q} \{ R(a, q, 0) - f(a) - c(q) \} - N \max_q \{ R(0, q, 0) - c(q) \},
\]

which is positive due to assumptions (a1)-(a4).
Thus, in all cases we have
\[ \Pi^E = \Pi \left( \left( t^E, 1 - t^E, t^E \right) \right) < \Pi \left( \left( 1 - t^E, 1 - t^E, t^E \right) \right) \leq \Pi^A. \]

If \( t^E = 1/2 \), then the same reasoning implies that \( \Pi^E \leq \Pi^A \).

Similarly, consider now the \( A \)-mode. If \( t^A > 1/2 \), then the firm could strictly increase profits by taking control over the transferable action \( a \) and keeping the variable fee unchanged, equal to \( t^A \). To see this, note that the change in profits is
\[ \Pi \left( t^A, 1 - t^A, t^A \right) - \Pi \left( 1 - t^A, 1 - t^A, t^A \right). \]

If \( t^A < 1 \), then this difference is positive by Lemma 3, because \( 1 - t^A > 1/2 \) implies \( (t^A, 1 - t^A, t^A) \in (0, 1)^3 \) and \( (t^A, 1 - t^A, t^A) > (1 - t^A, 1 - t^A, t^A) \). If \( t^A = 1 \), then the change in profits is
\[
\Pi \left( (1, 0, 1) \right) - \Pi \left( (0, 0, 1) \right) = \max_{a,Q} \left\{ N R(a,0,Q) - f(a) - C(Q) \right\} - \max_Q \left\{ N R(0,0,Q) - C(Q) \right\},
\]
which is positive because of assumptions (a1)-(a4).

Thus, in all cases we have
\[ \Pi^A = \Pi \left( 1 - t^A, 1 - t^A, t^A \right) < \Pi \left( t^A, 1 - t^A, t^A \right) \leq \Pi^E. \]

If \( t^A = 1/2 \), then the same reasoning implies that \( \Pi^A \leq \Pi^E \).

**Proof of Proposition 9**

Suppose there are \( M > 1 \) transferable actions. Denote by \( t^* \) the optimal variable fee and by \((D^*, \{1, \ldots, M\} \setminus D^*)\) the optimal allocation of control rights over the transferable actions. Suppose \( D^* \neq \emptyset \) and \( D^* \neq \{1, \ldots, M\} \). If \( t^* < 1 - t^* \) (i.e. \( t^* < 1/2 \)), then the firm could increase profits by giving up control over all actions \( a_j \) for \( j \in D \) to the agents and keeping \( t^* \) unchanged. To see this, note that the change in profits is
\[ \Pi \left( D^* \setminus \emptyset, t^* \right) - \Pi \left( D^*, t^* \right). \]

If \( t^* > 0 \), then this difference is positive by Lemma 3, because \( 1 - t^* > t^* > 0 \) and \( D^* \neq \emptyset \) imply
The function \( \mathcal{T} (D^*, t^*) \) can be written in the range \([0, 1]^{M+2}\) and \( \mathcal{T} (\emptyset, t^*) > \mathcal{T} (D^*, t^*) \). If \( t^* = 0 \), then the change in profits can be written

\[
\Pi (\mathcal{T} (\emptyset, 0)) - \Pi (\mathcal{T} (D^*, 0)) = N \max_{a^1, \ldots, a^{M+1}} \left\{ R (a^1, \ldots, a^{M+1}, 0) - \sum_{i=1}^{M+1} f^i (a^i) \right\} - N \max_{a^1, \ldots, a^{M+1}} \left\{ R (a^1, \ldots, a^{M+1}, 0) - \sum_{i \in \{1, \ldots, M+1\} \setminus D^*} f^i (a^i) \right\} \text{ s.t. } a^i = 0 \text{ if } i \in D^*.
\]

Clearly then, \( D^* \neq \emptyset \) and assumptions (a1)-(a4) (generalized to the current scenario with multiple transferable actions) imply \( \Pi (\mathcal{T} (\emptyset, 0)) - \Pi (\mathcal{T} (D^*, 0)) > 0 \).

Similarly, if \( t^* > 1 - t^* \), then the firm could increase profits by taking control over all actions \( j \in \{1, \ldots, M\} \setminus D^* \) and keeping \( t^* \) unchanged. To see this, note that the change in profits is

\[
\Pi (\mathcal{T} (\{1, \ldots, M\}, t^*)) - \Pi (\mathcal{T} (D^*, t^*)).
\]

If \( t^* < 1 \), then this difference is positive by Lemma 3, because \( 1 > t^* > 1 - t^* \) and \( D^* \neq \{1, \ldots, M\} \) imply \( \mathcal{T} (D^*, t^*) \in [0, 1]^{M+2} \) and \( \mathcal{T} (\{1, \ldots, M\}, t^*) > \mathcal{T} (D^*, t^*) \). If \( t^* = 1 \), then the change in profits can be written

\[
\Pi (\mathcal{T} (\{1, \ldots, M\}, 1)) - \Pi (\mathcal{T} (D^*, 1)) = N \max_{a^1, \ldots, a^M, a^{M+2}} \left\{ R (a^1, \ldots, a^M, 0, a^{M+2}) - \sum_{i \in \{1, \ldots, M, M+2\}} f^i (a^i) \right\} - N \max_{a^1, \ldots, a^M, a^{M+2}} \left\{ R (\tilde{a}^1, \ldots, \tilde{a}^M, 0, \tilde{a}^{M+2}) - \sum_{i \in D^* \cup \{M+2\}} f^i (a^i) \right\} \text{ s.t. } a^i = 0 \text{ if } i \in \{1, \ldots, M\} \setminus D^*.
\]

Clearly \( D^* \neq \{1, \ldots, M\} \) and assumptions (a1)-(a4) (generalized to the current scenario with multiple transferable actions) imply \( \Pi (\mathcal{T} (\{1, \ldots, M\}, 1)) - \Pi (\mathcal{T} (D^*, 1)) > 0 \).

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

**Proof of Proposition 4**

For part (a), if \( a \) is contractible, then the first constraint in (6) and the first constraint in (8) disappear, so the programs (5) and (7) become identical. If the actions \( a_i \) carry no cost \( (f = 0) \), then these first constraints remain distinct in the two modes, unless \( R_{\sigma} = 0 \). Suppose in addition that \( R (a, \sigma, q, Q) \) can be written as \( R1(a, \sigma) + R2(q, Q) \). Then, in stage 2, the equilibrium choices of \( (q, Q) \) as functions
of \( t \) are identical in both modes. Denote them by \((q(t), Q(t))\). The firm’s \( \mathcal{L} \)-mode profits are then

\[
\max_{t,a} \{ N(R_1(a, s(a)) + R_2(q(t), Q(t)) - c(q(t))) - C(Q(t)) \}
\]

subject to

\[
R_1(a, s(a)) + s_a(a) R_{1,\sigma}(a, S(a)) = 0,
\]

which is equal to

\[
\max_{t,a} \{ N(R_1(a, s(a)) + R_2(q(t), Q(t)) - c(q(t))) - C(Q(t)) \}.
\]

This is strictly higher than \( \mathcal{A} \)-mode profits

\[
\max_{t,a} \{ N(R_1(a, s(a)) + R_2(q(t), Q(t)) - c(q(t))) - C(Q(t)) \}
\]

subject to

\[
R_1(a, s(a)) = 0.
\]

For part (b), if the agents’ efforts are contractible or if \( R_q = 0 \), then the firm can achieve the first-best level of profits in \( \mathcal{L} \)-mode by setting \( t = 1 \), obtaining

\[
\Pi^{\mathcal{E}*} = \max_{a,q,Q} \{ N(R(a, s(a), q, Q) - f(a) - c(q)) - C(Q) \}.
\]

In \( \mathcal{A} \)-mode, we know the resulting profits are strictly lower because the choice of \( a \) is not first-best optimal (it does not account for spillovers).

If \( f = 0 \) and \( Q \) is contractible or \( R_Q = 0 \), then the firm can once again achieve the first-best level of profits in \( \mathcal{L} \)-mode, this time by setting \( t^E \) arbitrarily close to 0, obtaining

\[
\Pi^{\mathcal{E}*} = \max_{a,q,Q} \{ N(R(a, s(a), q, Q) - f(a) - c(q)) - C(Q) \}.
\]

In \( \mathcal{A} \)-mode it is also optimal to set \( t^A \) arbitrarily close to 0 but profits are less than first-best because the choice of \( a \) is not first-best optimal (it does not account for spillovers). As a result, \( \Pi^{\mathcal{E}*} > \Pi^{A*} \).