



Incentives for motivated experts in a partnership

Ting Liu^a, Ching-to Albert Ma^b, Henry Y. Mak^{c,*}

^a Department of Economics and College of Business, Stony Brook University, 100 Nicolls Road, Stony Brook, NY 11794, USA

^b Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA

^c Department of Economics, Indiana University–Purdue University Indianapolis, 425 University Boulevard, Indianapolis, IN 46202, USA

ARTICLE INFO

Article history:

Received 21 October 2017

Revised 7 May 2018

Accepted 8 May 2018

Available online 19 June 2018

JEL classification:

D00

D02

D80

D83

Keywords:

Motivated experts

Principal

Multiagent incentives

Partnership

Gatekeeping

Sharing rule

ABSTRACT

A Principal has a set of projects, each having different benefit potentials, and each requiring a basic technology from one of two experts and time inputs from both experts. Experts enjoy motivation utilities from production, but have private information of their own motivation preferences and project potentials. Technology and time-input choices are experts' private decisions. Experts form a Partnership, which designs a sharing rule and a gatekeeping protocol to determine experts' priority on technology choice. Using a linear cost-share contract that lets experts make minimum profits, the Principal implements the first best by delegating all decisions to the Partnership.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

We study a Principal contracting with a group of experts for service or production. The Principal has a set of projects, each having a different potential for producing benefits. Each project requires one expert to be responsible for the basic production technology, and the time inputs of all experts. The project potential determines which expert should provide the basic technology as well as experts' time inputs. Experts observe projects' potentials, but the Principal does not. They also obtain some motivation benefits from service or production, but experts' motivation benefits are unknown to the Principal. Experts decide on who should provide the basic technology, and each expert decides on his own time input.

An example will illustrate our study. Suppose the Principal is a payer who contracts with a generalist physician and a specialist physician to care for a group of chronically ill patients. Patients' illness severities correspond to the production potentials, and these are observed by the physicians, but unknown to the payer. Physicians are altruistic, and enjoy utilities from patients receiving medical benefits. However, their altruistic utilities from patients' benefits vary and are private information. Physicians decide on the basic care setting for each patient. For example, if the illness severity is mild, the generalist takes the role as the primary provider, so care will be at his private office with consultation by the specialist at a hospi-

* Corresponding author.

E-mail addresses: tingliu@stonybrook.edu (T. Liu), ma@bu.edu (C.-t. Albert Ma), makh@iupui.edu (H.Y. Mak).

tal; otherwise, the specialist takes the lead, and the setting will be reversed. The primary-provider decision corresponds to the basic technology choice. The specialist's technology is more advanced but also more costly than the generalist's. After the physicians have decided on the primary provider, they coordinate care by supplying medical services as a team. These services correspond to experts' time inputs.

Clearly, legal, accounting, and consulting professions have similar information and service structures. These are also common in construction, engineering, manufacturing, and even home building industries. In these professions and industries, the physicians' altruistic motivation in the medical field can be likened to professionalism, reputation, or both.

Our model posits two sources of missing information: project potentials and experts' motivation benefits. When the Principal must delegate production decisions to experts, distortions may result from missing information and experts' misaligned incentives of technology and time-input choices. In this paper we show that the Principal can implement the first best by delegating all decisions to an expert-partnership organization with a simple linear contract that partially reimburses the total production costs.

Why can the Principal solve all missing-information and misaligned-incentive problems with delegation and a linear cost contract? The heart of the solution lies in the way experts work together as members of a Partnership. Our concept of a Partnership consists of the following. First, a Partnership is a budget-neutral entity, receiving revenues from the Principal and sharing them among member experts. This is a standard assumption. Second, within the Partnership, experts are symmetrically informed about their motivational preferences and use this information to construct a sharing rule. Experts work closely together, so it is natural for them to know each other's preferences. Third, a Partnership keeps track of technology decisions and each expert's total time inputs, so the splitting of revenues can be made contingent on the collected information. This kind of internal accounting system is common in all firms. Finally, a Partnership can determine a gatekeeping protocol; this determines which among the experts has first priority to take on a project as the primary provider. Gatekeeping is common in the healthcare market, but any hierarchical processing of technology choice by professionals in other markets is similar.

Experts are economic agents, so we assume that each must earn a minimum profit to participate in the Partnership. The experts also enjoy private motivation benefits. In the case of physicians, their altruism is commonly recognized. In the case of professionals and other workers, the motivation benefits may actually come from future profits or their human-capital investment through their current work. In any case, our assumption is that these motivational benefits are private and cannot be capitalized. In other words, the Principal cannot extract any motivational benefits, but still must respect each expert's minimum-profit constraint.

Given the Principal's partial cost-reimbursement contract, for each project, the least motivated experts maximize their joint surplus by the first-best technology and time inputs, and also make the minimum profits. What about more motivated experts? Their higher motivation benefits drive them to maximize their joint surplus by over-utilizing technology and time inputs. However, this will violate experts' minimum-profit constraints. The Partnership then designs the sharing rule for the more motivated experts to deter over-utilization. The sharing rule then satisfies the binding minimum-profit constraint, and implements the first best again. Indeed, the sharing rule stipulates that as an expert becomes more motivated, he receives a smaller profit from adopting his technology and providing time input. The reduced financial benefit acts against the stronger motivation benefit, so more motivated experts are induced to choose the same first best as less motivated experts.

The first-best implementation is robust to many ways in which a Partnership is organized. In the basic model, a General Partnership chooses a profit sharing rule and a project gatekeeping protocol to maximize experts' joint surplus. In a Seniority Partnership, one expert contracts with the Principal, and chooses a sharing rule and a gatekeeping protocol to maximize his own payoff. The Principal can offer a single contract that implements the first best whether experts work together in a General Partnership or a Seniority Partnership.

Results here contrast sharply with the standard solution in the principal-multiagent mechanism design literature; see [Mookherjee \(2006\)](#) for a comprehensive survey. In the canonical model, a Principal designs a reward-punishment scheme which induces each agent to report truthfully his private information, then issues detailed instructions to each agent, and monitors the actions of each agent. We propose an alternative contracting paradigm. We show that the Principal can implement the first best by contracting with a partnership that retains private information and decision-making authority. Even more striking, the Principal achieves this implementation by a *single* contract.

Our model suggests that a Principal benefits from contracting with an expert organization. Partnerships can facilitate the match efficiency between projects and expertise. This is consistent with the empirical evidence in professional-service markets. [Epstein et al. \(2010\)](#) investigate obstetric practices and find that high-risk patients in group practices match with specialists more often than patients in solo practices. Furthermore, this improves patients' health outcomes.

Our analysis has policy bearings for the ongoing health care reform in the United States. As a bedrock of the reform, the Center for Medicare and Medicaid Services (CMS), the federal regulator managing Medicare and Medicaid has been encouraging health care providers to form Accountable Care Organizations (ACOs). In the ACO model, CMS contracts with a group of providers. An ACO is entrusted with coordinating patient care. This often requires sharing of information, consultation, and referrals between physicians within an ACO. Our model predicts that ACOs achieve a higher efficiency and a better care quality if they are paid by a mixture of capitation and cost-sharing contracts rather than by either pure capitation or pure

fee-for-service. Cost sharing between clients and service providers is also gaining popularity in the legal profession, which traditionally has billed by hours.¹

The paper belongs to the large principal-multiagent literature.² The more recent literature has recognized that agents often have diverse interests. Agents may be either intrinsically or extrinsically motivated in their effort choices or use of private information; see, for example, Akerlof and Kranton (2005), Benábou and Tirole (2003), Besley and Ghatak (2005), Francois (2000), Khalil et al. (2013), Makris and Siciliani (2013), Murdock (2002) and Prendergast (2007, 2008). This literature generally studies interaction between motivations and financial incentives. However, these papers assume that agents' motivational preferences are common knowledge.

In our paper, the Principal does not know experts' motivation. As far as we know, in an agency context, the missing motivation information assumption has only been studied by Choné and Ma (2011), Delfgaauw and Dur (2007, 2008), Jack (2005) and Liu and Ma (2013). All these show that missing information about an agent's motivation may lead to distortion, so they study second-best incentive contracts. However, all these models consider only a single agent, whereas here we have many agents, who are assumed to know about each other's motivation.

In Baron and Besanko (1992), Gilbert and Riordan (1995) and Mookherjee and Tsumagari (2004), an organization may consolidate many agents' hidden information of productivity, which may be used by a Principal to prescribe production plans. These papers show that a Principal can benefit from contracting with an informed organization when agents' inputs are complementary, but not when inputs are substitutes. In our model, experts' technologies are substitutes, but their time inputs are complements. However, the Principal can still implement the first best by delegating all production decisions to the Partnership.

Garicano and Santos (2004) and Grassi and Ma (2016) study experts' referrals under asymmetric information. Referrals in these papers are formally similar to expert gatekeeping in ours. In both papers, referrals lead to inefficiency because of adverse selection. Gatekeeping in our model is performed under complete information by motivated experts. We allow the Partnership to determine the equilibrium gatekeeping protocol, whereas in Garicano and Santos (2004) equilibria are derived under different given referral protocols.

The economics literature on Partnership is extensive; see, for example, Holmström (1982), Legros and Matthews (1993), Levin and Tadelis (2005) and Strausz (1999). The usual setup does not consider a Principal contracting with a Partnership. Rather, a Partnership consists of members who have joint ownership. A sharing rule that splits revenues among partners is a scheme to incentivize costly efforts. We go one step further and study how a Principal can influence a Partnership's choices. Furthermore, our concept of a Partnership includes the gatekeeping protocol, which has not been considered in the earlier literature.

The paper proceeds as follows. The next section presents technologies and the first best. Section 3 defines the Partnership's internal organization and the Principal's contract. Section 4 studies Partnership surplus maximization under minimum-profit constraints. Section 5 shows how sharing rules and gatekeeping protocols implement Partnership surplus maximization and the first best simultaneously. In Section 6, one expert becomes a senior partner, and sets up sharing rules and gatekeeping protocols. We show how the Principal continues to implement the first best. Finally, Section 7 draws some conclusions. The Appendix contains proofs of results.

2. Principal and expert services

A Principal has a continuum of production projects with a total mass normalized at one. Projects are defined by a benefit index, b , a random variable distributed on a strictly positive support $[b, \bar{b}]$ with distribution F and density f . The term "project b " means a project with benefit index b . For production, the Principal needs services from two experts; we call them *Expert 1* and *Expert 2*.

Each expert is identified by his production function that requires both experts' inputs, which can be each expert's time spent on the project. Expert j 's production function is defined by $R_j : \mathfrak{R}_+^2 \rightarrow \mathfrak{R}_+$, $j = 1, 2$, where R_j is a strictly increasing and strictly concave function. For each project, at most one expert will be chosen as the primary provider. The output or revenue when Expert j is the primary provider and when Experts 1 and 2, respectively, spend times t_{1j} and t_{2j} on project b is $R_j(t_{1j}, t_{2j})b$, $j = 1, 2$. (For the time input t_{ij} , the first subscript i indicates the identity of the expert who supplies the time, whereas the second subscript j indicates the identity of the primary provider.) Expert 2's production function generates more revenue from the same time inputs, and we will expand on the properties of R_j below. The technology costs of Experts 1 and 2 taking on the role of the primary provider are c_1 and c_2 , respectively, with $0 < c_1 < c_2$, so Expert 2 has a higher technology cost than Expert 1. We assume that experts have identical and constant per-unit time cost.³ We also ignore any capacity constraint, so impose no limits on the number of projects for which an expert can serve as the primary provider, or on the amount of time an expert can spend on a project.

¹ "Using Alternative Fee Arrangements to Increase New Business," Bloomberg Law, <http://about.bloomberglaw.com/practitioner-contributions/using-alternative-fee-arrangements-to-increase-new-business/>.

² Early contributions include (Demski and Sappington, 1984; Holmström, 1982; Ma, 1988; Mookherjee, 1984). The 30-plus years of literature is extensive, so any proper summary is beyond the scope of the current paper.

³ We can let experts' unit time costs be different, experts' time costs be increasing and strictly convex, or both. In either case, there will just be more notation, but no conceptual consequence.

In the Introduction, we have already provided an example in the health industry. There, the generalist physician is Expert 1, and the specialist physician is Expert 2. In other industries, junior lawyers, contractors, and manufacturers take on roles as Expert 1, whereas senior lawyers, architects, and product designers take on roles as Expert 2.

2.1. Complete information

We first consider the Principal's optimal allocation when each project's benefit index b and the experts' actions can be observed by the Principal. The Principal now can dictate the primary-provider assignment and time inputs for each project for maximum surplus. Suppose that Expert j is the primary provider, and Expert i spends time t_{ij} on project b , $i, j = 1, 2$. The surplus is $R_j(t_{1j}, t_{2j})b - c_j - t_{1j} - t_{2j}$. Let $V_j(b)$ be the Principal's maximum surplus from project b when Expert j is the primary provider:

$$V_j(b) \equiv R_j(t_{1j}^*(b), t_{2j}^*(b))b - c_j - t_{1j}^*(b) - t_{2j}^*(b), \tag{1}$$

$$\text{where } (t_{1j}^*(b), t_{2j}^*(b)) \equiv \underset{t_{1j}, t_{2j}}{\operatorname{argmax}} R_j(t_{1j}, t_{2j})b - c_j - t_{1j} - t_{2j} \tag{2}$$

are the optimal times experts should supply to project b when Expert j is the primary provider.⁴

For any (t_1, t_2) , we assume that $R_1(t_1, t_2)$ and $R_2(t_1, t_2)$ satisfy

$$R_2(t_1, t_2) > R_1(t_1, t_2) \tag{3}$$

$$\frac{\partial^2 R_i(t_1, t_2)}{\partial t_1 \partial t_2} \geq 0, \quad i = 1, 2 \tag{4}$$

$$R_2(t'_1, t'_2) - R_2(t_1, t_2) > R_1(t'_1, t'_2) - R_1(t_1, t_2), \quad \text{if } (t'_1, t'_2) \geq (t_1, t_2) \text{ and } (t'_1, t'_2) \neq (t_1, t_2). \tag{5}$$

Expert 1's technology is less powerful than Expert 2's. At any time inputs (t_1, t_2) , Expert 1's technology will generate less output (condition in (3)). Given the assumption on the experts' technology costs, $c_1 < c_2$, it is uninteresting to assume that Expert 1's technology is more powerful. Furthermore, the two experts' time inputs exhibit complementarity (condition (4)). Finally, the production functions R_1 and R_2 exhibit increasing differences (condition (5)). These assumptions allow us to derive useful comparative-static results (see Lemma 1).

An allocation assigns, for each project b , the identity of the primary provider, either Expert 1, Expert 2, or none at all, and the time each expert spends on the project if a primary provider is assigned to it. Projects without a primary provider will not be serviced. The Principal's optimal allocation maximizes its surplus at each b , i.e., $\max\{0, V_1(b), V_2(b)\}$ $b \in [\underline{b}, \bar{b}]$, and we call this the *first best*.⁵

We first present some properties of the optimal times $t_{ij}^*(b)$ and the surplus functions $V_j(b)$.

Lemma 1. *The optimal times $t_{ij}^*(b)$ are increasing in b , $i, j = 1, 2$, and at each b , $t_{11}^*(b) < t_{12}^*(b)$, $i = 1, 2$. Furthermore, the surplus functions $V_j(b)$ when Expert j is the primary provider are both increasing in b , and $V_1'(b) < V_2'(b)$ for all b . The surplus functions $V_j(b)$ are convex.*

Lemma 1 says that each expert's optimal time input increases when the project has a higher benefit index b , and when the primary provider has a more powerful technology. These results stem from production complementary and increasing differences, respectively, (4) and (5). Also, because Expert 1's technology is less powerful than Expert 2's, the surplus from Expert 1 acting as the primary provider, $V_1(b)$, increases at a lower rate than if Expert 2 acts as the primary provider, $V_2(b)$.

The first best may prescribe Expert 1 to be the primary provider for some projects, and Expert 2 for others. Or, it may prescribe only one expert to be the primary provider. Clearly, the first-best allocation in which each expert is the primary provider for some projects is more interesting. To ensure this, we make the following assumption.

Assumption 1. There are two benefit thresholds, b_1^* and b_2^* , such that

$$V_1(b_1^*) = 0 \quad \text{and} \quad V_1(b_2^*) = V_2(b_2^*), \quad \text{and} \quad \underline{b} < b_1^* < b_2^* < \bar{b}. \tag{6}$$

Under this assumption, in the first best i) if $\underline{b} \leq b < b_1^*$, the project is not serviced, ii) if $b_1^* \leq b < b_2^*$, Expert 1 is the primary provider, and iii) if $b_2^* \leq b \leq \bar{b}$, Expert 2 is the primary provider, (each expert also putting in the corresponding optimal time $t_{ij}^*(b)$ $i, j = 1, 2$, for each serviced project). An example of surplus-function configuration satisfying Assumption 1 is illustrated in Fig. 1. For given revenue functions R_j 's in (1) the assumption is satisfied when the fixed cost c_1 is small enough

⁴ Welfare benchmarks in the literature of motivated agents (see Choné and Ma (2011), Delfgaauw and Dur (2007, 2008), Jack (2005) and Liu and Ma (2013) do not include the motivation benefit in the principal's or social planner's objective function because it would have been a form of double counting (see, for example, Hammond (1987) and Milgrom (1993). Motivation benefits are defined below.

⁵ There is a continuum of projects, so all the optimality results refer to pointwise optimization at any b . The definitions of surplus, allocation, and the first best refer to any subset of projects with a positive measure. The same applies to Partnership's surplus and its own optimization. All our results apply to a discrete model when there are a fixed number of projects.

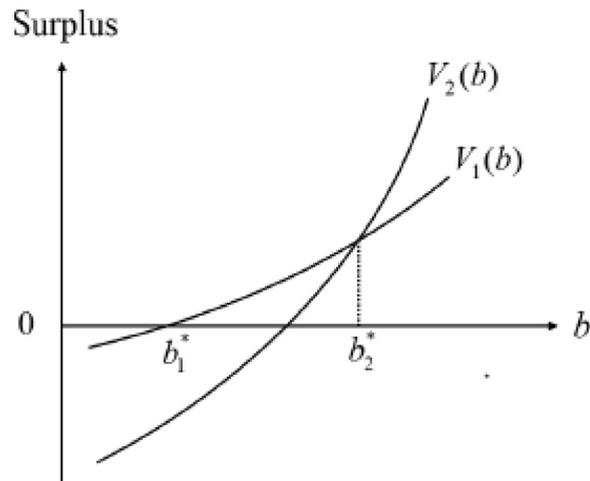


Fig. 1. First best.

and c_2 is not much larger than c_1 . In any case, for revenue functions R_j and costs c_j , (6) in Assumption 1 is equivalent to⁶

$$V_2(\underline{b}) < V_1(\underline{b}) < 0 < V_1(\bar{b}) < V_2(\bar{b}) \quad \text{and} \quad V_1(b_1) = 0 \Rightarrow V_2(b_1) < 0.$$

When Assumption 1 is violated, in the first best one expert will not be a primary provider for any project. In these cases, the surplus function V_2 may lie entirely above V_1 , or it may be entirely below. Expert 1 would be dominated if $b_1^* = b_2^*$, where $V_1(b_1^*) = 0$, so in Fig. 1, the part of V_1 that is above V_2 vanishes. Expert 2 would be dominated if $b_2^* = \bar{b}$, so in Fig. 1, the part of V_2 that is above V_1 disappears. However, our implementation results are robust to surplus-function configurations that violate Assumption 1, and we discuss these cases at the end of Section 4.

3. Motivated experts

We now describe the experts' preferences. Each expert enjoys motivation utility proportional to the output. If the output from using Expert j 's technology for project b is $R_j b$, Expert 1 and Expert 2, respectively, receive utilities $\alpha_1 R_j b$ and $\alpha_2 R_j b$. Here, the parameters α_1 and α_2 are, respectively, Expert 1's and Expert 2's degrees of motivation, which are distributed on strictly positive supports $[\underline{\alpha}_1, \bar{\alpha}_1]$ and $[\underline{\alpha}_2, \bar{\alpha}_2]$. (We do not need notation for distributions and densities.) Motivation parameters interacting linearly with revenues is a common assumption and makes our analysis tractable.⁷

The motivation utilities can be intrinsic or extrinsic. Intrinsic utilities are enjoyments from work. For instance, physicians are altruistic towards their patients. Here, the output $R_j b$ represents a patient's treatment benefit. The technologies R_1 and R_2 are the available medical services. Both physicians spend time with the patient, so each enjoys some utility. As another example, two lawyers working on a *pro bono* case enjoy providing public services. Extrinsic utilities refer to potential future earnings from work done with the Principal. A better outcome with the current Principal may bring in more businesses to each expert in the future. Hence, the motivation utilities may indicate the present values of these future returns.

In any case, utilities from motivation are to be distinguished from the monetary payoffs (profits) experts receive. Intrinsic or extrinsic utilities $\alpha_1 R_j b$ and $\alpha_2 R_j b$ do not count toward monetary profit. In the case of $\alpha_1 R_j b$ and $\alpha_2 R_j b$ representing enjoyment (intrinsic benefits), this is a natural interpretation. In the case of $\alpha_1 R_j b$ and $\alpha_2 R_j b$ representing future earnings (extrinsic benefits), this means that experts cannot borrow against them, which is also natural. We assume that any monetary payoff will add onto the motivation utility in an additive way.

Each expert must attain a reservation utility; this is a constraint universally adopted in agency problems. Suppose that Expert j 's motivation parameter is α_j , $j = 1, 2$ and the (total) output turns out to be Q , his motivation utility will be $\alpha_j Q$. If he also earns profit π_j , his total utility is $\alpha_j Q + \pi_j$. If the reservation utility is \bar{U}_j , the expert does not participate unless $\alpha_j Q + \pi_j \geq \bar{U}_j$. In addition, Expert j may earn a profit $\bar{\pi}_j$ if he does not participate. Because motivation utilities cannot be capitalized, for participation Expert j 's payoff and profit must satisfy $\alpha_j Q + \pi_j \geq \bar{U}_j$ and $\pi_j \geq \bar{\pi}_j$.

Following Liu and Ma (2013), Makris (2009), Makris and Siciliani (2013) and Olivella and Siciliani (2017), we assume that $\pi_j \geq \bar{\pi}_j$ is the only relevant constraint, so $\alpha_j Q + \pi_j \geq \bar{U}_j$ is slack. Effectively, we assume that the reservation utility level \bar{U}_j

⁶ First, because $V_1(\underline{b}) < 0 < V_1(\bar{b})$, by the continuity of V_1 there exists $\underline{b} < b_1 < \bar{b}$ at which $V_1(b_1) = 0$. We set this b_1 to b_1^* in Assumption 1. Second, because $V_2(b_1^*) < 0 = V_1(b_1^*)$ and $V_1(\bar{b}) < V_2(\bar{b})$, by the continuity of V_2 there exists $b_1^* < b_2^* < \bar{b}$ at which $V_1(b_2^*) = V_2(b_2^*)$.

⁷ Obviously, the use of a single parameter is without loss of generality. If Expert i 's motivation is $g_i(\alpha_i)$, where g_i is a continuous and increasing function, we will just redefine $g_i(\alpha_i)$ as the degree of motivation.

is not much higher than the minimum profit $\bar{\pi}_j$. The justification is simply that the expert's outside option rarely offers motivation utility; most jobs merely earn an expert the minimum profit.⁸ Furthermore, an expert considering contracting with the Principal obviously favors the Principal's projects over alternatives. That indicates a better motivation-utility prospect from the Principal's projects than those elsewhere.

3.1. Partnership, budget, and information

The first best requires, for each project, assigning an expert to be the primary provider, and each expert's optimal time given the primary provider. However, the Principal only knows the project benefit distribution and the range of the experts' motivation parameters but not their actual values. Moreover, the Principal can only verify the total cost (technology and time costs) incurred by the two experts after production. The Principal must delegate all the production decisions to the experts and compensate them according to the total cost. Nevertheless, experts work together, so they may be able to share information and incentivize the primary-provider assignment as well as time inputs. We model the work relationship between experts as a Partnership.

A Partnership is a fictitious player with preferences equal to the sum of the experts' payoffs. In addition, we restrict or empower a Partnership in three ways. First, a Partnership does not receive any new revenue other than what the Principal pays the experts, and it cannot dispose of received revenues other than through the experts. A Partnership must split revenues by way of a sharing rule, to be defined below. Second, a Partnership has information of each expert's degree of motivation, α_1 and α_2 , as well as records of how many projects each expert has served as the primary provider, and each expert's total cost (technology and time costs). The information can be used in setting up a sharing rule. Third, the Partnership can decide which of the two experts is a gatekeeper. The gatekeeping protocol is described next, in Stages 2 and 3 of the extensive form.

3.2. Gatekeeping protocol and extensive form

The Partnership decentralizes the primary-provider assignment and time inputs in the following extensive form. Certain details in the extensive form such as the Principal's contract and the Partnership sharing rule will be defined later.

- Stage 0 Nature draws the benefit indexes of the Principal's continuum of projects according to distribution F . Nature draws Expert j 's motivation parameter according to some distribution on $[\underline{\alpha}_j, \bar{\alpha}_j]$, $j = 1, 2$. Realizations of benefit indexes and motivation parameters are unknown to the Principal. The experts' motivation parameters are common knowledge among the experts. The Principal offers a contract to the Partnership.
- Stage 1 If the Partnership rejects the contract, the game ends. If the Partnership accepts the Principal's contract, it selects an expert, say Expert i , $i = 1, 2$, to be the gatekeeper, and sets up a sharing rule to maximize the experts' joint payoff subject to minimum profits.
- Stage 2 For each project, gatekeeper Expert i observes the benefit index b and decides whether to abandon the project, become the primary provider, or let Expert j , $j = 1, 2$ and $j \neq i$, take his turn.
- Stage 3 If gatekeeper Expert i passes on a project to Expert j for the primary-provider decision, Expert j observes the benefit index, and decides whether to abandon the project or become the primary provider.
- Stage 4 Knowing the benefit index, the two experts simultaneously choose the time inputs for each project that has a primary provider. The technology costs and time costs are incurred. The Partnership will be paid by the Principal according to the contract, and the experts will split profit according to the sharing rule.

3.3. Principal's contract and Partnership sharing rule

The Principal can verify the Partnership's total cost. Although complicated cost-sharing contracts can be used, we consider a simple, quasi-linear contract (Γ, s) which consists of a lump sum Γ and specifies that the Partnership is responsible for a fraction s , $0 \leq s \leq 1$, of the total cost.⁹ Let C_1 and C_2 be the total technology and time costs incurred by Experts 1 and 2, respectively. The Principal's contract pays the Partnership a total of $\Gamma + (1 - s)(C_1 + C_2)$. We use the accounting rule that the Partnership bears the experts' total production costs. After subtracting costs, the Partnership nets $\Gamma - s(C_1 + C_2)$.

Experts have more information than the Principal. They know each other's motivation parameters, the total cost incurred by each expert, and the mass of projects for which an expert acts as the primary provider. Experts use this information to design a profit sharing rule. Let $\pi_i(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2)$, $i = 1, 2$ denote the payment to Expert i , where M_1 and M_2 are

⁸ For instance, lawyers contracting with labor unions may enjoy motivation benefits from negotiating for better wage and working conditions, but they are less likely to receive motivation benefit from contracting with private business. Similarly, architects receive motivation benefits from working for the public good by contracting with the government but may not receive motivation benefit from working on a private project. The reservation utility \bar{U}_j is just the profits $\bar{\pi}_j$ an expert earns elsewhere.

⁹ The Principal's strategy set can be expanded to include a recommendation besides the contract (Γ, s) . The recommendation is simply the Partnership's equilibrium gatekeeping protocol and sharing rule (to be derived in Section 5 for the Principal's optimal contract). By definition, the Partnership will play the recommendation. However, this would require more notation, so we have not included it.

the masses of projects for which Expert 1 and 2 are, respectively, primary providers. A sharing rule is defined by the pair (π_1, π_2) such that for any $M_1, M_2, C_1, C_2, \alpha_1$, and α_2 , with $0 \leq M_1 + M_2 \leq 1$,

$$\Gamma - s(C_1 + C_2) = \pi_1(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2) + \pi_2(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2). \quad (7)$$

On the left-hand side of (7), payments from the Principal are independent of the motivation parameters and how the total cost is generated. On the right-hand side of (7), the split of the Principal's payments may depend on each expert's service decisions, but the split must always be equal to the net payment from the Principal's. Expert i 's profit is $\pi_i(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2)$, and his payoff is his profit together with the total motivation benefit.

4. Optimal contract and Partnership surplus

Our analysis proceeds in two steps. First, in this section, we study an auxiliary game which has only two stages. Stage 0 is the same as the extensive form in Section 3.2. In Stage 1, if the Partnership accepts the contract, it chooses an allocation to maximize its surplus subject to the Partnership's minimum-profit constraint, which is also the participation constraint. An allocation specifies which expert, if any, is the primary provider of each project b , and each expert's time input at every b , and experts simply execute the chosen allocation. The Partnership's maximum surplus in the auxiliary game is the upper bound of its surplus in the extensive form game in Section 3.2. This is because experts are obedient in the auxiliary game but need to be incentivized in the extensive form game.

Second, in the following section, we study the experts' incentives when the Partnership cannot dictate their actions. We will show that the Partnership can achieve its maximum surplus as in the auxiliary game by selecting an expert to be a gatekeeper and writing a sharing rule to align experts' incentives.

We define a Partnership's degree of motivation as the sum of those of the experts, namely $\alpha \equiv \alpha_1 + \alpha_2$. We will use the term "Partnership α " to denote a Partnership with degree of motivation α . Given the Principal's contract (Γ, s) , we define

$$W_j(b, \alpha; s) \equiv \max_{t_{1j}, t_{2j}} \alpha R_j(t_{1j}, t_{2j})b - s[c_j + t_{1j} + t_{2j}] \quad (8)$$

to be Partnership α 's surplus (gross of the transfer Γ) when Expert j is the primary provider for project b , and when Expert i with motivation parameter α_i puts in time input t_{ij} , $i, j = 1, 2$. For project b , the Partnership's maximum surplus is $\max\{0, W_1(b, \alpha; s), W_2(b, \alpha; s)\}$.

We begin with the relationship between the first best and Partnership surplus. In fact, the Partnership surplus function W_j in (8) is isomorphic to the first-best surplus function V_j defined in (1) because

$$\alpha R_j(t_{1j}, t_{2j})b - s[c_j + t_{1j} + t_{2j}] = s \left\{ R_j(t_{1j}, t_{2j}) \left[\frac{\alpha}{s} \right] b - c_j - t_{1j} - t_{2j} \right\},$$

so the optimal times for maximizing (8) are $t_{ij}^*([\frac{\alpha}{s}]b)$ from (2). Therefore, we have

$$W_j(b, \alpha; s) = sV_j\left(\left[\frac{\alpha}{s}\right]b\right). \quad (9)$$

It follows that Lemma 1 applies to (8). Partnership α 's maximum surplus is $\max\{0, W_1(b, \alpha; s), W_2(b, \alpha; s)\} = s \max\{0, V_1([\frac{\alpha}{s}]b), V_2([\frac{\alpha}{s}]b)\}$ at benefit index b . For an arbitrary contract (Γ, s) , Partnership α 's surplus gross of transfer Γ is proportional to the Principal's surplus with the project benefit index b modified to $[\frac{\alpha}{s}]b$.

We consider one particular Principal's contract. We will show that under this contract, any Partnership α chooses the first best for maximum surplus under the minimum-profit constraint. This contract is (Γ^*, s^*) :

$$\Gamma^* = \alpha \left[\int_{b_1^*}^{b_2^*} \{c_1 + t_{11}^*(b) + t_{21}^*(b)\} f(b) db + \int_{b_2^*}^{\bar{b}} \{c_2 + t_{12}^*(b) + t_{22}^*(b)\} f(b) db \right] + \bar{\pi}_1 + \bar{\pi}_2, \quad (10)$$

$$s^* = \alpha \equiv \alpha_1 + \alpha_2.$$

The contract says that a Partnership bears a fraction $s^* = \alpha \equiv \alpha_1 + \alpha_2$ of its cost, this fraction being the minimum motivation among all Partnerships; the transfer is its unreimbursed total cost at the first-best allocation and the minimum profits.

We have used $W_j(b, \alpha; s)$ in (8) to denote Partnership surplus when Expert j is the primary provider, and the maximum Partnership surplus is $\max\{0, W_1(b, \alpha; s), W_2(b, \alpha; s)\}$ for project b . However, this Partnership surplus has not considered the minimum-profit constraint. To do that, we use the following modification. Define assignment functions $\phi_i : [b, \bar{b}] \rightarrow [0, 1]$, $i = 1, 2$, each of these functions setting a probability that Expert i is to be the primary provider for project b . We impose the restriction $0 \leq \phi_1(b) + \phi_2(b) \leq 1$. Given the contract (Γ^*, s^*) , Partnership α chooses assignment functions $\phi_i(b)$, and time inputs $t_{ij}(b)$, $i, j = 1, 2$ to maximize:

$$\Gamma^* + \int_{\underline{b}}^{\bar{b}} \phi_1(b)[\alpha R_1(t_{11}(b), t_{21}(b))b - s^*\{c_1 + t_{11}(b) + t_{21}(b)\}]f(b)db + \int_{\underline{b}}^{\bar{b}} \phi_2(b)[\alpha R_2(t_{12}(b), t_{22}(b))b - s^*\{c_2 + t_{12}(b) + t_{22}(b)\}]f(b) db \tag{11}$$

subject to the minimum-profit constraint

$$\Gamma^* - s^* \left[\int_{\underline{b}}^{\bar{b}} \{ \phi_1(b)[c_1 + t_{11}(b) + t_{21}(b)] + \phi_2(b)[c_2 + t_{12}(b) + t_{22}(b)] \} f(b)db \right] \geq \bar{\pi}_1 + \bar{\pi}_2. \tag{12}$$

Our first result describes the surplus of the least motivated Partnership.

Lemma 2. *Given the contract (Γ^*, s^*) , the least motivated Partnership $\underline{\alpha} \equiv \underline{\alpha}_1 + \underline{\alpha}_2$ maximizes its total surplus and makes the minimum profit $\bar{\pi}_1 + \bar{\pi}_2$ by choosing the first-best allocation.*

From (9), Partnership surplus is isomorphic to the Principal’s surplus. However, a Partnership consists of experts who have different degrees of motivation. The contract (Γ^*, s^*) makes Partnership $\underline{\alpha}$ internalize the Principal’s surplus. Partnership $\underline{\alpha}$ values the output at $\underline{\alpha}R_j$ (where R_j is the production function when Expert j is the primary provider), so the Principal’s contract makes Partnership $\underline{\alpha}$ bear $s^* = \underline{\alpha}$ of the total cost. Then the lump sum Γ^* makes sure that Partnership $\underline{\alpha}$ makes the minimum profits. Lemma 2 simply says that the least motivated Partnership’s surplus is maximized by the first best. How about other Partnerships with higher motivation $\alpha \in (\underline{\alpha}, \bar{\alpha}]$? The next lemma says that for these Partnerships, the minimum-profit constraint must bind.

Lemma 3. *Given the contract (Γ^*, s^*) , the minimum-profit constraint (12) binds for each Partnership $\alpha > \underline{\alpha}$.*

Facing the same cost-share rate $s^* = \underline{\alpha}$, a Partnership values outputs more when its degree of motivation increases. In fact, if the minimum-profit constraint (12) does not bind, Partnership $\alpha > \underline{\alpha}$ must respond by serving more projects and investing more time inputs, and consequently making less profit.

Next, because the minimum-profit constraint (12) must bind, we write the objective function (11) as

$$\alpha \int_{\underline{b}}^{\bar{b}} \{ \phi_1(b)R_1(t_{11}(b), t_{21}(b))b + \phi_2(b)R_2(t_{12}(b), t_{22}(b))b \} f(b)db + \bar{\pi}_1 + \bar{\pi}_2. \tag{13}$$

Partnership α chooses ϕ_i and t_{ij} , $i, j = 1, 2$ to maximize (13) subject to the minimum-profit constraint (12), which, we emphasize, is independent of α . Obviously, an allocation is a solution of Partnership α ’s maximization of (13) subject to the binding constraint (12) if and only if it is a solution of Partnership $\hat{\alpha}$ ’s maximization of

$$\hat{\alpha} \int_{\underline{b}}^{\bar{b}} \{ \phi_1(b)R_1(t_{11}(b), t_{21}(b))b + \phi_2(b)R_2(t_{12}(b), t_{22}(b))b \} f(b)db + \bar{\pi}_1 + \bar{\pi}_2$$

subject to the same binding constraint (12), any α and $\hat{\alpha}$. But we know that at $\alpha = \underline{\alpha}$, the first best maximizes the Partnership’s surplus and the minimum-profit constraint binds. Lemma 3 says that the minimum-profit constraint of any Partnership $\hat{\alpha}$ with $\hat{\alpha} > \underline{\alpha}$ must also bind, so we have our main result in this section (proof omitted):

Proposition 1. *Given the contract (Γ^*, s^*) , a Partnership of any degree of motivation $\alpha \equiv \alpha_1 + \alpha_2$ maximizes the Partnership total surplus and makes the minimum profits by choosing the first-best allocation.*

Proposition 1 says that, under the contract (Γ^*, s^*) , the first-best allocation maximizes the surplus of all Partnerships α subject to the minimum-profit constraint. It is a striking result, yet is based simply on monotonicity. Higher motivation encourages more service provisions and time inputs. Each Partnership, however, must respect the minimum-profit constraint. When the least motivated Partnership just breaks even, a more motivated Partnership must find the minimum-profit constraint binding.

To close this section, we document the changes in the Principal’s contract when Assumption 1 is violated. Then we have the two cases that are described at the end of Section 2.1. First, when the surplus function V_1 is dominated by V_2 , we replace b_1^* by b_2^* (where $V_2(b_2^*) = 0$) in (10). Second, when the surplus function V_2 is dominated by V_1 , we replace b_2^* by \bar{b} for the contract in (10). All results in this section remain valid after these changes. When one expert is never the primary provider for any projects, replacing b_1^* by b_2^* or b_2^* by \bar{b} are all that are needed for adjusting the sharing rules for implementation.

5. Partnership sharing rule and implementation

By Proposition 1, if a Partnership could dictate experts’ actions upon accepting the Principal’s contract, it would prescribe the first-best allocation and achieve the first-best payoff. The Partnership, however, can only choose a gatekeeper and a sharing rule in the extensive form in Section 3.2. Consider a subgame starting at Stage 2, defined by a gatekeeper and a sharing rule. Suppose that this subgame has a unique subgame-perfect equilibrium first-best outcome with the Partnership making the minimum profits. Then any equilibrium of the extensive form must also be first best. In the next two subsections, we construct such a subgame, in which Expert 1 is the gatekeeper. There may well be other subgames that achieve the first best. In these other subgames, Expert 2 is the gatekeeper; we will come to these equilibria later.

5.1. Gatekeeper Expert 1 and sharing rule

The Principal's contract (Γ^*, s^*) pays a lump sum Γ^* and $(1 - s^*)$ of the total cost. A sharing rule splits the net revenue between the experts in terms of (M_1, M_2, C_1, C_2) , where $M_i, i = 1, 2$, are the masses of projects for which Expert i is the primary provider, and where C_i are Expert i 's total technology and time costs. The sharing rule must also be budget balanced, as in (7). Experts' preferences are common knowledge within the Partnership and their motivation parameters are used in the sharing rule.

Suppose that in Stage 1 the Partnership decides that Expert 1 is the gatekeeper (so Expert 1 gets priority to decide how projects are to be processed in Stage 2, while Expert 2 only gets his turn in Stage 3). We construct a budget-balanced sharing rule that implements the first best:

$$\pi_1(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2) = \Delta - \alpha_1 C_1 + \alpha_2 C_2 - s^* C_2 - \gamma_1 M_1 + \gamma_2 M_2 \tag{14}$$

$$\pi_2(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2) = \Gamma^* - \Delta + \alpha_1 C_1 - s^* C_1 - \alpha_2 C_2 + \gamma_1 M_1 - \gamma_2 M_2, \tag{15}$$

where

$$\gamma_1 \equiv (\alpha_1 + \alpha_2 - s^*) t_{21}^*(b_1^*) \tag{16}$$

$$\gamma_2 \equiv \alpha_2 V_2(b_2^*) + (\alpha_1 + \alpha_2 - s^*) t_{12}^*(b_2^*), \tag{17}$$

b_1^* and b_2^* are the first-best thresholds, and Δ is some constant.¹⁰

The sharing rule requires more information than the Principal's contract (which is based only on the total costs), but is quasi-linear (same as the Principal's contract). First, the sharing rule depends on experts' motivation parameters. Second, it keeps track of how many projects for which each expert has chosen to be the primary provider. Third, it depends on the total time and technology costs incurred by each expert. Under our accounting rule, the Partnership bears all the costs, and receives all the revenues from the Principal, so experts' net payments are entirely determined by the sharing rule. The sum of (14) and (15) is simply $\Gamma^* - s^*(C_1 + C_2)$, so it is always budget balanced as required by (7).

There are two basic components: each expert's total costs, C_1 and C_2 , and the project masses for which experts are primary providers, M_1 and M_2 . Experts' decisions on time inputs have no effect on M_1 and M_2 . However, if Expert j decides to be a primary provider, he changes both M_j and C_j .

Condition (14) asks Expert 1 to bear α_1 of any of his cost C_1 , so he bears α_1 of his time costs incurred in Stage 4. Similarly, condition (15) asks Expert 2 to bear α_2 of his time costs incurred in Stage 4. What about Expert 2's decisions in Stage 3? According to (15), if he is the primary provider for project b , he will raise the measure M_2 by one unit, so he incurs a cost γ_2 . Also, he will bear α_2 of his technology and time cost, but he receives $(\alpha_1 - s^*)$ of Expert 1's time input costs. Expert 2's net payment, if he takes on project b , is $-\alpha_2[c_2 + t_{22}(b)] - \gamma_2 + (\alpha_1 - s^*)t_{12}(b)$.

Expert 1's decisions in Stage 2 are more complicated. First, he has the option to take on project b . If he does so, he raises the measure M_1 by one, so pays γ_1 . He also has to bear the increment in technology and time costs C_1 at rate α_1 . The cost C_2 will also increase by Expert 2's time input, and Expert 1 receives $(\alpha_2 - s^*)$ of that increase. His net payment is $-\alpha_1[c_1 + t_{11}(b)] - \gamma_1 + (\alpha_2 - s^*)t_{21}(b)$.

Expert 1's second option is to pass on project b to Expert 2. Assume that Expert 2 will then take on the project. Condition (14) gets Expert 1 to bear his own time cost $\alpha_1 t_{12}(b)$, but it also gets him to receive an amount $\gamma_2 + (\alpha_2 - s^*)[c_2 + t_{22}(b)]$ due to Expert 2's cost and his being the primary provider. Expert 1's net payment is $-\alpha_1 t_{12}(b) + \gamma_2 + (\alpha_2 - s^*)[c_2 + t_{22}(b)]$. Finally, Expert 1 may just abandon the project and nets 0.

5.2. Gatekeeper Expert 1 and implementation

We have just laid out the experts' net payment consequences in each stage. Now we derive the subgame-perfect equilibria. In Stage 4, only time inputs are to be decided on any project b when an expert has chosen to be a primary provider. We have seen that the sharing rule makes Expert i bear α_i of his time cost. Because Expert i gets a motivational benefit at α_i of the output, he internalizes the social cost and benefit. The following result is immediate.

Lemma 4. *In the subgame-perfect equilibrium in Stage 4, each expert chooses the first-best time inputs $t_{ij}^*(b)$ for every project b when Expert j is the primary provider, $i, j = 1, 2$.*

Lemma 4 pins down the values of $t_{ij}(b)$ in the two experts' continuation equilibrium payoffs in Stage 4. It also implies that the first best is implemented when there are γ_1 and γ_2 so that Experts 1 and 2 take on projects as primary providers if and only if $b_1^* \leq b < b_2^*$ and $b \geq b_2^*$, respectively.

¹⁰ $\Delta = \bar{\pi}_1 + \int_{b_1^*}^{b_2^*} \{\alpha_1(c_1 + t_{11}^*(b)) + \gamma_1 - (\alpha_2 - s^*)t_{21}^*(b)\} f(b)db + \int_{b_2^*}^{\bar{b}} \{\alpha_1 t_{12}^*(b) - \gamma_2 - (\alpha_2 - s^*)(c_2 + t_{22}^*(b))\} f(b)db$.

In Stage 3, Expert 2 decides whether to become a primary provider or abandon the projects that Expert 1 has chosen to pass on. If Expert 2 chooses to abandon a project, his incremental payoff is zero. If Expert 2 chooses to be the primary provider for project b , he assumes the continuation equilibrium in Lemma 4. The net payment can be obtained from Section 5.1. Together with the motivation benefit, Expert 2's payoff will be $\alpha_2[R_2(t_{12}^*(b), t_{22}^*(b))b - c_2 - t_{22}^*(b)] - \gamma_2 + (\alpha_1 - s^*)t_{12}^*(b)$. We use the value of γ_2 in (17) and the continuation equilibrium payoffs to show the following.

Lemma 5. *In the subgame-perfect equilibrium in Stage 3, Expert 2 takes on a project b as the primary provider if and only if $b \geq b_2^*$.*

Our choice of γ_2 in (17) and the continuation equilibrium payoffs in Stage 4 simplify Expert 2's payoff from being the primary provider to

$$\alpha_2[V_2(b) - V_2(b_2^*)] + (\alpha_1 + \alpha_2 - s^*)[t_{12}^*(b) - t_{12}^*(b_2^*)], \tag{18}$$

and the steps for this expression are in the proof of the lemma. Given the sharing rule and the continuation equilibrium in Stage 4, Expert 2 internalizes any surplus difference between project b and the threshold project b_2^* (the first term), as well as any difference in Expert 1's time input costs (the second term). Because both V_2 and t_{12}^* are increasing in b and $s^* = \alpha_1 + \alpha_2$, Expert 2's decision in Stage 3 will implement the first best for all projects with benefit parameters above b_2^* .

In Stage 2, Expert 1 chooses between abandoning a project, becoming the primary provider, and passing it on. Then Expert 2's decision will result in a continuation equilibrium described by Lemma 5 in Stage 3, and then Lemma 4 in Stage 4. First, if Expert 1 chooses to take on project b , his payoff is $\alpha_1[R_1(t_{11}^*(b), t_{21}^*(b))b - c_1 - t_{11}^*(b)] - \gamma_1 + (\alpha_2 - s^*)t_{21}^*(b)$, which is the motivation benefit plus the net payment specified in Section 5.1. Second, if Expert 1 passes the project on to Expert 2, then, according to Lemma 5, Expert 2 will abandon it if $b < b_2^*$, but will provide service otherwise. Expert 1's payoff when $b > b_2^*$ is $\alpha_1[R_2(t_{12}^*(b), t_{22}^*(b))b - t_{12}^*(b)] + \gamma_2 + (\alpha_2 - s^*)[t_{22}^*(b) + c_2]$, where, again, we have added the motivation benefit to the net payment described in Section 5.1. We use the values of γ_1 in (16), γ_2 in (17) and the continuation equilibrium payoffs to show the following.

Lemma 6. *In the subgame-perfect equilibrium in Stage 2, Expert 1 serves project b as the primary provider if and only if $b_1^* \leq b < b_2^*$ and passes on project b to Expert 2 if and only if $b_2^* \leq b \leq \bar{b}$.*

We use γ_1 in (16) and the continuation equilibrium payoffs in Stage 4 to simplify Expert 1's payoff from being the primary provider to

$$\begin{aligned} &\alpha_1[R_1(t_{11}^*(b), t_{21}^*(b))b - c_1 - t_{11}^*(b)] - \gamma_1 + (\alpha_2 - s^*)t_{21}^*(b) \\ &= \alpha_1V_1(b) + (\alpha_1 + \alpha_2 - s^*)(t_{21}^*(b) - t_{21}^*(b_1^*)), \end{aligned} \tag{19}$$

where the steps for the equality are spelled out in the proof of the lemma. Payoff (19) says that Expert 1 fully internalizes the surplus from project b and the difference in Expert 2's time costs for project b and the cutoff project b_1^* when Expert 1 is the primary provider. Because $V_1(b_1^*) = 0$, and V_1, t_{21}^* are increasing in b , Expert 1's interest aligns with the first best for all projects with indexes below b_2^* .

Next, we use γ_2 in (17) and the continuation equilibrium in Stage 3 to simplify Expert 1's equilibrium payoff if he passes on a project $b > b_2^*$ to Expert 2:

$$\begin{aligned} &\alpha_1[R_2(t_{12}^*(b), t_{22}^*(b))b - t_{12}^*(b)] + (\alpha_2 - s^*)[t_{22}^*(b) + c_2] + \gamma_2 \\ &= \alpha_1V_2(b) + (\alpha_1 + \alpha_2 - s^*)[t_{12}^*(b_2^*) + t_{22}^*(b) + c_2] + \alpha_2V_2(b_2^*) \\ &> \alpha_1V_1(b) + (\alpha_1 + \alpha_2 - s^*)(t_{21}^*(b) - t_{21}^*(b_1^*)), \end{aligned} \tag{20}$$

where the steps for the equality are in the proof. Here, the choice of γ_2 and the more productive Expert 2 actually let Expert 1 gain more than what he can internalize from being the primary provider. For projects $b > b_2^*$, our choice of γ_2 in (17) aligns both Expert 1's incentive to pass on a project and Expert 2's incentive to be primary provider. In Fig. 2, we plot Expert 1's equilibrium payoffs in (19) for being the primary provider for all projects in the lower graph, and Expert 1's equilibrium payoffs in (20) when Expert 2 is the primary provider for projects $b > b_2^*$ in the upper graph.

Our preceding three lemmas completely characterize the subgame-perfect equilibrium outcomes in the subgame defined by gatekeeper Expert 1 and the sharing rule in (14) to (17). Projects $b < b_1^*$ will be abandoned by either expert. Projects with b between b_1^* and b_2^* will be served by Expert 1 (and rejected by Expert 2). Projects with b higher than b_2^* will be served by Expert 2 (and also served by Expert 1—off the equilibrium path). First-best time inputs will be supplied to those projects with a primary provider. Any subgame-perfect equilibrium outcome is first best. Finally, recall that we have used a constant Δ in the definition of the sharing rule. Its value is defined in footnote ¹⁰, and makes sure that in any subgame-perfect equilibrium, each expert earns the minimum profit. The three lemmas together say that gatekeeper Expert 1 and the sharing rule implement the allocation that maximizes Partnership surplus in any continuation equilibrium.¹¹ Hence, it is a best response for the Partnership to use Expert 1 as a gatekeeper and the sharing rule in (14) to (17). We conclude by stating the result formally.

¹¹ The lemmas do not say that the continuation equilibrium must be unique. Projects with benefit indexes below b_1^* must be rejected, but either expert will do that. Rather trivial multiple equilibria exist: in one equilibrium, Expert 1 rejects those projects; in another, Expert 2 does.

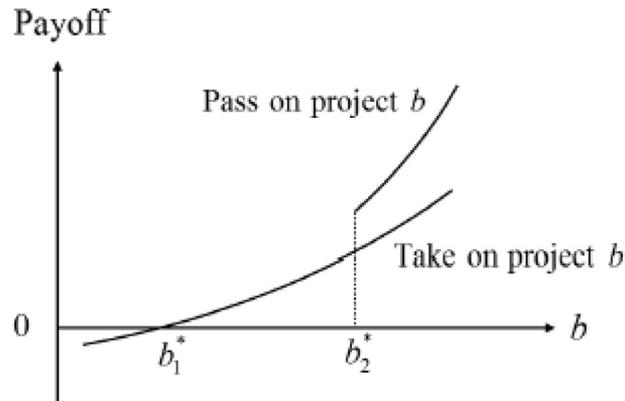


Fig. 2. Gatekeeper Expert 1's payoff in Stage 2.

Proposition 2. Suppose that the Principal offers the contract (Γ^*, s^*) . In the extensive form in Section 3.2, there is a subgame-perfect equilibrium in which the Partnership chooses Expert 1 as the gatekeeper and the budget-balanced sharing rule defined by (14)–(17), and the subgame-perfect equilibrium outcome is first best, with Expert 1's equilibrium profit being $\bar{\pi}_1$, and Expert 2's equilibrium profit being $\bar{\pi}_2$. Furthermore, because this equilibrium outcome also maximizes Partnership surplus, every subgame-perfect equilibrium outcome must be first best.

5.3. Gatekeeper Expert 2 and sharing rule

What about subgames in which Expert 2 is the gatekeeper? It turns out that the first best may be a continuation equilibrium when the Principal offers the same contract (Γ^*, s^*) under some conditions. In these cases, there are multiple subgame-perfect equilibria.

Lemma 7. Suppose that the Principal offers the contract (Γ^*, s^*) . If Expert 2 is the gatekeeper and $t_{12}^*(b) - t_{11}^*(b)$ is nondecreasing in b , there is a sharing rule so that the subgame-perfect equilibrium outcome is the first best with each expert making his minimum profit.

Any budget-balanced, quasi-linear sharing rule must take the form in (14) and (15). For first-best time inputs, each expert must internalize the time cost and motivational benefit. Experts 1 and 2 have motivation parameters α_1 and α_2 , so Expert 1 should bear α_1 of the cost C_1 in (14), and Expert 2 should bear α_2 of the cost C_2 in (15). The other terms that involve costs C_1 and C_2 in (14) and (15) are there to maintain budget balance. Then there remain only two other instruments for other incentives, namely γ_1 and γ_2 , and the terms involving project masses M_1 and M_2 must also maintain budget balance. Proposition 2 says that such instruments, γ_1 and γ_2 , are sufficient for first-best implementation when Expert 1 is the gatekeeper. In fact, the piece rate γ_2 in (17) ensures that Expert 2 will only provide service to projects $b > b_2^*$. Hence, Expert 1's decision is nontrivial only for projects $b > b_2^*$.

The same consideration does not hold when Expert 2 is the gatekeeper. For first-best implementation, Expert 1 must be willing to be the primary provider for projects with b between b_1^* and b_2^* . Because Expert 1's payoff is monotone increasing in b , Expert 1 must also be willing to be the primary provider for projects $b > b_2^*$. Therefore, Expert 2, as a gatekeeper, faces a nontrivial decision for a wider range of projects, all those with benefit indexes between b_1^* and \bar{b} .

When $t_{12}^*(b) - t_{11}^*(b)$ is nondecreasing in b , Expert 2's gain from taking on a project against passing on the project is increasing in b and the first best can be implemented. Here, the comparison between Expert 2's payoffs from taking on project b as a primary provider and passing on the project is illustrated in the left panel of Fig. 3. Nevertheless, when $t_{12}^*(b) - t_{11}^*(b)$ is decreasing in b , Expert 2's gain from being a primary provider for a project against passing it on may not be monotone increasing. As a result, Expert 2 may choose to be the primary provider for projects with very high and very low benefits, but pass on projects with intermediate benefits. This possibility is illustrated in the right panel of Fig. 3. In this case, the first best is not a continuation equilibrium.

6. Seniority Partnerships

We have used the Partnership as a fictitious player to represent the charter and governance set up by the experts when they form the organization. The first-best implementation results, however, are more general, and apply to alternate governance forms. Now we consider two other partnership designs. In each, one expert becomes the Senior Partner, whereas the other is the Junior Partner. (We will use the term General Partnership to refer to the Partnership-as-a-fictitious-player organization.)

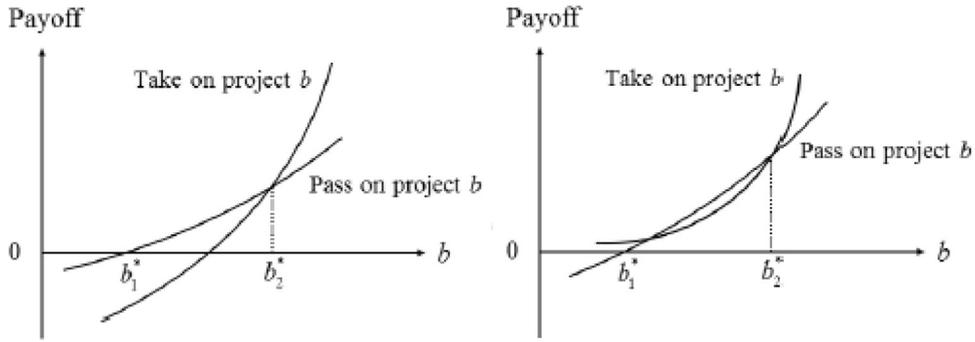


Fig. 3. Gatekeeper Expert 2's payoff in Stage 2.

The Senior Partner and the Junior Partner are joint owners, and they continue to split revenues and costs according to a sharing rule. However, the Senior Partner decides on accepting and rejecting the Principal's contract, the gatekeeping protocol, and the sharing rule. The Senior Partner will make these decisions based on his own preferences (instead of the aggregated preferences of experts). However, we maintain the same information structure and contractible states as in the implementation under General Partnership, so the functional form of the sharing rule remains the same as in (14) and (15).

We study, in turn, Expert 1 as the Senior Partner, and Expert 2 as the Senior Partner. In each case, the extensive form is identical to the one in Section 3.2 except that in Stage 1, the Senior Partner (rather than the fictitious Partnership) makes all the decisions. Can the Principal entrust a Seniority Partnership to implement the first best? How does the seniority governance affect the Principal's contract, gatekeeping protocol, and the sharing rule?

6.1. Senior Partner Expert 1

The first best can be implemented by the following Principal's cost-sharing contract $(\widehat{\Gamma}, \widehat{s})$:

$$\widehat{\Gamma} = \alpha_1 \left[\int_{b_1^*}^{b_2^*} \{c_1 + t_{11}^*(b) + t_{21}^*(b)\} f(b) db + \int_{b_2^*}^{\bar{b}} \{c_2 + t_{12}^*(b) + t_{22}^*(b)\} f(b) db \right] + \bar{\pi}_1 + \bar{\pi}_2, \tag{21}$$

$$\widehat{s} = \alpha_1.$$

This new contract is obtained by replacing $\alpha \equiv \alpha_1 + \alpha_2$ in (10) by α_1 ; this is the only difference. The Principal now uses the least motivated Expert 1 to construct the contract.

Given the Principal's contract $(\widehat{\Gamma}, \widehat{s})$, Senior Partner Expert 1 chooses a gatekeeping protocol and a sharing rule in Stage 1 to maximize his own payoff in the continuation equilibrium. As in Sections 4 and 5, we show implementation in two steps. First, we consider the surplus that Expert 1 can achieve if he can dictate an allocation. We obtain a result analogous to Proposition 1: given the Principal's contract in (21), Senior Partner Expert 1's payoff is maximized by the first best, with both experts making the minimum profits. Second, we show that Expert 1 can implement the first best in the extensive form game by choosing a gatekeeper and a sharing rule in Stage 1.

When Senior Partner Expert 1 can dictate an allocation, the technology cost and time inputs that Expert 2 is made to incur may be transferred to Expert 1. We assume, however, that Expert 2's motivation benefit remains private, and nontransferable. Hence, if Expert 1 lets Expert j be the primary provider for project b , Expert 1's payoff, net of any fixed payment to Expert 2, is $\alpha_1 R_j(t_{1j}, t_{2j})b - \widehat{s}(c_j + t_{1j} + t_{2j})$ for time inputs t_{1j} and t_{2j} . Let $W_j^1(b, \alpha_1; \widehat{s}) \equiv \max_{t_{1j}, t_{2j}} \alpha_1 R_j(t_{1j}, t_{2j})b - \widehat{s}(c_j + t_{1j} + t_{2j})$;

this is Expert 1's payoff from optimal time inputs when he chooses Expert j to be the primary provider, $j = 1, 2$. For any project b , Expert 1's surplus is $\max\{0, W_1^1(b, \alpha_1; \widehat{s}), W_2^1(b, \alpha_1; \widehat{s})\}$, net of any fixed payment. Expert 1's surplus is isomorphic to the Partnership surplus in (8), which has the form $\max\{0, W_1(b, \alpha; s), W_2(b, \alpha; s)\}$ (with $\alpha = \alpha_1 + \alpha_2$), so Expert 1 is like a General Partnership with (total) motivation parameter α_1 . The following lemma follows from an adaptation of results in Section 4.

Lemma 8. Suppose that Expert 1 chooses an allocation to maximize his own payoff subject to minimum profits. If the Principal offers the contract $(\widehat{\Gamma}, \widehat{s})$ in Stage 0, Expert 1 of any degree of motivation maximizes his payoff by choosing the first best.

Lemma 8 has the same logic as Lemmas 2 and 3 (which prove Proposition 1). If Senior Partner Expert 1 could dictate all decisions, he would internalize Expert 2's technology costs and time inputs, and would pay Expert 2 the minimum profit. Hence, Senior Partner Expert 1 with motivation parameter α_1 is like a General Partnership with motivation parameter $\alpha = \alpha_1 + 0$. But now the contract $(\widehat{\Gamma}, \widehat{s})$ from the Principal is simply adjusting for this lower degree of motivation from (Γ^*, s^*) , the one in (10). The first best maximizes the least motivated Expert 1's surplus, and his minimum-profit constraint just

binds. The binding minimum-profit constraint deters a more motivated Expert 1 from rendering services more generous than the first best.

Can Expert 1 choose a gatekeeping protocol and design a corresponding sharing rule in Stage 1 to implement the first best? The following proposition gives an affirmative answer.

Proposition 3. *Given the Principal's contract $(\widehat{\Gamma}, \widehat{s})$, there is a subgame-perfect equilibrium in which Senior Partner Expert 1 works as the gatekeeper and sets the following sharing rule*

$$\pi_1(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2) = \Upsilon - \alpha_1 C_1 + \alpha_2 C_2 - \widehat{s} C_2 - \gamma_1 M_1 + \gamma_2 M_2 \quad (22)$$

$$\pi_2(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2) = \widehat{\Gamma} - \Upsilon + \alpha_1 C_1 - \widehat{s} C_1 - \alpha_2 C_2 + \gamma_1 M_1 - \gamma_2 M_2, \quad (23)$$

where

$$\gamma_1 \equiv (\alpha_1 + \alpha_2 - \widehat{s}) t_{21}^*(b_1^*) \quad (24)$$

$$\gamma_2 \equiv \alpha_2 V_2(b_2^*) + (\alpha_1 + \alpha_2 - \widehat{s}) t_{12}^*(b_2^*), \quad (25)$$

for some constant Υ . Any subgame-perfect equilibrium outcome is first best, and gives each expert his minimum profit.

The sharing rule for implementation is essentially identical to the one in Section 5.1. There, the General Partnership uses the sharing rule (14)–(17) in response to the Principal's contract (Γ^*, s^*) . The piece rates γ_1 and γ_2 and the transfer Δ are functions of the cost share $s^* = \alpha_1 + \alpha_2$. When Expert 1 is the Senior Partner, the Principal adjusts the cost share to $\widehat{s} = \alpha_1$. In response, Expert 1 changes the sharing rule by replacing s^* in (14)–(17) by \widehat{s} . Because the sharing rule in (22)–(25) is essentially the same as (14)–(17), the two experts' incentives in Stages 3 and 4 remain unchanged. The equilibrium outcome must be the first-best allocation.

For brevity, we do not study those equilibria in which Senior Partner Expert 1 assigns Expert 2 as the gatekeeper. From Lemma 7, the equilibrium outcome may not always be the first best in the general Partnership. We expect the same difficulty when Expert 1 is the Senior Partner.

6.2. Senior Partner Expert 2

The case of Expert 2 being the Senior Partner works in the same logic. We simply reverse the experts' roles in the previous subsection. The analysis works in an analogous way. First, the Principal replaces α_1 in the contract $(\widehat{\Gamma}, \widehat{s})$ in (21) by α_2 , the lowest of Expert 2's degrees of motivation. Lemma 8 directly applies. Senior Partner Expert 2 of any degree of motivation chooses the first best to maximize his payoff and the two experts make their minimum profits. Implementation remains the same as when Expert 1 is the Senior Partner. Expert 2 chooses Expert 1 as the gatekeeper, switches the value of \widehat{s} in the sharing rule (22)–(25) from α_1 to α_2 , and adjusts the lump sum transfers accordingly.

6.3. Robust Principal's contract

We have shown that the first best can be implemented under either Expert 1 or Expert 2 Seniority Partnerships. The Principal has used different contracts for the first-best implementation. In fact, we now describe an even stronger result: a single Principal's contract can implement the first best whether experts form a General Partnership or a Seniority Partnership.

We have assumed that Expert i 's motivation parameter is drawn from support $[\alpha_i, \bar{\alpha}_i]$, a strictly positive interval. Suppose that the Principal, perhaps overly conservatively, thought that the support was really $[\alpha'_i, \bar{\alpha}_i]$, with $0 < \alpha'_i < \alpha_i$. That is, the Principal believed that the least motivated expert was less generous than Expert α_i . The Principal would have constructed the contracts using the lower motivation parameters, in General and Seniority Partnerships. However, Propositions 2 and 3 continue to hold! The reason is that for any General and Seniority Partnerships having larger motivation parameters, their minimum-profit constraints bind, and maximize surpluses by the first-best allocation. The following is immediate.

Proposition 4. *If the Principal offers the contract*

$$\widetilde{\Gamma} = \widetilde{\alpha} \left[\int_{b_1^*}^{b_2^*} \{c_1 + t_{11}^*(b) + t_{21}^*(b)\} f(b) db + \int_{b_2^*}^{\bar{b}} \{c_2 + t_{12}^*(b) + t_{22}^*(b)\} f(b) db \right] + \bar{\pi}_1 + \bar{\pi}_2, \quad (26)$$

$$\widetilde{s} = \widetilde{\alpha},$$

where $0 < \widetilde{\alpha} < \min\{\alpha_1, \alpha_2\} < \alpha_1 + \alpha_2 = \alpha$, any of the General or Senior Partnerships with motivation higher than $\widetilde{\alpha}$ makes the minimum profits and implements the first best as a subgame-perfect equilibrium outcome.

The Proposition implies a discontinuity at $\alpha = 0$. That is, if the support of an expert’s motivation parameter is $[0, \bar{\alpha}_j]$, a contract that sets $\tilde{s} = 0$ in (26) cannot achieve the first best. At $\tilde{s} = 0$ in (26), experts are just paid their minimum profits as lump sums, but do not bear any costs. The Partnership surplus in (8) now becomes $W_j(b, \alpha, 0) \equiv \max_{t_{1j}, t_{2j}} \alpha R_j(t_{1j}, t_{2j})b$.

Clearly, the Partnership surplus is no longer isomorphic to the first-best surplus. However, for any $\tilde{\alpha} > 0$, the isomorphism in (9) holds. A Partnership with any $\tilde{\alpha} > 0$ still internalizes the first-best surplus multiplied by $\tilde{\alpha} > 0$; the minimum-profit constraint still deters them from deviating from the first best. The Principal can get arbitrarily close to the first best by setting a positive \tilde{s} arbitrarily close to 0. Those Partnerships with $\alpha < \tilde{s}$ will not implement the first best.¹² However, Proposition 4 does imply that our results are robust. One simple contract can implement (or approximately implement) the first best at General and Seniority Partnerships. The Principal does not have to know the precise organization form.

7. Conclusion

We show how a Principal can successfully delegate project production decisions to experts who operate in a Partnership. Despite the lack of project information and experts’ preferences, the Principal uses a single cost share contract for the delegation. Our theory is predicated on how the Partnership is assumed to operate. First, within the Partnership, experts share information about projects and their own preferences. Second, the Partnership can set up a sharing rule that is based on each expert’s individual time inputs and technology choices. Third, the Partnership can set up a gatekeeping protocol to structure technology decision priorities.

The Principal’s contract is contingent on the Partnership’s total production cost. In fact, the Principal keeps track of nothing about each individual expert’s technology and time-input choices. Moreover, the contract’s reimbursement rate only depends on the preference of the least motivated Partnership. So, the Principal does not need to know experts’ motivation distribution. The Partnership sharing rule adjusts for differences in experts’ preferences. Also, the Partnerships weakly prefers to give technology-choice priority to the low-cost-low-productivity Expert 1. This gatekeeping protocol—generalist physician as gatekeeper referring to specialist physician—is commonly observed in the health care market.

Clearly, many extensions of our model are possible. An obvious one concerns projects that require more than two experts. In this case, Partnerships will also have more than two experts. The first best can be straightforwardly defined. However, gatekeeping can take place in a hierarchical fashion, and one expert can pass on a project either up or down the hierarchy. Sharing rules must also be enriched. Our preliminary work suggests a monotonicity result: efficiency can be achieved when hierarchical priority starts from the least-cost-least-productive expert and goes up.

We have assumed that experts’ reservation utilities are not much higher than their minimum profits, so the minimum-profit constraint is the participation constraint. This also means that experts’ outside options are independent of their types, a common assumption in agency models. The experience in the extant agency literature is that relaxing this assumption to allow type-dependent reservation utilities tends to be difficult, but is clearly of interest. Finally, experts play a game of complete information within the Partnership (although they do choose time inputs simultaneously). Any incomplete information about motivation benefits would necessarily complicate the model. Again, this possible exploration is of interest but for future research.

Acknowledgement

We thank Editor Daniel Houser, and a reviewer for their advice and suggestions. For their many comments, we thank seminar participants at Boston University, Columbia University, Hong Kong University of Science and Technology, Indiana University–Purdue University Indianapolis, Lehigh University, Michigan State University, Stony Brook University, University of Cologne, University of New South Wales, University of Quebec in Montreal, Virginia Tech University, 12th Annual International Industrial Organization Conference, and the 2014 North American Summer Meeting of the Econometric Society. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A

Proof of Lemma 1: Define a function $y : \mathfrak{N}_+^2 \times [b, \bar{b}] \times \{1, 2\} \rightarrow \mathfrak{N}_+$ by $y(t_1, t_2; b, j) = R_j(t_1, t_2)b - c_j - t_1 - t_2$. First, by (4), the function y is supermodular in (t_1, t_2) . Second, consider (t'_1, t'_2) and (t_1, t_2) with $(t'_1, t'_2) \geq (t_1, t_2)$ and $(t'_1, t'_2) \neq (t_1, t_2)$, and the difference $y(t'_1, t'_2; b, j) - y(t_1, t_2; b, j) = [R_j(t'_1, t'_2)b - c_j - t'_1 - t'_2] - [R_j(t_1, t_2)b - c_j - t_1 - t_2]$. Clearly, the difference is increasing in b . The difference is also increasing in j because $b[R_2(t'_1, t'_2) - R_2(t_1, t_2)] > b[R_1(t'_1, t'_2) - R_1(t_1, t_2)]$ by (5). We can now apply Theorem 5 in Milgrom and Shannon (1994), p.164. The optimal times $t_{ij}^*(b)$ are increasing in b and satisfy $t_{i1}^*(b) < t_{i2}^*(b)$, $i = 1, 2$.

¹² The discontinuity is reminiscent of the classic approximate first-best result of Mirrlees’ moral hazard model (Mirrlees, 1999). There the Principal can induce a costly effort by a scheme that almost always fully insures the Agent. But full insurance will only implement the lowest effort.

By the envelope theorem, we have $\frac{dV_j(b)}{db} = R_j(t_{1j}^*(b), t_{2j}^*(b)) > 0$. By (3) and the first part of the lemma, we have $R_1(t_1, t_2) < R_2(t_1, t_2)$ and $t_{i1}^*(b) < t_{i2}^*(b)$, $i = 1, 2$. Therefore,

$$R_1(t_{11}^*(b), t_{21}^*(b)) < R_2(t_{11}^*(b), t_{21}^*(b)) < R_2(t_{12}^*(b), t_{22}^*(b)), \text{ hence } V_1'(b) < V_2'(b).$$

The convexity of V_j follows from the positive partial derivatives of R_j and the optimal times $t_{ij}^*(b)$ increasing in b .

Proof of Lemma 2: First, we ignore the minimum-profit constraint for Partnership $\underline{\alpha}$. By (8), the unconstrained maximization of (11) yields the surplus $\max\{0, W_1(b, \underline{\alpha}; s^*), W_2(b, \underline{\alpha}; s^*)\}$ by pointwise optimization at b . By (9), we have

$$\begin{aligned} \max\{0, W_1(b, \underline{\alpha}; s^*), W_2(b, \underline{\alpha}; s^*)\} &= \max\left\{0, s^*V_1\left(\left[\frac{\underline{\alpha}}{s^*}\right]b\right), s^*V_2\left(\left[\frac{\underline{\alpha}}{s^*}\right]b\right)\right\} \\ &= \max\{0, s^*V_1(b), s^*V_2(b)\} = s^* \max\{0, V_1(b), V_2(b)\}. \end{aligned}$$

We conclude that Partnership $\underline{\alpha}$ chooses the first best for the unconstrained maximization of (11).

It remains to show that the omitted minimum-profit constraint is satisfied. Partnership $\underline{\alpha}$'s total cost is

$$\begin{aligned} &s^* \left[\int_b^{\bar{b}} \{\phi_1(b)[c_1 + t_{11}(b) + t_{21}(b)] + \phi_2(b)[c_2 + t_{12}(b) + t_{22}(b)]\} f(b) db \right] \\ &= s^* \left[\int_{b_1^*}^{b_2^*} \{c_1 + t_{11}^*(b) + t_{21}^*(b)\} f(b) db + \int_{b_2^*}^{\bar{b}} \{c_2 + t_{12}^*(b) + t_{22}^*(b)\} f(b) db \right]. \end{aligned}$$

According to the contract (Γ^*, s^*) , this total cost is completely reimbursed, so the Partnership nets $\bar{\pi}_1 + \bar{\pi}_2$. We conclude that the minimum-profit constraint is satisfied.

Proof of Lemma 3: Suppose not; that is, suppose that the minimum-profit constraint for Partnership $\alpha > \underline{\alpha}$ does not bind. Partnership $\alpha > \underline{\alpha}$ then chooses an allocation to solve the unconstrained maximization of (11). Using pointwise optimization, we have the maximum surplus at project b given by

$$\begin{aligned} \max\{0, W_1(b, \alpha; s^*), W_2(b, \alpha; s^*)\} &= \max\left\{0, s^*V_1\left(\left[\frac{\alpha}{s^*}\right]b\right), s^*V_2\left(\left[\frac{\alpha}{s^*}\right]b\right)\right\} \\ &= \max\left\{0, s^*V_1\left(\left[\frac{\alpha}{\underline{\alpha}}\right]b\right), s^*V_2\left(\left[\frac{\alpha}{\underline{\alpha}}\right]b\right)\right\} \\ &= s^* \max\left\{0, V_1\left(\left[\frac{\alpha}{\underline{\alpha}}\right]b\right), V_2\left(\left[\frac{\alpha}{\underline{\alpha}}\right]b\right)\right\}. \end{aligned}$$

Recall the definitions of V_j and t_{ij}^* in (1) and (2). By Lemma 1, t_{ij}^* is increasing in b , so $t_{ij}^*\left(\left[\frac{\alpha}{\underline{\alpha}}\right]b\right) > t_{ij}^*(b)$ because $\alpha \equiv \alpha_1 + \alpha_2 > \underline{\alpha} \equiv \underline{\alpha}_1 + \underline{\alpha}_2$. Furthermore, again from Lemma 1, $0 < V_1'(b) < V_2'(b)$. Therefore, if at b , we have $0 \leq V_1(b)$, we also have $0 < V_1\left(\left[\frac{\alpha}{\underline{\alpha}}\right]b\right)$. This means that Partnership α must provide service to more projects than in the first best. Also, if at b , we have $V_1(b) \leq V_2(b)$, we also have $V_1\left(\left[\frac{\alpha}{\underline{\alpha}}\right]b\right) < V_2\left(\left[\frac{\alpha}{\underline{\alpha}}\right]b\right)$. Expert 2 must be the primary provider for more projects than the first best.

The total cost incurred by Partnership α must be strictly more than the cost at first best

$$s^* \left[\int_{b_1^*}^{b_2^*} \{c_1 + t_{11}^*(b) + t_{21}^*(b)\} f(b) db + \int_{b_2^*}^{\bar{b}} \{c_2 + t_{12}^*(b) + t_{22}^*(b)\} f(b) db \right].$$

Therefore, the profit it earns from contract (Γ^*, s^*) is less than $\bar{\pi}_1 + \bar{\pi}_2$, which is a contradiction. We conclude that the minimum-profit constraint (12) binds.

Proof of Lemma 4 : Consider project b and suppose that Expert j is the primary provider. From the discussion in Section 5.1, in Stage 4, parts of experts' payoffs related to time inputs are $\alpha_1[R_j(t_{1j}, t_{2j})b - t_{1j}]$ and $\alpha_2[R_j(t_{1j}, t_{2j})b - t_{2j}]$. The equilibrium time inputs $(\hat{t}_{1j}(b), \hat{t}_{2j}(b))$ are mutual best responses:

$$\hat{t}_{1j}(b) = \max_{t_{1j}} \alpha_1[R_j(t_{1j}, \hat{t}_{2j}(b))b - t_{1j}] \quad \text{and} \quad \hat{t}_{2j}(b) = \max_{t_{2j}} \alpha_2[R_j(\hat{t}_{1j}(b), t_{2j})b - t_{2j}]. \tag{27}$$

Clearly, $(\hat{t}_{1j}(b), \hat{t}_{2j}(b))$ are best responses if and only if they solve

$$\max_{t_{1j}, t_{2j}} R_j(t_{1j}, t_{2j})b - t_{1j} - t_{2j} - c_j,$$

which is the definition of the first-best time inputs in (2).

Proof of Lemma 5: Expert 2's equilibrium payoff from taking on project b is

$$\begin{aligned} & \alpha_2[R_2(t_{12}^*(b), t_{22}^*(b))b - c_2 - t_{22}^*(b)] - \gamma_2 + (\alpha_1 - s^*)t_{12}^*(b) \\ & = \alpha_2[R_2(t_{12}^*(b), t_{22}^*(b))b - c_2 - t_{22}^*(b)] + (\alpha_1 - s^*)t_{12}^*(b) - \alpha_2V_2(b_2^*) - (\alpha_1 + \alpha_2 - s^*)t_{12}^*(b_2^*) \\ & = \alpha_2[R_2(t_{12}^*(b), t_{22}^*(b))b - c_2 - t_{12}^*(b) - t_{22}^*(b)] + (\alpha_1 + \alpha_2 - s^*)t_{12}^*(b) - \alpha_2V_2(b_2^*) - (\alpha_1 + \alpha_2 - s^*)t_{12}^*(b_2^*), \end{aligned}$$

where the first equality follows from the definition of γ_2 in (17), and the second is due to the term $\alpha_2 t_{12}^*(b)$ being added and subtracted. Next, we use the definition of V_2 in (1) to simplify the above to

$$\alpha_2[V_2(b) - V_2(b_2^*)] + (\alpha_1 + \alpha_2 - s^*)[t_{12}^*(b) - t_{12}^*(b_2^*)].$$

By the monotonicity of V_2 and t_{12}^* in b in Lemma 1, and the identity $s^* \equiv \alpha_1 + \alpha_2$, the above is positive if and only if $b \geq b_2^*$. We conclude that Expert 2 chooses to be a primary provider if and only if $b \geq b_2^*$.

Proof of Lemma 6: Suppose Expert 1 has project b . If he decides to be primary provider, his payoff is

$$\begin{aligned} & \alpha_1[R_1(t_{11}^*(b), t_{21}^*(b))b - c_1 - t_{11}^*(b)] - \gamma_1 + (\alpha_2 - s^*)t_{21}^*(b) \\ & = \alpha_1[R_1(t_{11}^*(b), t_{21}^*(b))b - c_1 - t_{11}^*(b) - t_{21}^*(b)] - \gamma_1 + (\alpha_1 + \alpha_2 - s^*)t_{21}^*(b) \\ & = \alpha_1V_1(b) + (\alpha_1 + \alpha_2 - s^*)(t_{21}^*(b) - t_{21}^*(b_1^*)), \end{aligned}$$

where the first equality follows from adding and subtracting $t_{21}^*(b)$, and where the second equality follows from our choice of γ_1 in (16). Clearly, from Assumption 1 and Lemma 1, we have $V_1(b) > 0$ and $t_{21}^*(b) - t_{21}^*(b_1^*) > 0$ if and only if $b > b_1^*$. In the continuation equilibrium in Lemma 5, Expert 2 will abandon projects with indexes below b_2^* . It follows that for any $b < b_2^*$, Expert 1's best response is that in the lemma.

Next, if $b > b_2^*$ and he passes the project on to Expert 2. According to Lemma 5, Expert 2 will be primary provider, so Expert 1's payoff is

$$\begin{aligned} & \alpha_1[R_2(t_{12}^*(b), t_{22}^*(b))b - t_{12}^*(b)] + (\alpha_2 - s^*)[t_{22}^*(b) + c_2] + \gamma_2 \\ & = \alpha_1[R_2(t_{12}^*(b), t_{22}^*(b))b - t_{12}^*(b) - t_{22}^*(b) - c_2] + (\alpha_1 + \alpha_2 - s^*)[t_{22}^*(b) + c_2] + \gamma_2 \\ & = \alpha_1V_2(b) + (\alpha_1 + \alpha_2 - s^*)[t_{12}^*(b_2^*) + t_{22}^*(b) + c_2] + \alpha_2V_2(b_2^*). \end{aligned}$$

This payoff is strictly bigger than $\alpha_1V_1(b) + (\alpha_1 + \alpha_2 - s^*)[t_{21}^*(b) - t_{21}^*(b_1^*)]$ for $b > b_2^*$, Expert 1's payoff from being primary provider (because $V_2(b) > V_1(b)$, and $t_{22}^*(b) > t_{21}^*(b)$). We conclude that Expert 1 strictly prefers to pass on any project $b > b_2^*$ to Expert 2. Expert 1's best response for project $b > b_2^*$ is as described in the lemma.

Proof of Lemma 7: We prove the lemma by showing that a sharing rule can implement the first best in the new extensive-form game if $t_{12}^*(b) - t_{11}^*(b)$ is nondecreasing in b . The sharing rule is

$$\tilde{\pi}_1(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2) = \tilde{\Delta} - \alpha_1C_1 + (\alpha_2 - s^*)C_2 - \tilde{\gamma}_1M_1 + \tilde{\gamma}_2M_2 \tag{28}$$

$$\tilde{\pi}_2(M_1, M_2, C_1, C_2; \alpha_1, \alpha_2) = \Gamma^* - \tilde{\Delta} + (\alpha_1 - s^*)C_1 - \alpha_2C_2 + \tilde{\gamma}_1M_1 - \tilde{\gamma}_2M_2, \tag{29}$$

where

$$\tilde{\gamma}_1 \equiv -(\alpha_1 + \alpha_2 - s^*)[t_{11}^*(b_1^*) + c_1] \tag{30}$$

$$\tilde{\gamma}_2 \equiv (\alpha_1 + \alpha_2 - s^*)[t_{12}^*(b_2^*) - t_{11}^*(b_2^*) + t_{11}^*(b_1^*)] \tag{31}$$

$$\begin{aligned} \tilde{\Delta} & \equiv \bar{\pi}_1 + \int_{b_1^*}^{b_2^*} \{\alpha_1(c_1 + t_{11}^*(b)) + \tilde{\gamma}_1 - (\alpha_2 - s^*)t_{21}^*(b)\}f(b)db \\ & \quad + \int_{b_2^*}^{\bar{b}} \{\alpha_1t_{12}^*(b) - \tilde{\gamma}_2 - (\alpha_2 - s^*)(c_2 + t_{22}^*(b))\}f(b)db. \end{aligned} \tag{32}$$

First, observe that (28), (29) are identical to (14), (15) except for the values of piece rates $\tilde{\gamma}_1, \tilde{\gamma}_2$ and constant $\tilde{\Delta}$. Hence, (28) and (29) must also be budget-balanced. Lemma 4 also straightforwardly applies to Stage 4 of the new extensive-form game. Given (28) and (29), each expert must choose the first-best time inputs $t_{ij}^*(b)$ for every project b .

Now consider Expert 1's decisions in Stage 3. Based on (28), Expert 1's incremental payoff from serving project b he has received from Expert 2 is

$$\begin{aligned} \tilde{u}_1(b) & \equiv \alpha_1R_1(t_{11}^*(b), t_{21}^*(b))b - \alpha_1(c_1 + t_{11}^*(b)) + (\alpha_2 - s^*)t_{21}^*(b) - \tilde{\gamma}_1 \\ & = \alpha_1V_1(b) + (\alpha_1 + \alpha_2 - s^*)t_{21}^*(b) - \tilde{\gamma}_1. \end{aligned} \tag{33}$$

By Lemma 1, both $V_1(b)$ and $t_{21}^*(b)$ are increasing in b . Now substitute the definition of $\tilde{\gamma}_1$ in (30) into (33) and evaluate $\tilde{u}_1(b)$ at b_1^* , we have $\tilde{u}_1(b_1^*) > 0$. Because $\tilde{u}_1(b)$ is continuous and increasing in b , there exists a cutoff value $\tilde{b}_1 < b_1^*$ such that $\tilde{u}_1(b) \geq 0$ if and only if $b \geq \tilde{b}_1$.

In Stage 2, Expert 2 chooses among abandoning a project, becoming the primary provider for the project, and passing on the project to Expert 1. By (29), Expert 2's incremental gain from taking on project b is

$$\tilde{u}_2(b) \equiv \alpha_2R_2(t_{12}^*(b), t_{22}^*(b))b - \alpha_2(t_{22}^*(b) + c_2) + (\alpha_1 - s^*)t_{12}^*(b) - \tilde{\gamma}_2$$

$$= \alpha_2 V_2(b) + (\alpha_1 + \alpha_2 - s^*) t_{12}^*(b) - \tilde{\gamma}_2, \tag{34}$$

whereas his incremental gain from passing on a project to Expert 1 is

$$\begin{aligned} \tilde{u}_2^p(b) &\equiv \alpha_2 R_1(t_{11}^*(b), t_{21}^*(b))b - \alpha_2 t_{21}^*(b) + (\alpha_1 - s^*)(t_{11}^*(b) + c_1) + \tilde{\gamma}_1 \\ &= \alpha_2 V_1(b) + (\alpha_1 + \alpha_2 - s^*)(c_1 + t_{11}^*(b)) + \tilde{\gamma}_1 \end{aligned} \tag{35}$$

if $b \geq \tilde{b}_1$, and $\tilde{u}_2^p(b) = 0$ if $b < \tilde{b}_1$. By Lemma 1, $V_2(b)$, $t_{12}^*(b)$, $V_1(b)$, and $t_{11}^*(b)$ are increasing in b , so both (34) and (35) are increasing in b .

We first consider $\tilde{u}_2^p(b)$. Substitute $\tilde{\gamma}_1$ defined in (30) into (35) and evaluate $\tilde{u}_2^p(b)$ at b_1^* , we have $\tilde{u}_2^p(b_1^*) = 0$ because $V_1(b_1^*) = 0$ by Assumption 1. We conclude that $\tilde{u}_2^p(b) > 0$ if $b > \tilde{b}_1$ and $\tilde{u}_2^p(b) = 0$ if $b \leq \tilde{b}_1$.

We next compare $\tilde{u}_2(b)$ with $\tilde{u}_2^p(b)$. For each project $b \geq \tilde{b}_1$, the difference between $\tilde{u}_2(b)$ and $\tilde{u}_2^p(b)$ is

$$\begin{aligned} \Delta \tilde{u}_2(b) &\equiv \tilde{u}_2(b) - \tilde{u}_2^p(b) \\ &= \alpha_2 (V_2(b) - V_1(b)) + (\alpha_1 + \alpha_2 - s^*)(t_{12}^*(b) - t_{11}^*(b) - c_1) - \tilde{\gamma}_1 - \tilde{\gamma}_2 \\ &= \alpha_2 (V_2(b) - V_1(b)) + (\alpha_1 + \alpha_2 - s^*)\{t_{12}^*(b) - t_{11}^*(b)\} - \{t_{12}^*(b_2^*) - t_{11}^*(b_2^*)\}. \end{aligned} \tag{36}$$

Because $V_2(b_2^*) = V_1(b_2^*)$ by Assumption 1, we have $\Delta \tilde{u}_2(b_2^*) = 0$. Differentiating (36) with respect to b yields

$$\frac{\partial \Delta \tilde{u}_2(b)}{\partial b} = \alpha_2 (V_2'(b) - V_1'(b)) + (\alpha_1 + \alpha_2 - s^*) \frac{\partial}{\partial b} [t_{12}^*(b) - t_{11}^*(b)].$$

By Lemma 1, we have $V_2'(b) - V_1'(b) > 0$. If $t_{12}^*(b) - t_{11}^*(b)$ is nondecreasing in b , $\Delta \tilde{u}_2(b)$ must be increasing in b for $b \geq \tilde{b}_1$.

Now suppose that $t_{12}^*(b) - t_{11}^*(b)$ is nondecreasing in b . We have

$$\max\{\tilde{u}_2(b), \tilde{u}_2^p(b), 0\} = \begin{cases} \tilde{u}_2(b) & \text{if } b_2^* \leq b \leq \bar{b} \\ \tilde{u}_2^p(b) & \text{if } b_1^* \leq b < b_2^* \\ 0 & \text{if } \underline{b} \leq b < b_1^* \end{cases}$$

Hence, Expert 2 maximizes his payoff by abandoning a project if and only if $b < b_1^*$, passing on a project to Expert 1 if and only if $b_1^* \leq b < b_2^*$, and leading the project as the primary provider if and only if $b_2^* \leq b$. The first best is a continuation equilibrium. Finally, it is easy to verify that given the value of $\tilde{\Delta}$ in (32), the two experts make the minimum profits by implementing the first best.

Proof of Lemma 8: We let Expert 1 be responsible for technology cost and time input costs incurred by Expert 2, and pay a lump sum equal to $\bar{\pi}_2$ to satisfy Expert 2's minimum-profit constraint. Given the contract $(\hat{\Gamma}, \hat{\delta})$, Expert 1's payoff from assigning Expert j , $j = 1, 2$, as the primary provider for project b is

$$\begin{aligned} W_j^1(b, \alpha_1; \hat{\delta}) &= W_1(b, \alpha_1 + 0; \hat{\delta}) \\ &= \hat{\delta} V_j\left(\left[\frac{\alpha_1}{\hat{\delta}}\right]b\right), \end{aligned}$$

where the first equality follows from (8) and the second equality follows from (9).

By the same argument in the proof of Lemma 2, Expert 1 with motivation parameter α_1 , the least motivated Senior Partner, maximizes his payoff by choosing the first-best allocation. Moreover, his minimum-profit constraint just binds,

$$\hat{\Gamma} - \hat{\delta} \left[\int_{b_1^*}^{b_2^*} \{c_1 + t_{11}^*(b) + t_{21}^*(b)\} f(b) db + \int_{b_2^*}^{\bar{b}} \{c_2 + t_{12}^*(b) + t_{22}^*(b)\} f(b) db \right] - \bar{\pi}_2 = \bar{\pi}_1.$$

Consider an Expert 1 who is more motivated, $\alpha_1 > \alpha_1$. Replacing s^* and $\alpha_1 + \alpha_2$ in the proof for Lemma 3 by $\hat{\delta}$ and α_1 , respectively, we see that Expert 1 with motivation parameter α_1 must have a binding minimum-profit constraint. The modified Lemmas 2 and 3 imply a modified Proposition 1, in which contract (Γ^*, s^*) is replaced by $(\hat{\Gamma}, \hat{\delta})$, and any degree of motivation $\alpha \equiv \alpha_1 + \alpha_2$ is replaced by α_1 . This is the statement of the lemma.

References

Akerlof, G.A., Kranton, R.E., 2005. Identity and the economics of organizations. *J. Econ. Perspect.* 19, 9–32.
 Baron, D.P., Besanko, D., 1992. Information, control, and organizational structure. *J. Econ. Manag. Strat.* 1, 237–275.
 Benábou, R., Tirole, J., 2003. Intrinsic and extrinsic motivation. *Rev. Econ. Stud.* 70, 489–520.
 Besley, T., Ghatak, M., 2005. Competition and incentives with motivated agents. *Am. Econ. Rev.* 95, 616–636.
 Choné, P., Ma, C.A., 2011. Optimal health care contract under physician agency. *Annales d'Économie et de Statistique* 229–256. 101/102.
 Delfgaauw, J., Dur, R., 2007. Signaling and screening of workers' motivation. *J. Econ. Behav. Organiz.* 62, 605–624.
 Delfgaauw, J., Dur, R., 2008. Incentives and workers' motivation in the public sector. *Econ. J.* 118, 171–191.
 Demski, J.S., Sappington, D., 1984. Optimal incentive contracts with multiple agents. *J. Econ. Theory* 33, 152–171.
 Epstein, A., Ketcham, J., Nicholson, S., 2010. Specialization and matching in professional services firms. *RAND J. Econ.* 41, 811–834.
 Francois, P., 2000. Public service motivation' as an argument for government provision. *J. Public Econ.* 78, 275–299.

- Garicano, L., Santos, T., 2004. Referrals. *Am. Econ. Rev.* 94, 499–525.
- Gilbert, R.J., Riordan, M.H., 1995. Regulating complementary products: A comparative institutional analysis. *RAND J. Econ.* 26, 243–256.
- Grassi, S., Ma, C.A., 2016. Information acquisition, referral, and organization. *RAND J. Econ.* 47, 935–960.
- Hammond, P.J., 1987. Altruism. In: Eatwell, J., Milgate, M., Newman, P. (Eds.), *The New Palgrave: A Dictionary of Economics*. Palgrave Macmillan, London, UK.
- Holmström, B., 1982. Moral hazard in teams. *Bell J. Econ.* 13, 324–340.
- Jack, W., 2005. Purchasing health care services from providers with unknown altruism. *J. Health Econ.* 24, 73–93.
- Khalil, F., Kim, D., Lawarrée, J., 2013. Contracts offered by bureaucrats. *RAND J. Econ.* 44, 686–711.
- Legros, P., Matthews, S.A., 1993. Efficient and nearly-efficient partnerships. *Rev. Econ. Stud.* 60, 599–611.
- Levin, J., Tadelis, S., 2005. Profit sharing and the role of professional partnerships. *Q. J. Econ.* 120, 131–171.
- Liu, T., Ma, C.A., 2013. Health insurance, treatment plan, and delegation to altruistic physician. *J. Econ. Behav. Organiz.* 85, 79–96.
- Ma, C.A., 1988. Unique implementation of incentive contracts with many agents. *Rev. Econ. Stud.* 55, 555–572.
- Makris, M., 2009. Incentives for motivated agents under an administrative constraint. *J. Econ. Behav. Organiz.* 71, 428–440.
- Makris, M., Siciliani, L., 2013. Optimal incentive schemes for altruistic providers. *J. Public Econ. Theory* 15, 675–699.
- Milgrom, P., 1993. Is sympathy an economic value? philosophy, economics, and the contingent valuation method. In: Hausman, J.A. (Ed.), *Contingent Valuation: A Critical Assessment*. North Holland, Amsterdam, the Netherlands.
- Milgrom, P., Shannon, C., 1994. Monotone comparative statics. *Econometrica* 62, 157–180.
- Mirrlees, J.A., 1999. The theory of moral hazard and unobservable behaviour: Part i. *Rev. Econ. Stud.* 66, 3–21.
- Mookherjee, D., 1984. Optimal incentive schemes with many agents. *Rev. Econ. Stud.* 51, 433–446.
- Mookherjee, D., 2006. Decentralization, hierarchies, and incentives: A mechanism design perspective. *J. Econ. Literature* 44, 367–390.
- Mookherjee, D., Tsumagari, M., 2004. The organization of supplier networks: Effects of delegation and intermediation. *Econometrica* 72, 1179–1219.
- Murdock, K., 2002. Intrinsic motivation and optimal incentive contracts. *RAND J. Econ.* 33, 650–671.
- Olivella, P., Siciliani, L., 2017. Reputational concerns with altruistic providers. *J. Health Econ.* 55, 1–13.
- Prendergast, C., 2007. The motivation and bias of bureaucrats. *Am. Econ. Rev.* 97, 180–196.
- Prendergast, C., 2008. Intrinsic motivation and incentives. *Am. Econ. Rev.* 98, 201–205.
- Strausz, R., 1999. Efficiency in sequential partnerships. *J. Econ. Theory* 85, 140–156.