Submitted to *Management Science*
manuscript

Evaluating the Efficacy of Providers' Compensation Contracts in Improving Participant Retention for Clinical Studies

Xueze Song
University of Illinois Urbana-Champaign, xuezes2@illinois.edu

Mili Mehrotra
University of Illinois Urbana-Champaign, milim@illinois.edu

Tharanga Rajapakshe
University of Florida, tharanga.rajapakshe@warrington.ufl.edu

In this work, we aim to analyze a clinical study sponsor's decisions regarding monetary payments to participants and compensation for providers (investigators and coordinators) for their efforts to improve participant retention in the study. To this end, we first consider a centralized model where the sponsor decides the monetary payments to participants and the providers' efforts. We then identify the optimal contracts for the providers under the two decentralized team structures: the sponsor-investigator (SI) model and the outsourcing (OM) model. We further analyze three widely adopted compensation contracts for the providers—fixed (FC), linear (LC), and conditional linear (CLC) given a decentralized structure. Our theoretical analysis shows that the expected retention cost with optimal contracts under decentralized structures is at most 40% higher than that under the centralized model. However, in practical instances, this cost increase is, at the most, 8% on average. A comparison of the FC, LC, and CLC contracts reveals that it is sufficient for the sponsor to choose between the FC and the LC contracts under the SI model, whereas, under the OM model, there exist cases where the sponsor is better off adopting the CLC contract. Further, the sponsor's expected retention cost when choosing the best of the three contracts is at most 6% (10%) higher on average relative to that for the optimal compensation contract under the SI (OM) model. Given a decentralized structure, we also identify cases where the optimal contract offers significant benefit over the three contracts observed in practice.

Key words: healthcare operations, clinical study, participant retention, compensation contracts

1. Introduction

A clinical study is a research investigation involving human subjects to evaluate interventions such as medical products, procedures, or changes to participants' behavior (ClinicalTrials.gov 2020). Ensuring the retention of a specified number of participants is pivotal for a study's scientific validity and economic feasibility. However, between 15% to 40% of participants enrolled in clinical studies drop out prematurely (Atlant Clinical 2020, Nuttall 2012). Reasons for participant dropout include but are not limited to financial costs, lack of understanding and engagement, and inconvenience (mdgroup

2020, National Research Council 2010). The stakeholders of a clinical study influence these reasons to drop out, and hence, in this work, we focus on analyzing the actions of various stakeholders who play an active role in improving participant retention.

The key stakeholders of a clinical study include the sponsor, investigator, clinical research coordinator, and participants. A sponsor is an entity that initiates and funds the clinical study and is responsible for contracting with qualified investigators and coordinators, providing information on treatment to the investigators, and ensuring proper study monitoring. An investigator designs, conducts, and manages the clinical study, while a coordinator is responsible for various tasks such as scheduling participant visits, explaining consent forms and study engagements to participants, and collecting data. Together, the investigator and the coordinator are also responsible for ensuring compliance with the study protocol and participants' safety. Participants have the right to withdraw their consent for participation and leave the study before the completion of all treatment and measurements.

The sponsor must collaborate with the investigator and the coordinator (*providers*) to conduct a clinical study. The structure of a clinical study team varies based on the study's scale, the sponsor's business focus, the geographical distribution of the eligible participant pools, etc. Depending on the relationship among the sponsor, the investigator, and the coordinator, we identify three team structures commonly observed in practice: (1) the centralized model where pharmaceutical companies, hospitals, universities, and other organizations may fund their employees to conduct clinical studies (e.g., Lin 2021, Tucker 2021, Schleider 2021), (2) the sponsor-investigator (SI) model where a sponsor may both initiate the study and investigate the treatment himself (i.e., a sponsor-investigator) but contract with an external entity to coordinate the study; one typical example is when the research faculty carries out clinical studies on their research projects (e.g., Hunsley 2020, McCabe 2020, Lutgendorf 2019), and (3) the outsourcing (OM) model where the sponsor may contract with other organizations to outsource the clinical study investigation and coordination (UCSF 2022, ElectroCore INC 2019); for instance, Pfizer has been collaborating with several organizations, including ICON and Parexel, to outsource its clinical study activities (Pfizer 2011).

The sponsor can provide monetary payment to participants who complete the study (e.g., Atsawasuwan 2020, Swanson 2020, Arora 2019). Such an approach to improve retention rate for a clinical study is generally acceptable and is a common practice (HHS.gov 2019, U.S. FDA 2018). The sponsor also compensates providers for exerting efforts towards reducing the inconvenience of participants during the study and improving participants' understanding of the study and engagement in the study. In practice, we observe three ways to compensate providers: (i) a set dollar amount with a requirement to exert at least a specified effort level, (ii) a payment per unit of effort, and (iii) a payment per unit of effort with a requirement to exert at least a specified effort level (Ingram

2021c,a, Office of Institutional Compliance 2021). We referred to the first form of payment as the Fixed Compensation (FC) Contract, the second as the Linear Compensation (LC) Contract, and the third as the Conditional Linear Compensation (CLC) Contract.

To the best of our knowledge, there does not exist an economic analysis that evaluates the cost performance of compensation contracts for providers to improve participant retention. Hence, our primary research objective in this study is *to analyze the compensation contracts mentioned above to encourage providers to exert appropriate efforts to improve participant retention*.

To address our research objective, we consider a clinical study with three stakeholders: a sponsor, an investigator, and a coordinator. The sponsor's objective is to minimize the cost of retaining a targeted number of participants until the completion of the study. We first examine the sponsor's problem of deciding the providers' effort levels and monetary payments to the participants under the centralized model. Second, we characterize the optimal compensation contracts for the providers under the SI and OM models. Third, we compare the three compensation contracts (FC, LC, and CLC) observed in practice with the optimal compensation contract to understand their relative cost performances.

Our theoretical analysis shows that the optimal effort levels always increase with the target retention rate regardless of the team structure. However, the sponsor may not find it optimal to offer higher monetary payments to the participants to achieve a higher retention rate. Given the optimal contract, we show that the optimal expected retention cost under the SI (resp., OM) model is at most 20% (resp., 40%) higher than that of the centralized model. Further, our analysis of the compensation contracts observed in practice shows that it is sufficient for the sponsor to optimally choose between the FC and the LC contracts under the SI model to minimize the expected retention cost. However, under the OM model, there exist cases where the sponsor is better off adopting the CLC contract.

We also conduct a computational study using a realistic test bed to derive insights into the cost performances of various contracts. First, we observe that given the optimal compensation contracts, the sponsor does not suffer much financially under the SI and the OM models in practical instances—the optimal expected retention cost under the SI (resp., OM) model is at most 6% (resp., 8%) higher than that of the centralized model in our numerical study. Second, we compare the optimal expected retention costs under the FC, LC, and CLC contracts as the specified requirements on effort levels change for a given clinical study. We find that the CLC (resp., LC) contract tends to be more cost-effective on average if the specified requirements on effort levels are small under the OM (resp., SI) model. In contrast, the FC and LC contracts are more cost-effective for medium and high values of the specified requirements, respectively, under both models. Further, as the target retention rate increases or as the effort becomes more efficient in retaining participants, the regions where the CLC and the FC contracts are more cost-effective than the LC contract expand.

Third, we compare the optimal expected retention costs under the three contracts with the costs of the optimal compensation contracts under the SI and the OM models. Our results show that the increase in the sponsor's cost for the best of these three contracts is at most 6% (resp., 10%) on average relative to that for the optimal compensation contract under the SI (resp., OM) model. Further, this increase is at most 10% in 83% (resp., 62%) of the instances under the SI (resp., OM) model, suggesting these three contracts do not perform poorly despite their simplistic forms. We also explore the instances under the SI (resp., OM) model where this increase in cost exceeds 10% to provide guidance on when it is significantly beneficial for the sponsor to adopt the optimal compensation contract. For example, we find that the sponsor may benefit from adopting the optimal compensation contract when the effort is more effective, the monetary payments are less effective, and the specified requirements on effort levels are small. Finally, we examine the cost performances of the FC, LC, and CLC contracts when the effectiveness of effort and monetary payment is uncertain. We find that the LC contract becomes less favorable than the FC and CLC contracts when this uncertainty exists.

2. Literature Review

Our research contributes to the growing operations management literature on clinical studies. Recent research on clinical studies in operations management focuses on issues such as the optimal design in adaptive clinical studies (Alban et al. 2023, Anderer et al. 2022, Tian et al. 2021, Ahuja and Birge 2020), statistical testing (Bertsimas and Sturt 2020, Goh et al. 2018), scheduling of participants visits (Colvin and Maravelias 2010), and clinical trial supply chain management (Fleischhacker et al. 2015, Fleischhacker and Zhao 2011).

Participant recruitment and retention in clinical studies have received more attention in recent years. Kouvelis et al. (2017) provide an optimal schedule to open testing sites and an optimal rate for recruiting participants to maximize the net present value of a drug. In contrast to our work, they focus solely on participant recruitment without considering the impact of dropouts. Tian et al. (2021) extend the above study by considering participant dropouts and uncertain drug quality. Tian et al. (2023) develop an optimal participant enrollment policy for effectively conducting late-stage clinical trials. Song et al. (2023) analyze the economic performance of commonly observed incentive schemes in improving participant retention. Our research differs from the above studies by examining compensation contracts that motivate the providers (i.e., the investigator and the coordinator) to exert appropriate effort to improve participant retention. Further, both Tian et al. (2023, 2021) consider the dropout rate to be an exogenous random variable, whereas we consider that the participation retention rate depends on the efforts exerted by the providers and the monetary incentive offered by the sponsor.

Our research also relates to the literature on incentive schemes for healthcare providers. This literature focuses on analyzing the performance of various payment contracts (e.g., fee-for-service, bundled payment, and gainsharing contract) among healthcare providers to improve health outcomes (Rajagopalan and Tong 2022, Ghamat et al. 2021, Gupta et al. 2021, Adida and Bravo 2019, Andritsos and Tang 2018, Jiang et al. 2012). Further, our problem of analyzing compensation contracts shares the mathematical underpinnings with the problems of designing contracts, where a principal contracts with multiple agents whose decisions affect an output. The extensive research in this domain has either focused on administering incentive schemes to the agents under moral hazard (Guillen et al. 2015, Dragon et al. 1996, McAfee and McMillan 1991, Holmstrom 1982) or analyzing team contracts when there is adverse selection, e.g., see Mookherjee (2006) for a detailed review of the relevant research.

Similar to the above literature, we also analyze compensation contracts for clinical study providers. However, according to the IRB guidelines, the compensation contracts cannot be based on clinical study outcomes such as efficacy or adverse effect of the drug or the number of participants retained (e.g., Ingram 2021b, Partners HealthCare 2017). Thus, the outcome-based contracts studied in the above literature cannot be offered in clinical studies. Further, recall that participation in a clinical study is voluntary and participants can drop out without completing the study. Hence, unlike a typical healthcare setting where compensation contracts are for healthcare providers alone, a sponsor of a clinical study needs to consider incentivizing participants along with compensating providers. Moreover, as mentioned in the Introduction, different clinical study team structures exist in practice. These features result in a setting that requires developing models and analyzing contracts specific to clinical studies.

3. Model Setting

We consider a sponsor who employs an investigator and a coordinator to conduct, oversee, and coordinate a clinical study. The study enrolls N participants who may choose to drop out before completion. Similar to the setting in Song et al. (2023), we assume that the target retention rate is $\bar{\delta} \in [0,1]$. Thus, the sponsor must retain at least $\bar{\delta}N$ participants till the end of the study for completion[1]. As mentioned in the Introduction, the sponsor can offer monetary payment to participants who complete the study to improve retention (e.g., Atsawasuwan 2020, Swanson 2020, Retinal Consultants of Houston 2019). In addition, the investigator and the coordinator can exert effort to improve participant retention. The investigator's effort may include providing adequate resources (hiring and training staff, time, facilities) for the study, assuring easy access to staff, appropriate distribution of investigational agents (e.g., drugs), and ensuring participants' safety, etc. (Feehan and

[1] Song et al. (2023) provide a detailed explanation of how $\bar{\delta}$ is determined in practice.

Garcia-Diaz 2020, Baer et al. 2011). The coordinator's effort may include engaging and interacting with participants to understand their concerns, spending time to explain the consent form, study requirements, and time commitments to participants, sending thank-you notes to encourage continued participation, etc. (Genesis Research Services 2021, Chhatre et al. 2018, Institute of Medicine 2015).

Let $a, e \in [0,1]$, be the investigator's and coordinator's effort levels, respectively. Such an approach to model effort has been widely adopted in the OM literature (e.g., Bellos and Kavadias 2021, Adida and Bravo 2019, Hu et al. 2016, Corbett and DeCroix 2001). Let $f \geq 0$ be the per-participant monetary payment. Previous research has shown that providers' effort and the monetary payment to participants are effective in improving retention (Parkinson et al. 2019, Booker et al. 2011, National Research Council 2010). Thus, we assume that the retention rate δ achieved in the study is jointly affected by the providers' effort levels and the monetary payment. Further, collaborative efforts between different stakeholders are beneficial to retain participants (e.g., Fouad et al. 2014), which suggests a complementary relationship between the providers' effort levels. The Operations Management and Economics literature models effort complementarity using a supermodular function (e.g., Roels 2014, Milgrom and Roberts 1995). Therefore, we let *achieved* retention rate $\delta(a,e,f) = \lambda(a+e) + \lambda_J a\, e + \lambda_f f$ where $\lambda, \lambda_J, \lambda_f \in [0,1]$. The parameters λ, λ_J, λ_f reflect the effectiveness of the individual provider's effort, the joint effectiveness of providers' efforts, and the effectiveness of the monetary payment f on participant retention, respectively. We consider these parameters to be deterministic. However, we relax this assumption in Section 6. Typically, for a given clinical study, the sponsor can empirically derive the values for these parameters based on historical data from similar studies.

Let the effort cost be $\Theta_I(a;\theta_I)$ and $\Theta_C(e;\theta_C)$ for the investigator and the coordinator, respectively, where θ_I, θ_C are their respective effort cost parameters. We assume $\Theta_I(a;\theta_I)$ (resp., $\Theta_C(e;\theta_C)$) is an increasing convex function in a (resp., e). Specifically, let $\Theta_I(a;\theta_I) = \dfrac{1}{2}\theta_I a^2$ and $\Theta_C(e;\theta_C) = \dfrac{1}{2}\theta_C e^2$. This functional form is commonly adopted as exerting higher effort can increase the cost disproportionately (Hu et al. 2016, Lafontaine and Slade 1996, Hauser et al. 1994). We assume both the investigator and the coordinator can be of two types—high (H) and low (L). The high-type provider incurs a higher cost to exert the same amount of effort as compared to the low-type provider. That is, $\theta_I^H \geq \theta_I^L$ and $\theta_C^H \geq \theta_C^L$. Note that depending on the clinical study team structures, the sponsor may or may not know the providers' type. Figure 1 summarizes the three team structures we discussed in the Introduction. We next analyze these structures in detail. Table 1 summarizes the commonly used notation.

3.1. The Centralized (CM) Model

Here the sponsor, the investigator, and the coordinator belong to the same organization (see Figure 1), and the sponsor knows the providers' type. Hence, given type $k \in \{L, H\}$ of the investigator and

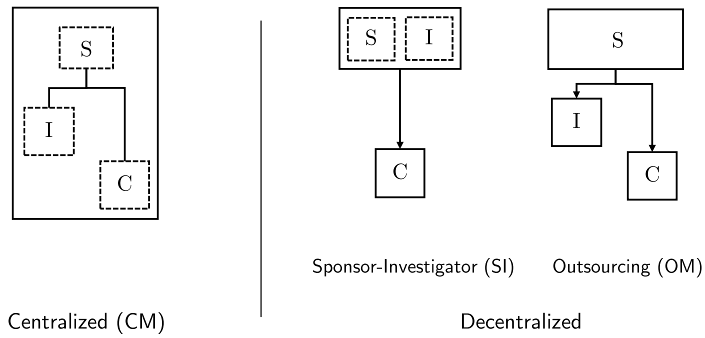


Figure 1 Clinical Study Team Structures

Notation	Definition
Parameters	
N	sample size
$\bar{\delta}$	target retention rate
θ_I^k	investigator's effort cost parameter, $k \in \{L, H\}$
θ_C^l	coordinator's effort cost parameter, $l \in \{L, H\}$
λ	effectiveness of provider's effort
λ_f	effectiveness of monetary payment
λ_J	joint effectiveness of providers' effort
p	probability that the investigator is high type
q	probability that the coordinator is high type
Functions	
$\delta(a, e, f)$	achieved retention rate
$\Theta_I(a; \theta_I^k)$	investigator's effort cost
$\Theta_C(e; \theta_C^l)$	coordinator's effort cost
Π	expected retention cost
Decision Variables	
a	investigator's effort level
e	coordinator's effort level
f	monetary payment to the participant
r	compensation to the investigator
s	compensation to the coordinator

Table 1 Notation

$l \in \{L, H\}$ of the coordinator, the sponsor determines the effort levels a_{kl}, e_{kl} for the providers and the monetary payment to the participants f_{kl} to minimize the retention cost while ensuring the achieved retention rate is at least $\bar{\delta}$ (Constraint (2)). The retention cost (Π_{kl}^{CM}) consists of the investigator's effort cost, the coordinator's effort cost, and monetary payment to each participant who completes the study. The sponsor's decision problem for a given type combination (k, l) is as follows:

$$\min_{a_{kl}, e_{kl} \in [0,1], f_{kl} \geq 0} \quad \Pi_{kl}^{CM} = \Theta_I(a_{kl}; \theta_I^k) + \Theta_C(e_{kl}; \theta_C^l) + \delta(a_{kl}, e_{kl}, f_{kl}) N f_{kl}, \tag{1}$$

$$\text{s.t.} \quad \delta(a_{kl}, e_{kl}, f_{kl}) \geq \bar{\delta}. \tag{2}$$

To understand the interaction between the providers' efforts and monetary payment to participants, we focus on the parameter settings where the optimal solution is interior, i.e., $a_{kl}^*, e_{kl}^* \in (0,1), f_{kl}^* > 0$, $k, l \in \{L, H\}$, where superscript $*$ represents the optimal value. Assumption 1 represents these parameter settings.

ASSUMPTION 1. *The effort cost parameters satisfy* $\min\{\theta_I^L, \theta_C^L\} > \frac{\max\{\bar{\delta}N(\lambda_J+\lambda), 2N(\bar{\delta}\lambda_J+\lambda^2)\}}{\lambda_f}$.

Next, we provide the sponsor's optimal decisions under the centralized model. Proofs of all technical results are available in Appendix EC.1.

PROPOSITION 1. *Given type* $k \in \{L, H\}$ *of the investigator and type* $l \in \{L, H\}$ *of the coordinator, the sponsor's optimal decisions are:* $a_{kl}^* = \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J+\lambda_f\theta_C^l)}{\lambda_f^2\theta_C^l\theta_I^k-\bar{\delta}^2N^2\lambda_J^2}$, $e_{kl}^* = \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J+\lambda_f\theta_I^k)}{\lambda_f^2\theta_C^l\theta_I^k-\bar{\delta}^2N^2\lambda_J^2}$, *and* $f_{kl}^* = \frac{\bar{\delta}-\lambda(a_{kl}^*+e_{kl}^*)-\lambda_J a_{kl}^* e_{kl}^*}{\lambda_f}$.

Proposition 1 helps us understand the trade-off between offering monetary payment to participants and exerting effort in achieving the target retention rate. It is straightforward to observe that higher the effectiveness of implementing effort (λ and/or λ_J), ceteris paribus, higher the optimal values of effort levels and lower the monetary payment. On the contrary, higher the effectiveness of monetary payment (λ_f), lower the optimal values of the effort levels and higher the optimal value of the monetary payment.

Note that when $\lambda_J = 0$, as the cost of implementing effort increases for one provider, it is intuitive that the sponsor decreases the corresponding effort level and increases the monetary payment to maintain the target retention rate, and the effort level of the other provider remains unchanged. However, when there exists a positive complementarity between the efforts (i.e., $\lambda_J > 0$), the first result in Corollary 1 below suggests that when one provider's effort cost parameter increases, the sponsor decreases the effort levels of both providers.

COROLLARY 1. *We have the following results:*

- ***Impact of*** θ_I^k ***and*** θ_C^l: *The optimal values* a_{kl}^*, e_{kl}^* *are decreasing and* f_{kl}^* *is increasing in* θ_I^k *and* $\theta_C^l, k, l \in \{L, H\}$.
- ***Impact of*** $\bar{\delta}$: *The optimal effort* a_{kl}^*, e_{kl}^* *are increasing in* $\bar{\delta}$. *The optimal monetary payment* f_{kl}^* *is increasing at* $\bar{\delta} = 0$ *and unimodal in* $\bar{\delta} \in [0, 1]$.

Corollary 1 offers a guideline to the sponsor for allocating his funds towards increasing effort or incentivizing participants. Specifically, the first result implies that when it is expensive for the sponsor to require higher effort from a provider to influence participant retention, he should provide a higher monetary payment to participants for achieving the target retention rate rather than requiring higher effort from the other provider. The second result suggests that as the target retention rate ($\bar{\delta}$) increases, the sponsor should require higher optimal efforts from the providers. However, the sponsor may not need to increase the optimal monetary payment to a participant to achieve higher $\bar{\delta}$. When $\bar{\delta}$ is low, the sponsor can easily increase the monetary payment without significantly increasing the overall cost as the number of participants who receive payments (i.e., $\bar{\delta}N$) is low. However, when $\bar{\delta}$ is high, $\bar{\delta}N$ also becomes high. Hence, increasing the payment to achieve high $\bar{\delta}$ may result in a higher overall cost. In such situations, the sponsor may consider focusing on increasing providers' efforts and reducing the monetary payment to a participant.

Next, we describe the details of decision-making for participant retention under the sponsor-investigator model.

3.2. The Sponsor-Investigator (SI) Model

Under this model, a sponsor also acts as an investigator; hence, he initiates and conducts the clinical study. The coordinator is an external entity. Thus, the sponsor knows the type of investigator but does not know the type of coordinator. Instead, he knows that with probability $q \in [0,1]$ the coordinator is high type (θ_C^H) and with probability $\bar{q} = 1 - q$ the coordinator is low type (θ_C^L), where $0 < \theta_C^L \leq \theta_C^H$.

The sponsor designs a compensation contract specifying the coordinator's payment based on his effort level. Using the revelation principle, we restrict our attention to a direct mechanism consisting of at most two contracts—one for each coordinator type (Myerson 1981). In particular, given type k of the investigator, the sponsor offers a menu of effort-payment pairs $\{(e_{kH}, s_{kH}), (e_{kL}, s_{kL})\}$ to the coordinator. The coordinator decides whether to accept and, if so, which contract to select. The sponsor then determines the effort level a_{kl} for the investigator and the monetary payment f_{kl} for the participants.

Given type k of the investigator, the sponsor's problem, then, is

$$\begin{aligned}
\min_{a_{kl}, e_{kl} \in [0,1]; f_{kl}, s_{kl} \geq 0} \quad & \Pi_k^{SI} = \mathbb{E}_l[\Theta_I(a_{kl}; \theta_I^k) + s_{kl} + f_{kl} N \delta(a_{kl}, e_{kl}, f_{kl})], && && (3)\\
\text{s.t.} \quad & \delta(a_{kl}, e_{kl}, f_{kl}) \geq \bar{\delta}, && \forall l \in \{L, H\}, && (4)\\
& s_{kl} - \Theta_C(e_{kl}; \theta_C^l) \geq s_{k(-l)} - \Theta_C(e_{k(-l)}; \theta_C^l), && \forall l \in \{L, H\}, && (5)\\
& s_{kl} - \Theta_C(e_{kl}; \theta_C^l) \geq 0, && \forall l \in \{L, H\}. && (6)
\end{aligned}$$

The objective function (3) captures the expected retention cost, which consists of (i) the investigator's effort cost, (ii) the compensation to the coordinator, and (iii) the monetary payment to the participants who complete the study. Constraint (4) ensures that the achieved retention rate is at least $\bar{\delta}$. Constraints (5) are the incentive compatibility (IC) constraints that guarantee truth-telling from the coordinator. That is, the sponsor designs the effort levels and corresponding compensation $(e_{kl}, s_{kl}), l \in \{L, H\}$ such that the coordinator's monetary benefit, $[s_{kl} - \Theta_C(e_{kl}; \theta_C^l)]$, is maximized by revealing his true type. Constraints (6) are the individual rationality (IR) constraints that ensure the coordinator receives non-negative benefits from the study. To simplify expressions for optimal solutions, we consider $\hat{\theta}_C^H = \theta_C^H + \frac{\bar{q}}{q}(\theta_C^H - \theta_C^L)$. We summarize the optimal values of the sponsor's decision variables in the proposition below:

PROPOSITION 2. *Given an investigator's type $k \in \{L, H\}$, the optimal solution to the SI model is as follows:*

1. $a_{kL}^{\circ} = a_{kL}^{*}$, $a_{kH}^{\circ} = \frac{\bar{\delta}\lambda N(\bar{\delta}N\lambda_J + \lambda_f \hat{\theta}_C^H)}{(\lambda_f^2 \theta_I^k \hat{\theta}_C^H - \bar{\delta}^2 N^2 \lambda_J^2)}$, *where a_{kL}^{*} is the centralized optimal solution.*
2. $e_{kL}^{\circ} = e_{kL}^{*}$, $e_{kH}^{\circ} = \frac{\bar{\delta}\lambda N(\bar{\delta}N\lambda_J + \lambda_f \theta_I^k)}{(\lambda_f^2 \theta_I^k \hat{\theta}_C^H - \bar{\delta}^2 N^2 \lambda_J^2)}$, *where e_{kL}^{*} is the centralized optimal solution.*
3. $s_{kL}^{\circ} = \Theta_C(e_{kL}^{\circ}; \theta_C^L) + \Theta_C(e_{kH}^{\circ}; \theta_C^H) - \Theta_C(e_{kH}^{\circ}; \theta_C^L), s_{kH}^{\circ} = \Theta_C(e_{kH}^{\circ}; \theta_C^H)$.
4. $f_{kl}^{\circ} = \frac{\bar{\delta} - \lambda(a_{kl}^{\circ} + e_{kl}^{\circ}) - \lambda_J a_{kl}^{\circ} e_{kl}^{\circ}}{\lambda_f}$, $l \in \{L, H\}$.

In the optimal solution to the sponsor's problem, the IR constraint for the high-type coordinator is tight. The sponsor's cost is minimized when the IC constraint is tight for the low-type coordinator. Together, these characteristics of the optimal solution imply that the low-type coordinator receives an information rent. Further, note that the information rent $\Theta_C(e_{kH}; \theta_C^H) - \Theta_C(e_{kH}; \theta_C^L)$ for the low-type coordinator does not depend on his effort level. Consequently, when the coordinator's type is low, the optimal effort levels for the providers under the SI model are the same as those under the centralized model. The following result summarizes the impact of model parameters on the optimal effort levels and the optimal monetary payment to a participant.

COROLLARY 2. *Given an investigator's type $k \in \{L, H\}$, when the coordinator is high type, the providers' optimal effort levels (a_{kH}°, e_{kH}°) are decreasing in θ_C^H while increasing in θ_C^L, q and $\bar{\delta}$. The optimal monetary payment f_{kH}° is (i) increasing in θ_C^H while decreasing in θ_C^L and q, (ii) increasing at $\bar{\delta} = 0$ and unimodal in $\bar{\delta} \in [0, 1]$. When the coordinator is low type, the findings are the same as in Corollary 1.*

As θ_C^H increases, naturally, the effort level e_{kH}° decreases, and a_{kH}° reduces accordingly from the positive complementarity. Therefore, f_{kH}° increases. It is intuitive that an increase in θ_C^L results in a decrease in e_{kL}°. However, an increase in e_{kH}° as θ_C^L increases is not straightforward. The reduction in the gap $\theta_C^H - \theta_C^L$ due to an increase in θ_C^L allows the sponsor to offer smaller information rent which in turn helps increase e_{kH}°. Then, using the similar argument above, the insights regarding a_{kH}° and f_{kH}° follow immediately.

Further, as q increases, the sponsor's chance to contract with a high-type coordinator increases and hence, reducing the need for an information rent, which helps increase e_{kH}°. Hence, the directional changes in the optimal effort levels and optimal monetary payments to participants immediately follow from a similar argument as above. Finally, the impact of $\bar{\delta}$ on the optimal effort levels and the monetary payment are similar to that under Corollary 1. We next summarize the relationship between the optimal effort levels of the coordinator and the investigator, which follows immediately from the relative magnitudes of their effort cost parameters.

COROLLARY 3. *Given an investigator's type $k \in \{L, H\}$, the optimal effort (i) $e^{\circ}_{kL} \geq a^{\circ}_{kL}$ if and only if $\theta^k_I \geq \theta^L_C$ and (ii) $e^{\circ}_{kH} \geq a^{\circ}_{kH}$ if and only if $\theta^k_I \geq \hat{\theta}^H_C$.*

We next analyze the sponsor's problem under the OM model.

3.3. The Outsourcing (OM) Model

In this section, we analyze the OM model, where both the investigator and the coordinator are external entities appointed by the sponsor. Thus, the sponsor does not know the types of the investigator and the coordinator but knows the distribution of their types. Similar to the SI model, we assume that the coordinator is the high type with probability $q \in [0,1]$ and low type with probability $\bar{q} = 1 - q$. Further, the investigator is the high type with probability $p \in [0,1]$ and low type with probability $\bar{p} = 1 - p$. We further assume that the providers' type distributions are independent. Using the revelation principle, the sponsor designs a mechanism that asks the providers to reveal their types. Given the announced types, the sponsor offers the effort levels and the corresponding compensations for the investigator and the coordinator.

The sponsor also decides the monetary payment to participants given each type combination of the investigator and coordinator. In particular, the sponsor solves the following problem such that for a provider, revealing their true type is the optimal strategy (regardless of the type of the other provider).

$$\begin{aligned}
\min_{a_{kl}, e_{kl}, r_{kl}, s_{kl}} \quad & \Pi^{OM} = \mathbb{E}_{k,l}\left[r_{kl} + s_{kl} + f_{kl} N \delta(a_{kl}, e_{kl}, f_{kl})\right], && && (7)\\
\text{s.t.} \quad & \delta(a_{kl}, e_{kl}, f_{kl}) \geq \bar{\delta}, && \forall k, l \in \{L, H\}, && (8)\\
& r_{kl} - \Theta_I(a_{kl}; \theta^k_I) \geq r_{-kl} - \Theta_I(a_{-kl}; \theta^k_I), && \forall k, l \in \{L, H\}, && (9)\\
& r_{kl} - \Theta_I(a_{kl}; \theta^k_I) \geq 0, && \forall k, l \in \{L, H\}, && (10)\\
& s_{kl} - \Theta_C(e_{kl}; \theta^l_C) \geq s_{k(-l)} - \Theta_C(e_{k(-l)}; \theta^l_C), && \forall k, l \in \{L, H\}, && (11)\\
& s_{kl} - \Theta_C(e_{kl}; \theta^l_C) \geq 0, && \forall k, l \in \{L, H\}, && (12)\\
& a_{kl}, e_{kl} \in [0,1], r_{kl}, s_{kl} \geq 0, && \forall k, l \in \{L, H\}. && (13)
\end{aligned}$$

The objective function (7) describes the expected retention cost. Constraints (8) ensures that the target retention rate is at least $\bar{\delta}$. Constraints (9) and (11) are the IC constraints for the investigator and the coordinator, respectively, guaranteeing that the providers reveal their true types. Constraints (10) and (12) are IR constraints for the investigator and the coordinator.

In the optimal solution to the sponsor's problem presented in Proposition 3, IR constraints for high-type providers are tight and hence, they receive zero benefits. The sponsor's cost is minimized when IC constraints are tight for the low-type providers. This implies that the low-type providers receive information rents. We use superscript $\Diamond$ to denote the optimal values of effort levels, compensation amounts, and monetary payments to the participant under the OM model. Consider $\hat{\theta}_I^L = \theta_I^L, \hat{\theta}_I^H = \theta_I^H + \frac{\bar{p}}{p}(\theta_I^H - \theta_I^L), \hat{\theta}_C^L = \theta_C^L$, and $\hat{\theta}_C^H = \theta_C^H + \frac{\bar{q}}{q}(\theta_C^H - \theta_C^L)$.

Proposition 3. *Given type $k \in \{L,H\}$ of the investigator and type $l \in \{L,H\}$ of the coordinator, the optimal compensation to the providers and optimal the monetary payment to the participants are as follows:*

1. $a_{kl}^{\Diamond} = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J + \lambda_f \hat{\theta}_C^l)}{\lambda_f^2 \hat{\theta}_I^k \hat{\theta}_C^l - \bar{\delta}^2 N^2 \lambda_J^2}$,
2. $e_{kl}^{\Diamond} = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J + \lambda_f \hat{\theta}_I^k)}{\lambda_f^2 \hat{\theta}_I^k \hat{\theta}_C^l - \bar{\delta}^2 N^2 \lambda_J^2}$,
3. $s_{kH}^{\Diamond} = \Theta_C(e_{kH}^{\Diamond}; \theta_C^H)$, $s_{kL}^{\Diamond} = \Theta_C(e_{kL}^{\Diamond}; \theta_C^L) + \Theta_C(e_{kH}^{\Diamond}; \theta_C^H) - \Theta_C(e_{kH}^{\Diamond}; \theta_C^L)$,
4. $r_{Hl}^{\Diamond} = \Theta_I(a_{Hl}^{\Diamond}; \theta_I^H)$, $r_{Ll}^{\Diamond} = \Theta_I(a_{Ll}^{\Diamond}; \theta_I^L) + \Theta_I(a_{Hl}^{\Diamond}; \theta_I^H) - \Theta_I(a_{Hl}^{\Diamond}; \theta_I^L)$,
5. $f_{kl}^{\Diamond} = \frac{\bar{\delta} - \lambda(a_{kl}^{\Diamond} + e_{kl}^{\Diamond}) - \lambda_J a_{kl}^{\Diamond} e_{kl}^{\Diamond}}{\lambda_f}$.

Proposition 3 implies that the optimal effort levels of the providers are lower than those under the centralized model. This result follows from a similar explanation as provided in Section 3.2. We next summarize the behavior of optimal effort levels with respect to various model parameters they depend on. The explanation of the results is similar to those provided for Corollaries 1 and 2 and hence, avoided for brevity.

Corollary 4. *Under the OM model, the providers' optimal effort levels (i) $a_{kl}^{\Diamond}, e_{kl}^{\Diamond}$ are decreasing in θ_I^k and θ_C^l while increasing in $\bar{\delta}$, (ii) $a_{Hl}^{\Diamond}, e_{Hl}^{\Diamond}$ are increasing in θ_I^L and p, (iii) $a_{kH}^{\Diamond}, e_{kH}^{\Diamond}$ are increasing in θ_C^L and q. The sponsor's optimal monetary payment (i) $f_{kl}^{\Diamond}$ is increasing in θ_I^k and θ_C^l, (ii) $f_{Hl}^{\Diamond}$ is decreasing in θ_I^L and p, (iii) $f_{kH}^{\Diamond}$ is decreasing in θ_C^L and q, and (iv) $f_{kl}^{\Diamond}$ is increasing at $\bar{\delta} = 0$ and unimodal in $\bar{\delta} \in [0,1]$, for $k,l \in \{L,H\}$.*

3.4. A Comparative Analysis

In this section, we analyze the optimal contracts under the SI and OM models relative to the CM model. Specifically, we aim to understand the impact of these three models on the sponsor's cost of retaining the target number of participants, the providers' compensations, the participants' payments, and the providers' efforts to retain participants. We first comment on how the expected optimal effort levels of providers change with the three models. The proposition below establishes that the providers' expected optimal effort levels are the largest when the clinical study team structure is centralized. The result follows from the fact that information rent increases due to decentralization, and hence, the sponsor desires smaller effort levels.

Proposition 4. *The provider's expected optimal effort levels are such that $\mathbb{E}_{k,l}\left(a_{kl}^{\Diamond}\right) \le \mathbb{E}_{k,l}\left(a_{kl}^{\circ}\right) \le \mathbb{E}_{k,l}\left(a_{kl}^{*}\right)$ and $\mathbb{E}_{k,l}\left(e_{kl}^{\Diamond}\right) \le \mathbb{E}_{k,l}\left(e_{kl}^{\circ}\right) \le \mathbb{E}_{k,l}\left(e_{kl}^{*}\right)$; $k,l \in \{L,H\}$.*

We next analyze the sponsor's expected retention costs, expected total compensation to the providers, and monetary payment to participants under the three models.

Define $\Pi_1 = E_{k,l}[\Pi_{kl}^{CM*}], \Pi_2 = E_k[\Pi_k^{SI\circ}], \Pi_3 = \Pi^{OM\Diamond}$ for comparison purposes. Let $P_1 = \mathbb{E}_{k,l}\left[\Theta_I(a_{kl}^*;\theta_I^k)+\Theta_C(e_{kl}^*;\theta_C^k)\right], P_2 = \mathbb{E}_{k,l}\left[\Theta_I(a_{kl}^\circ;\theta_I^k)+s_{kl}^\circ\right], P_3 = \mathbb{E}_{k,l}\left[r_{kl}^\Diamond+s_{kl}^\Diamond\right]$, represent the expected optimal total compensation to the providers under the CM, SI, and OM models, respectively. Let $F_m = \Pi_m - P_m$ be the expected optimal monetary payments to the participants who complete the study, $m = 1,2,3$, under the CM, SI, and OM models, respectively.

Proposition 5. *We have* $\Pi_3 \geq \Pi_2 \geq \Pi_1$, $P_3 \leq P_2 \leq P_1$, *and* $F_3 \geq F_2 \geq F_1$.

The first result above reveals that the expected retention cost is lowest when the clinical study team is centralized, which is intuitive. Note that under the OM model, the total information rent is highest, whereas Proposition 4 implies that the providers' efforts and, therefore, the costs of their efforts are lowest. The second result in the proposition above suggests that the negative impact of increased information rents on the sponsors' cost is lower than the positive impact of decreased effort costs. Hence, the total compensation to the providers is the lowest under the OM model. This result, together with the first result, implies that the monetary payment to the participants is highest under the OM model. Our next proposition further analyzes the difference between the expected retention cost under the three models.

Proposition 6. *We have the following results:*

- ***Impact of*** θ_I^L ***and*** θ_C^L: $(\Pi_m - \Pi_1)$ *is decreasing in* θ_I^L *and* θ_C^L, $m = 2,3$.
- ***Impact of*** θ_C^H: *Under the SI model,* $(\Pi_2 - \Pi_1)$ *is (i) increasing in* θ_C^H *for* $\theta_C^H \in (\theta_C^L, \theta_1)$, *and (ii) decreasing in* θ_C^H *for* $\theta_C^H > \theta_2$ *where thresholds* $\theta_C^L < \theta_1 < \theta_2$. *Similar observations hold under the OM model for* $q > \hat{q}$, *where* $\hat{q} \in (0,1)$.
- ***Impact of*** θ_I^H: *Under the SI model,* $(\Pi_2 - \Pi_1)$ *is decreasing in* θ_I^H. *Under the OM model, for* $p > \hat{p} \in (0,1)$, $(\Pi_3 - \Pi_1)$ *is (i) increasing in* θ_I^H *for* $\theta_I^H \in (\theta_I^L, \underline{\theta})$, *and (ii) decreasing in* θ_I^H *for* $\theta_I^H \geq \bar{\theta}$ *where* $\theta_I^L < \underline{\theta} < \bar{\theta}$.
- ***Impact of*** λ, λ_J ***and*** λ_f: $(\Pi_m - \Pi_1)$ *is increasing in* λ, λ_J *while decreasing in* λ_f, $m = 2,3$.
- ***Impact of*** δ: $(\Pi_m - \Pi_1)$ *is increasing in* $\bar{\delta}$, $m = 2,3$.

Proposition 6 examines how different parameters affect the increase in the expected retention cost from adopting either the SI or the OM model instead of the CM model. First, as the effort cost parameter increases for the low-type providers (θ_I^L and θ_C^L), the high- and low-type providers become similar, and hence, the difference between the expected retention cost under the SI (resp., the OM) model and that under the CM model reduces. However, the impact of increasing θ_C^H is two-fold. On the one hand, information rent to the low-type provider increases under the SI/OM model. On the other hand, implementing efforts becomes less desirable as compared to providing monetary payment to the participants. Therefore, $(\Pi_m - \Pi_1)$, $m = 2,3$, first increases with θ_C^H due to a larger information rent, while decreases when the sponsor shifts to the monetary payment.

Under the OM model, the impact of an increase in θ_I^H is similar to that of θ_C^H following a similar explanation as above. However, under the SI model, the sponsor knows the investigator's type, and thus, the investigator does not receive any information rent. Further, as θ_I^H increases, the sponsor prefers incentivizing participants as compared to compensating providers for higher efforts. Hence, the information rent of the coordinator decreases under the SI model, and $(\Pi_2 - \Pi_1)$ decreases. Finally, when the providers' effort becomes more effective in retaining participants (i.e., λ, λ_J increases or λ_f decreases) or the target retention rate ($\bar{\delta}$) increases, the sponsor desires providers to exert higher efforts, which in turn, increases the information rent under the SI/OM model. Hence, $(\Pi_m - \Pi_1)$, $m = 2, 3$, increases in $\lambda, \lambda_J, \bar{\delta}$ while decreases in λ_f. We conclude this section by providing upper bounds on the relative *performances* (i.e., the sponsor's expected retention costs) of the SI and the OM models.

Proposition 7. *$\frac{\Pi_2}{\Pi_1} \leq 1.2$ and $\frac{\Pi_3}{\Pi_1} \leq 1.4$.*

In practice, the performances of the two models can be better than the theoretical bounds established above. For example, using our test bed in Section 5.1, we observe that in our practical instances maximum value of $\frac{\Pi_2}{\Pi_1}$ (resp.,$\frac{\Pi_3}{\Pi_1}$) is 1.06 (resp., 1.08). These observations suggest that by utilizing optimal compensation contracts under a decentralized structure, clinical studies' sponsors can enjoy the resources and expertise of the service providers to improve participant retention without incurring much financial penalty.

Note that the optimal compensation contracts for providers are menu contracts, whereas, as mentioned in the Introduction, we observe the usage of contracts based on fixed or linear payment (regardless of the types) for providers' efforts in practice. Hence, a natural question arises regarding the performances of these contracts relative to optimal contracts for the decentralized structures. To address this question, we first theoretically analyze them and then comment on their performances in the next section.

4. Compensation Contracts in Practice

We consider three contracts mentioned in the Introduction: (1) the fixed compensation (FC) contract, (2) the linear compensation (LC) contract, and (3) the conditional linear compensation (CLC) contract. We discuss the design and implementation of each contract and derive the sponsor's optimal decisions under the SI and OM models. The detailed formulations of the sponsor's decision problems are in Appendix EC.2.

4.1. The Fixed Compensation Contract

Under this contract, the sponsor offers fixed compensations to the investigator and the coordinator and requires them to exert at least a specified effort level to ensure that the providers deliver adequate services to facilitate the study. In practice, the sponsor may determine the specified effort level based on historical data from similar studies and his experience with providers. Thus, under the SI model, for given type k of the investigator, the sponsor decides the investigator's effort a_k, monetary payment to the participant f_k, fixed compensation s_k for the coordinator, and requires the coordinator to exert at least an effort level $\underline{e}_k$, $k \in \{L, H\}$. Under the OM model, the sponsor decides the monetary payment to the participant f, fixed compensations r to the investigator, and s to the coordinator and requires the investigator (resp., coordinator) to exert at least an effort level $\underline{a}$ (resp., $\underline{e}$). Henceforth, we referred to these specified effort levels as "lower bounds" on effort. The following proposition characterizes the optimal solutions under the SI and OM models. We use superscript $\circ$ (resp., $\Diamond$) to represent optimal values for the SI (resp., OM) model.

Proposition 8. *When the sponsor adopts the FC contract,*

- *Under the SI model, given the investigator's type k, the sponsor's optimal decisions are: $s_k^\circ = \frac{1}{2}\theta_C^H \underline{e}_k^2$, $a_k^\circ = \frac{\bar{\delta}N(\lambda+\lambda_J \underline{e}_k)}{\lambda_f \theta_I^k}$, $f_k^\circ = \frac{(\bar{\delta}-\lambda \underline{e}_k)\lambda_f \theta_I^k - \bar{\delta}N(\lambda+\lambda_J \underline{e}_k)^2}{\lambda_f^2 \theta_I^k}$, $k \in \{L, H\}$. The coordinator's optimal effort is: $e_k^\circ = \underline{e}_k$ regardless of his type.*
- *Under the OM model, the sponsor's optimal decisions are: $r^\Diamond = \frac{1}{2}\theta_I^H \underline{a}^2$, $s^\Diamond = \frac{1}{2}\theta_C^H \underline{e}^2$, and $f^\Diamond = \frac{\bar{\delta}-\lambda(\underline{a}+\underline{e})-\lambda_J \underline{a}\underline{e}}{\lambda_f}$. The providers' optimal efforts, regardless of their types, are $a^\Diamond = \underline{a}$, $e^\Diamond = \underline{e}$.*

Proposition 8 shows that under the SI model, the coordinator exerts an effort level equal to the lower bound. This result follows from the nature of the FC contract. Thus, the sponsor minimizes his expected retention cost by setting the compensation as the high-type coordinator's effort cost (which guarantees the participation of both types) while choosing the investigator's effort a_k and the monetary payment f_k to ensure $\bar{\delta}N$ number of participants complete the study. The optimal solution under the OM model follows from similar arguments.

Let Π_{2k}^F, Π_3^F denote the sponsor's optimal expected retention costs for the SI (for given k) and OM models using the FC contract. The following result characterizes the performance of the FC contract relative to the optimal contracts in Sections 3.2 and 3.3.

Corollary 5.
- *For a given k under the SI model, we have*

$$\frac{\Pi_{2k}^F}{\Pi_k^{SI^\circ}} \leq 1 + \frac{2}{\theta_C^L {e_{LL}^\circ}^2}\mathbb{E}_l\left[\left(\theta_C^l - \frac{\bar{\delta}^2 N^2 \lambda_J^2}{\lambda_f^2 \theta_I^k}\right)(e_{kl}^\circ - \underline{e}_k)(e_{kl}^* - \underline{e}_k)\right] + \bar{q}\left(\frac{\theta_C^H}{\theta_C^L} - 1\right)\frac{(\underline{e}_k^2 - {e_{kH}^\circ}^2)}{{e_{LL}^\circ}^2}. \quad (14)$$

- *Under the OM model,*

$$\frac{\Pi_3^F}{\Pi^{OM^\Diamond}} \leq 1 + \frac{2}{\theta_I^L {a_{LL}^\Diamond}^2}\mathbb{E}_{k,l}\Big[(\underline{a} - a_{kl}^\Diamond)(\underline{a}\theta_I^k - \frac{\bar{\delta}N(\lambda+\lambda_J \underline{e})}{\lambda_f}) + \frac{1}{2}(\theta_I^H - \theta_I^k)(\underline{a}^2 - {a_{Hl}^\Diamond}^2)\Big] + \frac{2}{\theta_C^L {e_{LL}^\Diamond}^2}\mathbb{E}_{k,l}\Big[(\underline{e} - e_{kl}^\Diamond)(\underline{e}\theta_C^l - \frac{\bar{\delta}N(\lambda+\lambda_J \underline{a})}{\lambda_f}) + \frac{1}{2}(\theta_C^H - \theta_C^l)(\underline{e}^2 - {e_{kH}^\Diamond}^2)\Big]. \quad (15)$$

where e_{kl}^*, e_{kl}°, and $a_{kl}^\Diamond, e_{kl}^\Diamond$ are the optimal effort values under the centralized model (Section 3.1), the SI model (Section 3.2), and the OM model (Section 3.3), respectively, $k, l \in \{L, H\}$.

The above result implies that the performance of the FC contract depends on the lower bounds on efforts. We explain this further via an example below:

Example 1. Consider a case where both providers have a single type, i.e., $\theta_C^H = \theta_C^L = \theta_C$ and $\theta_I^H = \theta_I^L = \theta_I$. Under the SI model, we can rewrite the LHS in (14) for a single type by dropping the superscripts and the subscripts for types as follows:

$$\frac{\Pi_2^F}{\Pi^{SI^\circ}} \leq 1 + \frac{2}{\theta_C e^{\circ 2}} \left[\left(\theta_C - \frac{\bar{\delta}^2 N^2 \lambda_J^2}{\lambda_f^2 \theta_I} \right) (e^\circ - \underline{e})(e^* - \underline{e}) \right] = 1 + \frac{2}{\theta_C} \left[\left(\theta_C - \frac{\bar{\delta}^2 N^2 \lambda_J^2}{\lambda_f^2 \theta_I} \right) \left(\frac{e^* - \underline{e}}{e^*} \right)^2 \right].$$

Last equality above follows from Propositions 1-2 as $e^\circ = e^*$. Thus, if the lower bound $\underline{e} = e^\circ = e^*$, the FC contract can achieve the target retention rate at the optimal retention cost Π^{SI°. A similar observation holds for the OM model. Further, if the lower bounds for the SI and the OM models with the FC contracts are strictly lower than the respective optimal efforts (as specified in Sections 3.2 and 3.3), then $\frac{\Pi_2^F}{\Pi^{SI^\circ}} \leq 3$ and $\frac{\Pi_3^F}{\Pi^{OM^\diamond}} \leq 5$; otherwise, these bounds can be arbitrarily bad. □

We analyze the general case under the FC contract in Section 5. Given the FC contract, the providers (the coordinator under the SI model and both providers under the OM model) exert effort levels equal to the respective lower bounds regardless of their types. Next, we consider the LC contract, which allows the providers to exert effort based on their types.

4.2. The Linear Compensation Contract

In contrast to the FC contract, the LC contract compensates the providers based on their efforts. We denote the linear compensation per unit of effort to the coordinator (resp., investigator) as $\beta \geq 0$ (resp., $\nu \geq 0$). We further add the subscript k to the coordinator's compensation for the SI model to represent its dependency on the investigator's type k.

Under the SI model, given the investigator's type k, the sponsor decides the investigator's effort a_k, monetary payment to the participants f_k, and the per unit of effort compensation β_k. Given β_k, the coordinator decides his effort e_k. Under the OM model, the sponsor decides the per unit of effort compensations β and ν, and the monetary payment f. The providers decide their efforts a and e. The following results characterize the optimal solutions under the LC contracts for the SI and the OM models.

Proposition 9. *When the sponsor adopts the LC contract,*

- *Under the SI model, given the investigator's type k, the sponsor's optimal decisions are:* $f_k^\circ = \frac{\bar{\delta} - \lambda(a_k^\circ + e_{kH}^\circ) - \lambda_J a_k^\circ e_{kH}^\circ}{\lambda_f}$ *and* β_k°, a_k° *are the optimal solutions to* Problem P_{LSI}. *The optimal effort of type* l *coordinator is* $e_{kl}^\circ = \frac{\beta_k^\circ}{\theta_C^l}$, $l \in \{L, H\}$.
- *Under the OM model, the sponsor's optimal decisions are:* $f^\diamond = \frac{\bar{\delta} - \lambda(a_H^\diamond + e_H^\diamond) - \lambda_J a_H^\diamond e_H^\diamond}{\lambda_f}$ *and* $\nu^\diamond, \beta^\diamond$ *are the optimal solution to* Problem P_{LOM}. *The optimal efforts of providers are:* $a_k^\diamond = \frac{\nu^\diamond}{\theta_I^k}, e_l^\diamond = \frac{\beta^\diamond}{\theta_C^l}, k, l \in \{L, H\}$.

See Appendix EC.2.3.1 for the detailed expressions of Problems P_{LSI} and P_{LOM}.

Let Π_{2k}^N, Π_3^N denote the sponsor's optimal expected retention costs for the SI (for given k) and OM models using the LC contract. When there are multiple types of providers, it is notationally complex to present a parametric performance bound for the LC contract similar to the one obtained for the FC contract. Hence, we resort to computational study to explain its performance in general. However, in Corollary 6, we focus on a setting where both providers have a single type to analytically understand the performance of the LC contract relative to the optimal contracts in Sections 3.2 and 3.3. We drop the type index for this result.

Corollary 6. *When both providers have only a single type, we have* $\frac{\Pi_2^N}{\Pi^{SI^\circ}} < \frac{3}{2}, \frac{\Pi_3^N}{\Pi^{OM^\diamondsuit}} < 2.$

4.3. The Conditional Linear Compensation Contract

In this section, we consider a *conditional linear compensation* (CLC) contract, which extends the LC contract discussed above by imposing a lower bound on the effort level. In particular, the CLC contract compensates the coordinator (resp., investigator) β (resp., ν) amount per unit of effort and requires to exert an effort level more than or equal to a lower bound $\underline{e}$ (resp., $\underline{a}$).

Proposition 10. *When the sponsor adopts the CLC contract,*

- *Under the SI model, given the investigator's type k, the sponsor's optimal decisions are:* $f_k^\circ = \frac{\bar{\delta}-\lambda(a_k^\circ+e_{kH}^\circ)-\lambda_J a_k^\circ e_{kH}^\circ}{\lambda_f}$ *and* β_k°, a_k° *are the optimal solutions to* Problem P_{CLSI}. *The optimal effort of type l coordinator is* $e_{kl}^\circ = \max\{\underline{e}_k, \frac{\beta_k^\circ}{\theta_C^l}\}$, $l \in \{L, H\}$.
- *Under the OM model, the sponsor's optimal decisions are:* $f^\diamondsuit = \frac{\bar{\delta}-\lambda(a_H^\diamondsuit+e_H^\diamondsuit)-\lambda_J a_H^\diamondsuit e_H^\diamondsuit}{\lambda_f}$ *and* $\nu^\diamondsuit, \beta^\diamondsuit$ *are the optimal solution to* Problem P_{CLOM}. *The optimal efforts of providers are:* $a_k^\diamondsuit = \max\{\underline{a}, \frac{\nu^\diamondsuit}{\theta_I^k}\}, e_l^\diamondsuit = \max\{\underline{e}, \frac{\beta^\circ}{\theta_C^l}\}, k, l \in \{L, H\}$.

The detailed expressions of Problem P_{CLSI} and Problem P_{CLOM} are available in Appendix EC.2.4.1. When both providers have a single type, it is straightforward to verify that $\beta = \frac{\underline{e}\,\theta_C}{2}$ is a feasible solution to the sponsor's problem under the CLC contract. Further, the CLC contract with this value of β reduces to an FC contract with the fixed payment $\beta\underline{e}$. Therefore, the sponsor's optimal expected retention cost when adopting the CLC contract is at most that of this FC contract. Hence, the performance guarantee for the FC contract obtained in Example 1 also applies to the CLC contract. In the following section, we analyze the relative performance of the three contracts for general settings.

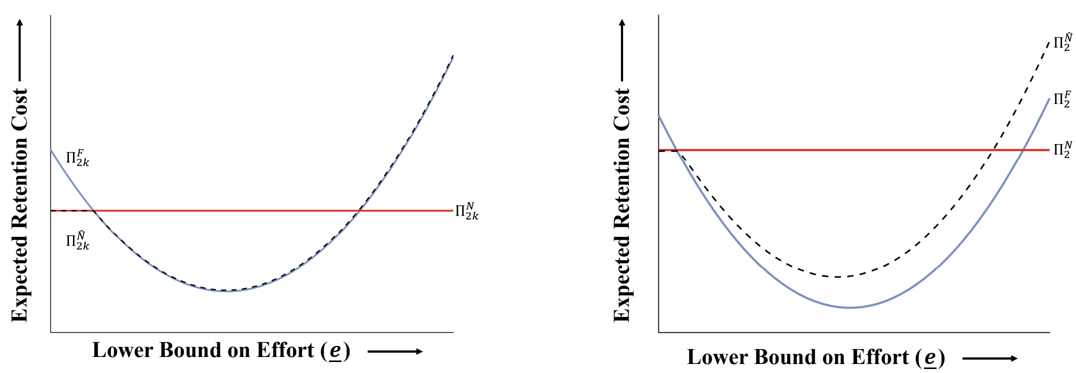
4.4. A Theoretical Comparison of the FC, LC, and CLC contracts

Our goal here is to understand the relative performance of the three (FC, LC, and CLC) contracts given lower bounds ($\underline{e}$, $\underline{a}$) on the effort levels. To compare these contracts, we consider that the FC and the CLC contracts have the same lower bounds. Proposition 11 characterizes their relative performances under the SI model. Let $\Pi_{2k}^{\hat{N}}$, $\Pi_3^{\hat{N}}$ denote the sponsor's optimal expected retention costs for the SI (for given type k of the investigator) and OM models using the CLC contract.

PROPOSITION 11. *Under the SI Model, for a given k and $\underline{e}_k$, we have the following results:*

- *When $\theta_C^H \le 2\theta_C^L$: If $\underline{e}_k \in [0, e_1]$, we have $\Pi_{2k}^{\hat{N}}(\underline{e}_k) = \Pi_{2k}^N \le \Pi_{2k}^F(\underline{e}_k)$; otherwise, $\Pi_{2k}^{\hat{N}}(\underline{e}_k) = \Pi_{2k}^F(\underline{e}_k)$, where e_1 is the lowest effort level that $\Pi_{2k}^F(\underline{e}_k) = \Pi_{2k}^N$.*
- *When $\theta_C^H > 2\theta_C^L$: If $\underline{e}_k \in [0, e_1]$, we have $\Pi_{2k}^{\hat{N}}(\underline{e}_k) = \Pi_{2k}^N \le \Pi_{2k}^F(\underline{e}_k)$; otherwise, $\Pi_{2k}^F(\underline{e}_k) \le \Pi_{2k}^{\hat{N}}(\underline{e}_k)$, where e_1 is the lowest effort level that $\Pi_{2k}^F(\underline{e}_k) = \Pi_{2k}^N$.*

Figure 2 illustrates the results in the above proposition. First, note that under the SI model, either the LC or the FC contract performs better than the CLC contract. Second, notice that it is beneficial for the sponsor to use the LC contract instead of using the FC contract with a low or high $\underline{e}_k$. To explain this result, we rely on the computation study performed in Section 5 as we do not have a closed-form solution for the LC contract. Under the FC contract, the optimal effort level is the same as the lower bound on effort $\underline{e}_k$. Analyzing the instances where the LC contract is the best-performing contract among the three contracts, we notice that the optimal effort levels are higher (resp., lower) than those under the FC contract with low (resp., high) $\underline{e}_k$. Furthermore, for these instances, the optimal effort levels under the LC contract are closer to those under the CM model. These observations together imply that the FC contract with a low (resp., high) $\underline{e}_k$ performs worse than the LC contract as the sponsor underestimates (resp., overestimates) the value of $\underline{e}_k$ relative to optimal effort level under the CM model.



(a) $\theta_C^H \le 2\theta_C^L$ (b) $\theta_C^H > 2\theta_C^L$

Figure 2 Relative Performance of the FC, LC, and CLC Contracts Under the SI Model

As mentioned above, the cost of the CLC contract is no less than that of the FC or LC contract under the SI model. However, our next result shows that the CLC contract can strictly outperform both the FC and LC contracts under the OM model.

PROPOSITION 12. *When $\theta_C^H \le 2\theta_C^L$ and $\theta_I^H \le 2\theta_I^L$, there exist*

1. $\alpha_0, \sigma_1, \sigma_2 \in (0,1]$ *such that* $\Pi_3^{\hat{N}}(\underline{a}, \underline{e}) < \min\{\Pi_3^F(\underline{a}, \underline{e}), \Pi_3^N\}, \forall\ 0 \le \underline{a} \le \alpha_0, \sigma_1 \le \underline{e} \le \sigma_2$.
2. $\alpha_1, \alpha_2, \sigma_0 \in (0,1]$ *such that* $\Pi_3^{\hat{N}}(\underline{a}, \underline{e}) < \min\{\Pi_3^F(\underline{a}, \underline{e}), \Pi_3^N\}, \forall\ \alpha_1 \le \underline{a} \le \alpha_2, 0 \le \underline{e} \le \sigma_0$.

Deriving the relative performances of the three contracts for general situations under the OM model is theoretically challenging, hence, we resort to numerical analysis. Using the test bed described in the next section, we observe that the CLC contract performs better on average than both the FC and LC contracts when $\underline{e} = \underline{a} = 0.025$ and is dominated by the FC contract as $\underline{e}, \underline{a}$ increase. Finally, for high values of $\underline{e}, \underline{a}$, the FC and CLC contracts perform worse than the LC contract. We further observe that when $\underline{e} = \underline{a} = 0.025$, the CLC contract strictly outperforms the FC and the LC contracts in about 19% of the instances and reduces the expected retention cost by 2% on average over the best of the FC and the LC contracts (maximum 18% reduction).

The above results suggest that the relative performance of the three contracts in terms of the sponsor's costs may not be unidirectional and may vary based on the lower bounds of effort levels. Further, notice that in our theoretical results above, the thresholds on $\underline{e}, \underline{a}$ are functions of problem parameters, implying that for given $\underline{e}, \underline{a}$ values, the outperforming contract depend on other problem parameters as well. Hence, in the following section, we present an extensive computational analysis to understand the relative performance of the three contracts under different parameter settings.

5. Computational Study

In this computational study, our objective is to understand (1) the relative performance of the FC, LC, and CLC contracts and (2) when adopting the optimal contracts developed in Sections 3.2 and 3.3 can be significantly beneficial relative to the FC, LC, and CLC contracts. To this end, we calibrate our parameters using publicly available data. We also explore a wide range of possible values for a parameter that has no public data. We provide the details of the test bed next.

5.1. The Test Bed

Similar to Song et al. (2023), we consider a three-month clinical study with the number of participants $N = 100$. This value of N is consistent with NIH-funded studies (Gresham et al. 2018). Next, we discuss how we derive the parameter values. We summarize the ranges of parameters in Table 2.

Effort Cost Parameter: Following Wong et al. (2014) and Song et al. (2023), we consider the payment to clinical site staff (RN/CRA Costs) and physician (Physician Costs) to estimate the investigator's effort cost parameter. This results in a cost estimate of \$3000. For computing various effort cost parameters in our models, we first set the cost of implementing maximum effort by the low-type investigator $\frac{1}{2}\theta_I^L = \$3000$, which implies $\theta_I^L = \$6000$. Using this value of θ_I^L as a base, we then estimate the effort cost parameter of the high-type investigator as follows: $\frac{\theta_I^H}{\theta_I^L} \in \{1.05, 1.15, 1.25, 1.35\}$. We further consider $\frac{\theta_C^L}{\theta_I^L} \in \{0.4, 0.6, 0.8, 1.0\}$. Finally, we choose the values of θ_C^H such that $\frac{\theta_I^H - \theta_I^L}{\theta_C^H - \theta_C^L} \in \{0.5, 1.5, 2.5, 3.5\}$.

Parameter	Range	Parameter	Range
N	100	λ	$\{0.4, 0.55, 0.7, 0.85, 1.0\}$
$\bar{\delta}$	$\{0.6, 0.7, 0.8, 0.9\}$	λ_f	$\{0.03, 0.045, 0.06, 0.075\}$
θ_I^L	\$6000	λ_J	$\{0.05, 0.15, 0.25\}$
$\frac{\theta_I^H}{\theta_I^L}$	$\{1.05, 1.15, 1.25, 1.35\}$	p	$\{0.1, 0.3, 0.5, 0.7, 0.9\}$
$\frac{\theta_C^L}{\theta_I^L}$	$\{0.4, 0.6, 0.8, 1.0\}$	q	$\{0.1, 0.3, 0.5, 0.7, 0.9\}$
$\frac{\theta_I^H-\theta_I^L}{\theta_C^H-\theta_C^L}$	$\{0.5, 1.5, 2.5, 3.5\}$	$\underline{a}=\underline{e}$	$\{0.025, 0.125, 0.225, 0.325, 0.425\}$

Table 2 Parameter Settings

Effort and Payment Effectiveness: We consider different levels of effectiveness from the providers' effort and the sponsor's monetary payment by choosing (1) five levels for providers' effort effectiveness: $\lambda \in \{0.4, 0.55, 0.7, 0.85, 1.0\}$, (2) three levels for the joint impact of the providers' effort: $\lambda_J \in \{0.05, 0.15, 0.25\}$, (3) four levels for the effectiveness of the monetary payment: $\lambda_f \in \{0.03, 0.045, 0.06, 0.075\}$. We further consider the minimum effort requirement under the FC contract to be five different values $\underline{a}$ (and $\underline{e}$) $\in \{0.025, 0.125, 0.225, 0.325, 0.425\}$.

Type Distributions for the Investigator and the Coordinator: We consider five values of $p, q \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ to capture the impact of the distribution of the providers' effort cost parameters.

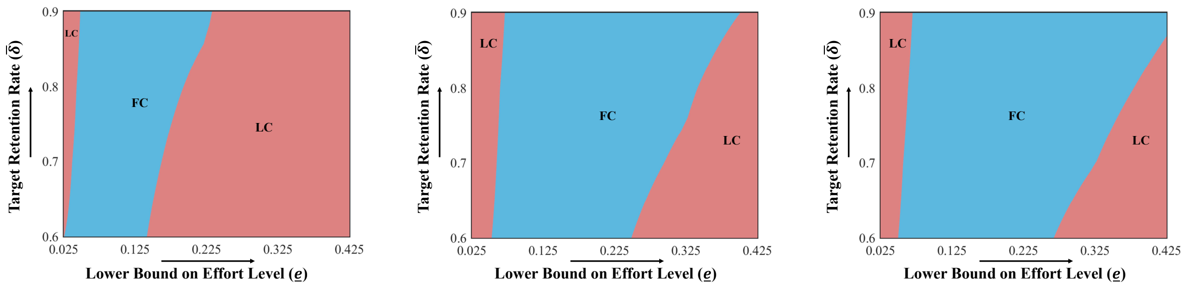
Retention Rate: We consider $\bar{\delta} \in \{0.6, 0.7, 0.8, 0.9\}$ to capture studies with different target retention rates (Evangelista 2013).

In summary, the above parameter settings result in a total of $1,920,000$ instances. Next, we describe the key findings from our computational study. Our first result compares the FC, LC, and CLC contracts.

5.2. Understanding the Relative Performance of the FC, LC, and CLC Contracts

Figure 3 illustrates the impact of the effort lower bound, the target retention rate, and the effectiveness of effort on the best-performing contract among the FC and LC contracts under the SI model. We do not include the CLC contract in this figure as it is dominated by the best of the FC and LC contracts (see Proposition 11). We observe that the FC contract performs better on average for medium values of $\underline{e}$ and the LC contract performs better otherwise. These observations are consistent with our findings in Section 4.4. The results under the OM model are structurally similar and are presented in Appendix EC.3.1.

To understand the observations in Figure 3, we first explain the benefits and drawbacks of each contract. Recall from Proposition 9, under the LC contract, the optimal effort given the provider effort cost parameter θ_C^l is $\frac{\beta}{\theta_C^l}$. Therefore, the optimal effort changes faster with respect to the payment per unit of effort β for the low-type provider than for the high-type. A potential drawback of the



(a) $\lambda = 0.4$ (b) $\lambda = 0.7$ (c) $\lambda = 1.0$

Figure 3 Comparing the Performances of the FC, LC, and CLC Contracts Under the SI Model

LC contract is that when the sponsor increases the value of β to ensure the retention constraint is achieved for the high-type provider, the low-type provider's effort level increases at a faster rate, which results in much higher compensation for the effort. A benefit of the LC contract is that it has the flexibility to adjust both the effort and the monetary payment as parameters change. In contrast, under the FC contract, the effort level from the provider equals the lower bound on effort $\underline{e}$ and does not change with other problem parameters. Hence, for a given $\underline{e}$, the sponsor can only change monetary payment to the participants as other parameters change. This property of the FC contract results in a drawback or benefit depending on the values of $\underline{e}$ and other parameters.

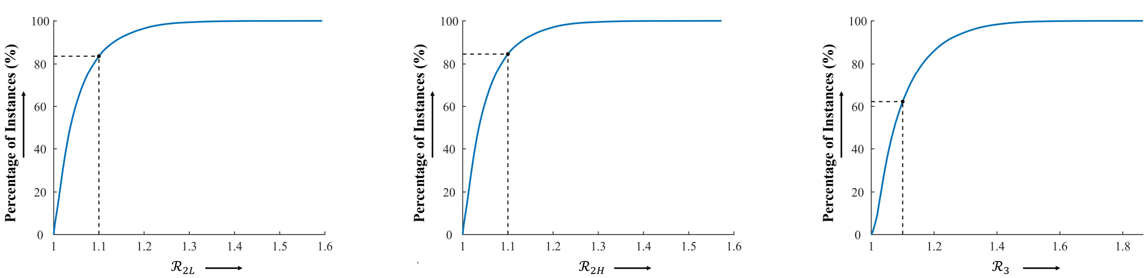
Using the above statements, we next explain the impact of various parameter changes on the relative performance of the two contracts. When $\bar{\delta}$ increases, the sponsor requires a higher effort and offers a higher β to satisfy the retention constraint under the LC contract and hence, it suffers from its drawback. As $\bar{\delta}$ increases, under the FC contract, at lower values of $\underline{e}$, the increase in total monetary payment to the participants (due to an increase in $\bar{\delta}$ and already high monetary payment) is higher than that under higher values of $\underline{e}$. In our computational analysis, we observe that when the sponsor sets a low $\underline{e}$ (resp., high), the negative impact of the increase in $\bar{\delta}$ on expected retention cost under the FC contract is higher (resp., lower) than that under the LC contract. Thus, the LC contract performs better.

As λ increases, the LC contract becomes more preferred over the FC contract with a low $\underline{e}$ value. Note that ceteris paribus, as λ increases, the effort becomes more efficient in improving retention as compared to monetary payment. Thus, the sponsor may benefit from decreasing the monetary payment and increasing providers' payments to increase effort. The LC contract offers this flexibility whereas the FC contract cannot change the effort level. When $\underline{e}$ is low, the FC contract suffers from its drawback. Hence, as λ increases, the resulting decrease in the expected retention cost under the FC contract is lower than that under the LC contract.

When the sponsor sets $\underline{e}$ to be medium or high, the regions in which the LC contract outperforms the FC contract become smaller as λ increases. When λ is small, the FC contract suffers from its drawback whereas the LC contract benefits from its flexibility and hence, performs better than the FC contract. As λ increases, under the FC contract, the sponsor can achieve the same retention rate with less monetary payment and the expected retention cost decreases. For the LC contract, as λ increases, the sponsor reduces the monetary payment while relying more on effort. However, as explained earlier, the LC contract suffers from the drawback that the optimal effort changes at a different rate for different types of providers. Therefore, this transition from monetary payment to exerting higher effort yields less cost reduction under the LC contract as compared to the FC contract.

5.3. When to Adopt the Optimal Compensation Contracts

In this section, we compare the best of the FC, LC, and CLC contracts in terms of cost with the optimal contracts under the SI and OM models discussed in Sections 3.2 and 3.3. In particular, given a parameter setting, we consider the following ratios: $\mathcal{R}_{2k} = \dfrac{\min_{t\in\{F,N,\hat{N}\}} \Pi_{2k}^t}{\Pi_k^{SI^\circ}}$ for the SI model, $k \in \{L, H\}$ and $\mathcal{R}_3 = \dfrac{\min_{t\in\{F,N,\hat{N}\}} \Pi_3^t}{\Pi^{OM^\diamond}}$ for the OM model. We observe that this ratio is 1.06 (resp., 1.10) on average under the SI (resp., OM) model. Figure 4 shows the cumulative distributions of the values of the above two ratios in our computational study. As illustrated in the figure, $\mathcal{R}_{2k}, k \in \{L, H\}$, (resp., $\mathcal{R}_3$) is less than 1.10 under the SI (resp., OM) model in about 83% (resp., 62%) of the instances, implying that the performance of the best of three contracts in terms of cost is generally satisfactory. However, this ratio could reach 1.60 (resp., 1.86) under the SI (resp., OM) model and hence, necessitates identifying parametric settings where the compensation contracts observed in practice may underperform significantly. In these instances, the sponsor should consider adopting the optimal compensation contract. We next explore these instances for the SI model to identify parametric settings where using the optimal compensation contract provided in Section 3.2 can be significantly beneficial. The insights into the parametric settings are similar for the OM model and are provided in Appendix EC.3.1.

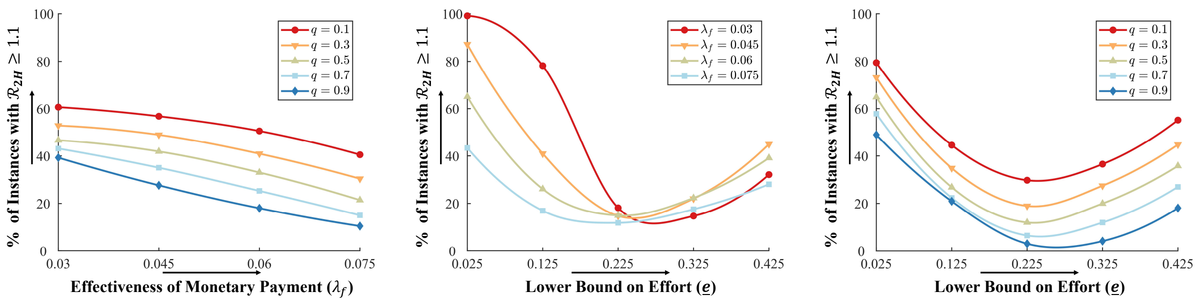


(a) SI Model, $k = L$ (b) SI Model, $k = H$ (c) OM Model

Figure 4 **Cumulative Distribution of $\mathcal{R}_{2k}$ and $\mathcal{R}_3$**

Among the instances where $\mathcal{R}_{2k} \geq 1.10$, about 66% instances have effectiveness of effort parameter $\lambda \geq 0.85$. Recall that a higher λ means the effort becomes more effective in retaining participants.

Hence, the costs of both the optimal contract and the best-performing (among the FC, LC, and CLC) contract decrease as λ increases. However, the cost of the optimal contract decreases at a faster rate than that of the best-performing contract, and hence the ratio of the two costs is higher at higher λ. Our study of the computation results further suggests that the effect of higher λ on the ratio is compounded (i.e., the ratio typically exceeds 1.10) when at least two of the following conditions are satisfied: (1) the sponsor sets the lower bound on effort ($\underline{e}$) to be small, (2) the effectiveness of monetary payment (λ_f) is low, and (3) the probability of having a high-type coordinator (q) is low. For instances with $\lambda \geq 0.85$, we illustrate the impact of $\underline{e}, \lambda_f, q$ on the ratio in Figure 5 and explain the impact below:



(a) (b) (c)

Figure 5 Impact of $\underline{e}, \lambda_f$, and q on the Ratio $\mathcal{R}_{2H}$

1. λ_f and q are low (Figure 5a): Using the results in Section 3.2, it is easy to verify that the cost of the optimal contract is affected by these two factors with opposite impacts. On the one hand, adopting monetary payment becomes less cost-effective with a low λ_f, and hence, the expected retention cost increases. On the other hand, a low q suggests that the sponsor has a higher likelihood to contract with a coordinator who is less costly in exerting effort, which reduces the expected retention cost. Thus, the cost of the optimal contract can increase or decrease as both λ_f and q decrease. For the compensation contracts observed in practice, a low λ_f also increases the cost of achieving the target retention rate. Further, the cost under the FC contract does not depend on q (see Section 4.1) and it is easy to show that the costs under the LC and CLC contracts become higher as q decreases. Hence, a decrease in both parameters drives up the cost of the compensation contracts observed in practice. When the cost of the optimal contract decreases as both λ_f and q decrease, the impact on the ratio is obvious. When the cost of the optimal contract increases, this increase is lower than that for the compensation contracts observed in practice and thus, the ratio becomes higher when both λ_f and q are low.

2. $\underline{e}$ and λ_f are low (Figure 5b): Observe from Proposition 11, the LC is the best-performing contract when the sponsor sets a low value of $\underline{e}$ under the FC or the CLC contract. Notice that the

costs of the optimal contract and the LC contract do not depend on $\underline{e}$. Further, both costs increase as λ_f decreases. However, the increase in cost under the LC contract is larger (owing to the drawback of the LC contract mentioned in Section 5.2) compared to the optimal contract, resulting in the ratio increasing as both $\underline{e}$ and λ_f decrease.

3. q and $\underline{e}$ are low (Figure 5c): the observations under this case follow directly from the arguments mentioned above.

6. Impact of Uncertainty in the Retention Rate on the Performance of the FC, LC, and CLC contracts

In our analysis thus far, we considered the effectiveness of providers' efforts (λ, λ_J) and the sponsor's payment in retaining participants (λ_f) to be deterministic. However, one could argue that in practice, these effectiveness parameters can be uncertain. Hence, in this section, we consider these parameters as random variables: $\tilde{\lambda}, \tilde{\lambda}_J, \tilde{\lambda}_f$. Further, unobserved factors may influence retention, which we capture by introducing an additional random variable, ϵ. Consequently, we have $\tilde{\delta}(a,e,f) = \tilde{\lambda}a + \tilde{\lambda}e + \tilde{\lambda}_J ae + \tilde{\lambda}_f f + \epsilon$, where $\phi := (\tilde{\lambda}, \tilde{\lambda}_J, \tilde{\lambda}_f, \epsilon)$ is a random vector. We assume that ϕ follows a multivariate distribution with mean μ, and covariance matrix Σ.

Before proceeding with the analysis in this section, it is important to note that these uncertainties do not impact the problem descriptions and, therefore, the structure of the optimal decisions of the providers (the coordinator under the SI model and both providers under the OM model). Further, we can modify the decision-making problem for the sponsor (studied in previous sections) by writing the retention constraint for each random scenario as follows: $\tilde{\delta}(a,e,f) \geq \bar{\delta}$. To solve the sponsor's problem, we relax the retention constraint and require it to be satisfied for only ζ percentage of scenarios. That is,

$$Prob\left(\tilde{\delta}(a,e,f) \geq \bar{\delta}\right) \geq \zeta. \tag{16}$$

Solving the sponsor's problem with retention constraint (16) is analytically challenging, in general. However, it is straightforward to observe that when the uncertainty is due to the unobserved factors only (i.e., $\lambda, \lambda_J, \lambda_f$ are deterministic), the sponsor's problem can be solved by replacing $\bar{\delta}$ with $\bar{\delta} - \Phi^{-1}(1-\zeta)$ in the retention constraint considered in each of the previous sections. Here, Φ is the CDF of ϵ. Below, we present a numerical study to derive insights into the impact of uncertainties in $\lambda, \lambda_J, \lambda_f$, and due to the unobserved factors, on the performances of various compensation contracts.

6.1. Computational Analysis

We consider the test bed described in Section 5.1 and set the mean of ϕ as $\mu = (\lambda, \lambda_J, \lambda_f, 0)$, with the covariance matrix $\Sigma = \text{diag}(\alpha^2\lambda^2, \alpha^2\lambda_J^2, \alpha^2\lambda_f^2, \alpha^2)$ for each instance, where values of $\lambda, \lambda_J, \lambda_f$ are

α	Centralized Model	SI Model Optimal Contract	SI Model FC contract	SI Model LC contract	SI Model CLC contract	OM Model Optimal Contract	OM Model FC contract	OM Model LC contract	OM Model CLC contract
0.01	5.26%	5.28%	4.79%	5.57%	4.89%	4.80%	3.67%	11.41%	3.60%
0.02	10.72%	10.75%	9.75%	11.24%	9.81%	10.26%	8.08%	17.89%	7.94%
0.03	16.47%	16.53%	14.97%	17.31%	14.98%	16.03%	12.60%	24.81%	12.39%

Table 3 Percentage Increase in Expected Retention Cost for Different Levels of Uncertainty

specified in Table 2. We then consider three values of α as $\{0.01, 0.02, 0.03\}$ such that the probability of realizing negative values of random variables is significantly small.

Table 3 displays the increase in the optimal expected retention cost because of uncertainty for the models analyzed in previous sections. For a given α, the value in each cell of the table is an average number over all instances in our test bed. We observe that as α increases, the cost increase due to uncertainty becomes more significant, which is as expected. Further, notice that the FC and the CLC contracts are relatively more robust to the uncertainty than the LC contract. As discussed in Section 5.2, the drawback of the LC contract is that when the sponsor increases the value of β (and ν) to ensure the retention constraint is achieved for the high-type provider, the low-type provider's effort level increases at a faster rate, which results in much higher compensation for the effort. As compared to the LC contract, the FC contract does not suffer from this drawback as a provider exerts the same level of effort regardless of the type. Hence, the FC contract becomes more robust to uncertainty relative to the LC contract. For the CLC contract, although the provider's effort also depends on the payment per unit of effort, the existence of a lower bound on effort ensures the difference in the effort level between the low-type and the high-type provider is not as significant as that under the LC contract. Therefore, the CLC contract is more robust to uncertainty than the LC contract. Finally, under the OM model, the sponsor contracts with both providers, where the respective low-type provider's effort level increases at a faster rate than the high-type provider. Therefore, the sponsor suffers from the elevated cost increase from both providers; thus, the cost increase is much higher for the OM model than the SI model for the LC contract.

We also explore the impact of uncertainty on the performance of the FC and CLC contracts as the lower bound of effort changes (see Table 4). For a given value(s) of the lower bound(s), the value

$\underline{a} = \underline{e}$	FC Contract SI Model	FC Contract OM Model	CLC Contract SI Model	CLC Contract OM Model
0.025	8.52%	8.81%	8.77%	8.39%
0.125	8.23%	8.19%	8.19%	8.14%
0.225	7.49%	6.67%	7.45%	6.64%
0.325	6.49%	4.67%	6.46%	4.66%
0.425	5.36%	1.32%	5.43%	1.31%

Table 4 Percentage Increase in Expected Retention Cost Due to Uncertainty for Different Lower Bounds on Effort

in each cell of the table is an average number over all instances in our test bed. We observe that the robustness of both contracts increases for the higher values of $\underline{e}$. The reason is as follows: for the FC contract, note that the providers exert an effort that is equal to their respective lower bounds. Under the deterministic case, when the lower bounds on effort levels are small, the effort levels from the providers are limited and the sponsor relies largely on the monetary payment to satisfy the retention constraint. On the other hand, when the lower bounds on effort levels are high, the providers' effort levels are also high, while the monetary payments to the participants are small. Since providers' effort levels are fixed at the lower bound for the FC contract, to satisfy Constraint (16) under the uncertainty case, the sponsor must increase the monetary payment to the participants relative to that under the deterministic case. This increase in monetary payment decreases with lower bounds on effort levels, resulting in a lower increase in the expected retention cost for higher values of $\underline{e}$ (and $\underline{a}$). The increase in monetary payment as lower bounds on effort levels increase is further lower under the OM model as compared to the SI model. Hence, we also observe that the FC contract is more robust under the OM model. For the CLC contract, when the lower bounds on effort levels are small, the provider's effort is generally higher than the lower bound values, and the CLC contract performs similarly to the LC contract, which is less robust to uncertainty. On the other hand, when the lower bounds on effort levels are high, the provider's effort is close to the lower bound values, and the CLC contract performs similarly to the FC contract. Hence, the percentage cost increase due to uncertainty is smaller for higher values of $\underline{e}$ (and $\underline{a}$) following a similar argument as for the FC contract.

We conclude this section by stating that the insights regarding the relative performance of the three contracts and when to adopt the optimal contracts provided in Sections 5.2–5.3 hold under uncertainty as well. Hence, for brevity, we do not repeat the discussion about those insights here.

7. Conclusion

In this work, we study various compensation contracts for providers to address the participant retention challenges faced by clinical studies. We identify three clinical study team structures in practice: (i) the centralized model where the sponsor decides monetary payment to participants and effort levels for the providers, (ii) the sponsor-investigator (SI) model, and (iii) the outsourcing (OM) model. Given a decentralized structure of the clinical study team, our analysis provides the optimal compensation contract that minimizes the expected retention cost of achieving a target retention rate. We also examine the impact of different problem primitives on the values of contract parameters. We then identify the sponsor's optimal decisions under the three compensation contracts observed in practice, namely, the fixed compensation (FC), the linear compensation (LC), and the conditional linear compensation (CLC) contracts. We compare the relative performance of the above contracts.

Finally, we extend our analysis to understand the impact of uncertainty in retention (due to various factors) on the relative performance of the compensation contracts.

Outsourcing clinical investigations is a common strategy among pharmaceutical companies where more than half of the clinical studies are outsourced (Spinner 2021). Our theoretical analysis reveals that the additional expenditure associated with the SI (resp., OM) model is at most 20% (resp., 40%). However, we observe that, in practical settings, this additional expenditure is typically much lower (maximum 6% and 8% for the SI and OM model, respectively) than these theoretical bounds. Hence, given a decentralized team structure, by optimally designing the contracts for providers, the sponsor can achieve the target retention rate in a cost-effective manner under most practical settings. Our analysis further suggests that providers' efforts and corresponding payments are higher under a centralized team structure, whereas participants' payments are higher under the decentralized team structures.

For better quality and safety management, sponsors of clinical studies seeking outsourced services often set standards for the effort levels of the providers. Consider, for example, clinical studies on genetic therapeutics which require approvals from both the Institutional Biosafety Committee and Institutional Review Board (see Office of Biotechnology Activities Oversight 2023, NIH Office of Biotechnology Activities 2018) and need extensive participant education and consent processes. For such studies, the standards may be moderate to high, and the FC or the CLC contract will be equivalent under the SI model and the FC contract will be beneficial under the OM model. When the standards are low, the FC contract is beneficial under the SI model whereas the CLC contract could outperform the FC contract under the OM model. For example, observational studies on healthy participants generally have simpler protocols and fewer logistics (e.g., Chiofalo 2023, Bradbury 2019) and hence, the standards may typically be low. Sponsors' of such studies will benefit from offering the FC or CLC contract depending on the structure of the team. Note, however, that if these standards on effort levels face a challenge of overestimation or underestimation, the sponsor may consider offering the LC contract instead of imposing a lower bound and offering the FC or the CLC contract.

We also identify parameter settings for which the sponsor benefits significantly by adopting the optimal contracts over the three contracts observed in practice. In clinical studies that involve behavioral changes or sensitive topics, continuous support from investigators or coordinators can be more effective, and monetary payments are usually less effective for retaining participants. For such studies, if the standards on effort levels are low or the possibility of having providers with low costs of implementing efforts is high, the sponsor should adopt the optimal contracts. Consider another example of clinical studies on vaccines for infectious diseases that involve fewer medical procedures, while requiring more follow-up reports from the participants (Janssen 2021, Konopnicki 2018). For such studies, the effort can have high effectiveness as appreciation and communication can have a huge impact on participant retention whereas the standards on effort levels are generally low. In addition to these characteristics, if the possibility of having providers with low costs of implementing efforts is high, adopting the optimal contract is significantly more advantageous.

Finally, we comment on the assumptions and limitations of our analysis. First, for ease of exposition, we consider there is only one coordinator. Our analysis can be easily extended if there are multiple coordinators. Second, under the OM model, we only consider the situation where the sponsor offers the same type of contract to both providers because our goal is to understand the relative performance of the three contracts observed in practice. However, if the goal is simply to minimize the retention cost for the sponsor, one may consider the possibility that the sponsor can provide different types of contracts to different providers.

References

Adida, E. and Bravo, F. (2019). Contracts for healthcare referral services: Coordination via outcome-based penalty contracts. *Management Science*, 65(3):1322–1341.

Ahuja, V. and Birge, J. R. (2020). An approximation approach for response-adaptive clinical trial design. *INFORMS Journal on Computing*, 32(4):877–894.

Alban, A., Chick, S. E., and Forster, M. (2023). Value-based clinical trials: Selecting recruitment rates and trial lengths in different regulatory contexts. *Management Science*, 69(6):3516–3535.

Anderer, A., Bastani, H., and Silberholz, J. (2022). Adaptive clinical trial designs with surrogates: When should we bother? *Management Science*, 68(3):1982–2002.

Andritsos, D. A. and Tang, C. S. (2018). Incentive programs for reducing readmissions when patient care is co-produced. *Production and Operations Management*, 27(6):999–1020.

Arora, V. (2019). Inpatient sleep loss: Educating and empowering patients (ISLEEP). Available at `https://www.clinicaltrials.gov/ct2/show/NCT04151251`.

Atlant Clinical (2020). Patient retention in clinical trials: Problems, costs and solutions. Available at `https://www.atlantclinical.com/patient-retention-in-clinical-trials`.

Atsawasuwan, P. (2020). Gingival crevicular fluid characterization during orthodontic treatment. Identifier NCT04226469, Available at `https://clinicaltrials.gov/ct2/show/NCT04226469`.

Baer, A. R., Devine, S., Beardmore, C. D., and Catalano, R. (2011). Clinical investigator responsibilities. *Journal of Oncology Practice*, 7(2):124–128.

Bellos, I. and Kavadias, S. (2021). Service design for a holistic customer experience: A process framework. *Management Science*, 67(3):1718–1736.

Bertsimas, D. and Sturt, B. (2020). Computation of exact bootstrap confidence intervals: Complexity and deterministic algorithms. *Operations Research*, 68(3):949–964.

Booker, C. L., Harding, S., and Benzeval, M. (2011). A systematic review of the effect of retention methods in population-based cohort studies. *Bmc Public Health*, 11(1):1–12.

Bradbury, M. (2019). A study of a device used to take pictures of skin blood flow. Available at `https://clinicaltrials.gov/ct2/show/NCT03912376`.

Chhatre, S., Jefferson, A., Cook, R., Meeker, C. R., Kim, J. H., Hartz, K. M., Wong, Y.-N., Caruso, A., Newman, D. K., Morales, K. H., and Jayadevappa, R. (2018). Patient-centered recruitment and retention for a randomized controlled study. *Trials*, 19(1):1–10.

Chiofalo, B. (2023). Endometrial changes in breast cancer women. Available at `https://clinicaltrials.gov/ct2/show/NCT05717634`.

ClinicalTrials.gov (2020). Learn about clinical studies. Available at `https://clinicaltrials.gov/ct2/about-studies/learn`.

Colvin, M. and Maravelias, C. T. (2010). Modeling methods and a branch and cut algorithm for pharmaceutical clinical trial planning using stochastic programming. *European Journal of Operational Research*, 203(1):205–215.

Corbett, C. J. and DeCroix, G. A. (2001). Shared-savings contracts for indirect materials in supply chains: Channel profits and environmental impacts. *Management Science*, 47(7):881–893.

Dragon, R., Garvey, G. T., and Turnbull, G. K. (1996). A collective tournament. *Economics Letters*, 50(2):223–227.

ElectroCore INC (2019). Vagus nerve stimulation for management of bronchoconstriction in patients hospitalized with copd exacerbations. Available at `https://clinicaltrials.gov/ct2/show/NCT01612793`.

Evangelista, E. (2013). Patient recruitment: The most recognized silent driver of clinical development costs. *International Pharmaceutical Industry*, 5(1).

Feehan, A. K. and Garcia-Diaz, J. (2020). Investigator responsibilities in clinical research. *Ochsner Journal*, 20(1):44–49.

Fleischhacker, A., Ninh, A., and Zhao, Y. (2015). Positioning inventory in clinical trial supply chains. *Production and Operations Management*, 24(6):991–1011.

Fleischhacker, A. J. and Zhao, Y. (2011). Planning for demand failure: A dynamic lot size model for clinical trial supply chains. *European Journal of Operational Research*, 211(3):496–506.

Fouad, M. N., Johnson, R. E., Nagy, M. C., Person, S. D., and Partridge, E. E. (2014). Adherence and retention in clinical trials: A community-based approach. *Cancer*, 120:1106–1112.

Genesis Research Services (2021). Clinical trials – the role of the coordinator. Available at `https://genesisresearchservices.com/clinical-trials-the-role-of-the-coordinator/`.

Ghamat, S., Zaric, G. S., and Pun, H. (2021). Care-coordination: Gain-sharing agreements in bundled payment models. *Production and Operations Management*, 30(5):1457–1474.

Goh, J., Bayati, M., Zenios, S. A., Singh, S., and Moore, D. (2018). Data uncertainty in markov chains: Application to cost-effectiveness analyses of medical innovations. *Operations Research*, 66(3):697–715.

Gresham, G. K., Ehrhardt, S., Meinert, J. L., Appel, L. J., and Meinert, C. L. (2018). Characteristics and trends of clinical trials funded by the national institutes of health between 2005 and 2015. *Clinical Trials*, 15(1):65–74.

Guillen, P., Merrett, D., and Slonim, R. (2015). A new solution for the moral hazard problem in team production. *Management Science*, 61(7):1514–1530.

Gupta, D., Mehrotra, M., and Tang, X. (2021). Gainsharing contracts for CMS' episode-based payment models. *Production and Operations Management*, 30(5):1290–1312.

Hauser, J. R., Simester, D. I., and Wernerfelt, B. (1994). Customer satisfaction incentives. *Marketing Science*, 13(4):327–350.

HHS.gov (2019). Attachment A - addressing ethical concerns offers of payment to research participants. Available at `https://www.hhs.gov/ohrp/sachrp-committee/recommendations/attachment-a-september-30-2019/index.html`.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics*, 13(2):324–340.

Hu, Y., Shin, J., and Tang, Z. (2016). Incentive problems in performance-based online advertising pricing: Cost per click vs. cost per action. *Management Science*, 62(7):2022–2038.

Hunsley, J. (2020). An online post-adoption family intervention using trust-based relational intervention (TBRI) to improve understanding, well-being, and connection in adoptive families. Identifier NCT04443504, Available at `https://clinicaltrials.gov/ct2/show/NCT04443504`.

Ingram, M. (2021a). Combine compensation models for more precise investigator payments. Available at `https://cms.centerwatch.com/articles/25827-combine-compensation-models-for-more-precise-investigator-payments`.

Ingram, M. (2021b). Compensating investigators: Core principles of compensation and contracting. Available at `https://cms.centerwatch.com/articles/25811-compensating-investigators-core-principles-of-compensation-and-contracting`.

Ingram, M. (2021c). Methods for calculating PI compensation: Pros and cons. Available at `https://cms.centerwatch.com/articles/25818-methods-for-calculating-pi-compensation-pros-and-cons`.

Institute of Medicine (2015). *Sharing Clinical Trial Data*. National Academies Press.

Janssen (2021). A study of Ad26.COV2.S and influenza vaccines in healthy adults. Available at `https://clinicaltrials.gov/ct2/show/NCT05091307`.

Jiang, H., Pang, Z., and Savin, S. (2012). Performance-based contracts for outpatient medical services. *Manufacturing & Service Operations Management*, 14(4):654–669.

Konopnicki, D. (2018). Vaccination against human papillomavirus (HPV) with the 9-valent vaccine in hiv-positive women (the papillon study) (papillon). Available at `https://clinicaltrials.gov/ct2/show/NCT03391921`.

Kouvelis, P., Milner, J., and Tian, Z. (2017). Clinical trials for new drug development: Optimal investment and application. *Manufacturing & Service Operations Management*, 19(3):437–452.

Lafontaine, F. and Slade, M. E. (1996). Retail contracting and costly monitoring: Theory and evidence. *European Economic Review*, 40(3-5):923–932.

Lin, R. S. Y. (2021). Effects of an empowerment-based psycho-behavioral program on persons with mild cognitive impairment. Identifier NCT04723667, Available at `https://clinicaltrials.gov/ct2/show/NCT04723667`.

Lutgendorf, S. (2019). Biobehavioral-cytokine interactions in ovarian cancer (VEGF). Identifier NCT01113112, Available at `https://clinicaltrials.gov/ct2/show/NCT01113112`.

McAfee, R. P. and McMillan, J. (1991). Optimal contracts for teams. *International Economic Review*, 32(3):561–577.

McCabe, C. (2020). Clinical investigation of the dextenza 0.4mg intracanalicular insert delivery system used in cataract surgery. Identifier NCT04687800, Available at `https://clinicaltrials.gov/ct2/show/NCT04687800`.

mdgroup (2020). Why do patients drop out of clinical trials? Available at `https://mdgroup.com/blog/why-do-patients-drop-out-of-clinical-trials/`.

Milgrom, P. and Roberts, J. (1995). Complementarities and fit strategy, structure, and organizational change in manufacturing. *Journal of Accounting and Economics*, 19(2-3):179–208.

Mookherjee, D. (2006). Decentralization, hierarchies, and incentives: A mechanism design perspective. *Journal of Economic Literature*, 44(2):367–390.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73.

National Research Council (2010). *The prevention and treatment of missing data in clinical trials*. National Academies Press.

NIH Office of Biotechnology Activities (2018). Requirements for DAIDS-supported research involving recombinant DNA. Available at `https://rsc.niaid.nih.gov/sites/default/files/IBCReqsRecombinantDNA.pdf`.

Nuttall, A. (2012). Considerations for improving patient recruitment into clinical trials. *Clinical Leader Newsletter*.

Office of Biotechnology Activities Oversight (2023). Clinical trials and the IBC. Available at `https://research.umn.edu/units/obao/research-oversight-areas/clinical-trials-and-ibc`.

Office of Institutional Compliance (2021). Fixed price contracts, including clinical trials. Available at `https://policy.umn.edu/research/clinicaltrials`.

Parkinson, B., Meacock, R., Sutton, M., Fichera, E., Mills, N., Shorter, G. W., Treweek, S., Harman, N. L., Brown, R. C., Gillies, K., et al. (2019). Designing and using incentives to support recruitment and retention in clinical trials: a scoping review and a checklist for design. *Trials*, 20(1):1–14.

Partners HealthCare (2017). Bonus payments in clinical trial agreements. Available at `https://www.partners.org/Assets/Documents/Medical-Research/Clinical-Research/Bonus-Payments-in-CTA.pdf`.

Pfizer (2011). Pfizer announces new strategic partnerships with icon and parexel international corporation. Available at `https://www.pfizer.com/news/press-release/press-release-detail/pfizer_announces_new_strategic_partnerships_with_icon_and_parexel_international_corporation`.

Rajagopalan, S. and Tong, C. (2022). Payment models to coordinate healthcare providers with partial attribution of outcome costs. *Manufacturing & Service Operations Management*, 24(1):600–616.

Retinal Consultants of Houston (2019). Treat & extend treatment with 0.5mg ranibizumab vs monthly treatment with 0.5mg ranibizumab (T-REX). Identifier NCT01748292, Available at `https://clinicaltrials.gov/ct2/show/NCT01748292`.

Roels, G. (2014). Optimal design of coproductive services: Interaction and work allocation. *Manufacturing & Service Operations Management*, 16(4):578–594.

Schleider, J. (2021). Evaluating the telehealth single session consultation service. Identifier NCT04725578, Available at `https://clinicaltrials.gov/ct2/show/NCT04725578`.

Song, X., Mehrotra, M., and Rajapakshe, T. (2023). An analysis of incentive schemes for participant retention in clinical studies. *Manufacturing & Service Operations Management*, 25(3):1033–1050.

Spinner, J. (2021). Cros fighting for the future of clinical research: report. Available at `https://www.outsourcing-pharma.com/Article/2021/03/29/CROs-fighting-for-the-future-of-clinical-research-report`.

Swanson, G. (2020). Circadian misalignment in inflammatory bowel disease. Identifier NCT04637399, Available at `https://clinicaltrials.gov/ct2/show/NCT04637399`.

Tian, Z., Han, W., and Powell, W. B. (2021). Adaptive learning of drug quality and optimization of patient recruitment for clinical trials with dropouts. *Manufacturing & Service Operations Management*, 24(1):580–599.

Tian, Z., Hazen, G. B., and Li, H. (2023). Optimal enrollment in late-stage new drug development with learning of drug's efficacy for group-sequential clinical trials. *Manufacturing & Service Operations Management*, 25(1):88–107.

Tucker, R. (2021). Evaluation of sleep program. Identifier NCT04727619, Available at `https://clinicaltrials.gov/ct2/show/NCT04727619`.

UCSF (2022). A study of long-term al001 dosing in frontotemporal dementia (FTD) patients (INFRONT-2). Available at `https://memory.ucsf.edu/research-trials/clinical-trials/infront-2`.

U.S. FDA (2018). Payment and reimbursement to research subjects: Guidance for institutional review boards and clinical investigators. Available at `https://www.fda.gov/regulatory-information/search-fda-guidance-documents/payment-and-reimbursement-research-subjects`.

Wong, H.-H., Jessup, A., Sertkaya, A., Birkenbach, A., Berlind, A., and Eyraud, J. (2014). Examination of clinical trial costs and barriers for drug development final. Available at `https://aspe.hhs.gov/report/examination-clinical-trial-costs-and-barriers-drug-development`.

E-Companion: Online Appendix for "Evaluating the Efficacy of Providers' Compensation Contracts in Improving Participant Retention for Clinical Studies"

EC.1. Proofs of Technical Results

We first establish the following inequalities that are used in subsequent proofs of our theoretical results.

LEMMA EC.1. *For any given set of parameters we have,*

1. $\bar{q}\theta_C^H - (2\bar{q}-1)\theta_C^L \le \theta_C^H - \bar{q}\theta_C^L$,
2. $q\theta_C^H \le \theta_C^H - \bar{q}\theta_C^L$,
3. $\lambda_f^2\theta_I^L(\theta_C^H - \bar{q}\theta_C^L) - q\bar{\delta}^2N^2\lambda_J^2 > 0$,
4. $\lambda_f^2\theta_C^L\theta_I^L(\theta_C^H - \bar{q}\theta_C^L) - \bar{\delta}^2N^2\lambda_J^2(\bar{q}\theta_C^H - (2\bar{q}-1)\theta_C^L) > 0$.

EC.1.1. Proof of Lemma EC.1

Note that

1. $\bar{q}\theta_C^H - (2\bar{q}-1)\theta_C^L = \theta_C^H - \bar{q}\theta_C^L - q(\theta_C^H - \theta_C^L) \le \theta_C^H - \bar{q}\theta_C^L$,
2. $\theta_C^H - \bar{q}\theta_C^L = q\theta_C^H + \bar{q}(\theta_C^H - \theta_C^L) \ge q\theta_C^H$,
3. $\lambda_f^2\theta_I^L(\theta_C^H - \bar{q}\theta_C^L) - q\bar{\delta}^2N^2\lambda_J^2 \ge q(\lambda_f^2\theta_I^L\theta_C^H - \bar{\delta}^2N^2\lambda_J^2) > 0$,
4. $\lambda_f^2\theta_C^L\theta_I^L(\theta_C^H - \bar{q}\theta_C^L) - \bar{\delta}^2N^2\lambda_J^2(\bar{q}\theta_C^H - (2\bar{q}-1)\theta_C^L) \ge (\bar{q}\theta_C^L - (2\bar{q}-1)\theta_C^L)(\lambda_f^2\theta_I^L\theta_C^H - \bar{\delta}^2N^2\lambda_J^2) > 0$.

The proof is now completed. □

EC.1.2. Proof of Proposition 1

We first make the following claim.

CLAIM EC.1. *Under the optimal solution, constraint* (2) *is binding.*

Proof We prove this by contradiction. Let the optimal solution be $(a_{kl}^*, e_{kl}^*, f_{kl}^*)$, $k,l \in \{L,H\}$. Suppose the constraint (2) is not binding, i.e., $\delta(a_{kl}^*, e_{kl}^*, f_{kl}^*) > \bar{\delta}$. Since $\delta(a_{kl}, e_{kl}, f_{kl})$ increases in a_{kl}, e_{kl}, and f_{kl}, there exists $\gamma \in (0,1)$ such that $\delta(\gamma a_{kl}^*, \gamma e_{kl}^*, \gamma f_{kl}^*) \ge \bar{\delta}$. Further, the objective function is increasing in a_{kl}, e_{kl}, and f_{kl}. Therefore, $\Pi_{kl}^{CM}(\gamma a_{kl}^*, \gamma e_{kl}^*, \gamma f_{kl}^*) \le \Pi_{kl}^{CM}(a_{kl}^*, e_{kl}^*, f_{kl}^*)$ which contradicts the optimality of the solution $(a_{kl}^*, e_{kl}^*, f_{kl}^*)$. The result now follows. □

From Claim EC.1, we have $f_{kl} = \frac{\bar{\delta} - \lambda(a_{kl}+e_{kl}) - \lambda_J a_{kl}e_{kl}}{\lambda_f}$. Thus, for $k,l \in \{L,H\}$, the sponsor's problem reduces to the following

$$\min_{a_{kl},e_{kl}\in[0,1]} \Pi_{kl}^{CM} = \frac{\theta_C^l {e_{kl}}^2}{2} + \frac{\theta_I^k {a_{kl}}^2}{2} + \frac{\bar{\delta}N(\bar{\delta} - \lambda(a_{kl}+e_{kl}) - \lambda_J a_{kl}e_{kl})}{\lambda_f}, \tag{EC.1}$$

$$\text{s.t.} \quad \bar{\delta} - \lambda(a_{kl}+e_{kl}) - \lambda_J a_{kl}e_{kl} \ge 0. \tag{EC.2}$$

Let $\bar{a}_{kl}$ and $\bar{e}_{kl}$ be the optimal solution to the unconstrained optimization problem. Note that the Hessian matrix, $H = \begin{bmatrix} \theta_I^k & -\frac{\bar{\delta}N\lambda_J}{\lambda} \\ -\frac{\bar{\delta}N\lambda_J}{\lambda} & \theta_C^l \end{bmatrix}$. It is easy to show that H is positive definite. Thus, using first order conditions we have, $\bar{a}_{kl} = \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J+\lambda_f\theta_C^l)}{\lambda_f^2\theta_C^l\theta_I^k-\bar{\delta}^2N^2\lambda_J^2}$, and $\bar{e}_{kl} = \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J+\lambda_f\theta_I^k)}{\lambda_f^2\theta_C^l\theta_I^k-\bar{\delta}^2N^2\lambda_J^2}$.

We next prove that the solution $(\bar{a}_{kl}, \bar{e}_{kl})$ is the optimal solution to the sponsor's constrained problem as well. To this end, it suffices to show that $0 < \bar{a}_{kl}, \bar{e}_{kl} < 1$ and constraint EC.2 is satisfied.

Using inequalities from Lemma EC.1 and Assumption 1, we can show that $0 < \bar{a}_{kl}, \bar{e}_{kl} < 1$. Also, using Assumption 1, it is straightforward to show that the constraint (EC.2) is satisfied at the solution $(\bar{a}_{kl}, \bar{e}_{kl})$. □

EC.1.3. Proof of Corollary 1

The proofs of results related to the impact of θ_I^k, θ_C^l p, and q are straightforward and hence, omitted for brevity. We next establish the impact of $\bar{\delta}$ on f_{kl}^* under the centralized model.

Note that

$$\frac{\partial f_{kl}^*}{\partial \bar{\delta}} = \frac{1}{\lambda_f(\lambda_f^2\theta_C^l\theta_I^k - \bar{\delta}^2N^2\lambda_J^2)^3}\Bigg(-\bar{\delta}^6N^6\lambda_J^6 - 2\bar{\delta}^3N^4\lambda^2\lambda_f^2\lambda_J^3\theta_C^l\theta_I^k + 3\bar{\delta}^4N^4\lambda_f^2\lambda_J^4\theta_C^l\theta_I^k - 6\bar{\delta}N^2\lambda^2\lambda_f^4\lambda_J{\theta_C^l}^2{\theta_I^k}^2 + \lambda_f^5{\theta_C^l}^2{\theta_I^k}^2\left(\lambda_f\theta_C^l\theta_I^k - N\lambda^2(\theta_C^l+\theta_I^k)\right) - 3\bar{\delta}^2N^2\lambda_f^3\lambda_J^2\theta_C^l\theta_I^k\left(\lambda_f\theta_C^l\theta_I^k + N\lambda^2(\theta_C^l+\theta_I^k)\right)\Bigg).$$

From Assumption 1, $\lambda_f(\lambda_f^3\theta_C^l\theta_I^k - \bar{\delta}^2N^2\lambda_J^2)^3$ is positive. Denote the numerator in the above expression on the right hand side as $\mathcal{L}_{kl}$. Then $\frac{\partial f_{kl}^*}{\partial \bar{\delta}} = \frac{\mathcal{L}_{kl}}{\lambda_f(\lambda_f^3\theta_C^l\theta_I^k-\bar{\delta}^2N^2\lambda_J^2)^3}$. Further,

$$\frac{\partial \mathcal{L}_{kl}}{\partial \bar{\delta}} = -6N^2\lambda_J\left(\lambda^2\lambda_f^2\theta_C^l\theta_I^k(\bar{\delta}N\lambda_J+\lambda_f\theta_C^l)(\bar{\delta}N\lambda_J+\lambda_f\theta_I^k) + \bar{\delta}\lambda_J(\bar{\delta}^2N^2\lambda_J^2 - \lambda_f^2\theta_I^k\theta_C^l)^2\right) \le 0.$$

Therefore, $\mathcal{L}_{kl}$ is decreasing in $\bar{\delta}$. When $\bar{\delta}=0$, the value of $\mathcal{L}_{kl}|_{\bar{\delta}=0} = \lambda_f^5{\theta_C^l}^2{\theta_I^k}^2(\bar{\delta}N\lambda_J - N\lambda^2(\theta_C^l+\theta_I^k)) \ge 0$. If $\mathcal{L}_{kl} \ge 0$ at $\bar{\delta}=1$, we have $\frac{\partial f_{kl}^*}{\partial \bar{\delta}} \ge 0, \forall \bar{\delta} \in [0,1]$, and f_{kl}^* is increasing in $\bar{\delta}$. Otherwise, f_{kl}^* is first increasing, then decreasing in $\bar{\delta}$ for $\bar{\delta} \in [0,1]$. □

EC.1.4. Proof of Proposition 2

We first make the following claim:

Claim EC.2. *For any value a_{kl}, $k,l \in \{L,H\}$, all IR and IC constraints are satisfied for the coordinator and the sponsor's cost is minimized when IC constraint* (5) *is binding for the low type coordinator and IR constraint* (6) *is binding for the high-type coordinator.*

Proof of Claim EC.2 For any value a_{kl}, consider a solution $\bar{e}_{kl}$ such that $s_{kL} = \Theta_C(\bar{e}_{kH};\theta_C^H) + \Theta_C(\bar{e}_{kL};\theta_C^L) - \Theta_C(\bar{e}_{kH};\theta_C^L)$ and $s_{kH} = \Theta_C(\bar{e}_{kH};\theta_C^H)$. We next show the above values give a feasible solution to the sponsor's problem by verifying (6) is satisfied for $l=L$ and (5) is satisfied for $l=H$.

$$s_{kL} - \Theta_C(\bar{e}_{kL};\theta_C^L) \ge s_{kH} - \Theta_C(\bar{e}_{kH};\theta_C^L) \ge s_{kH} - \Theta_C(\bar{e}_{kH};\theta_C^H) \ge 0,$$

$$
\begin{aligned}
s_{kL} - \Theta_C(\bar{e}_{kL};\theta_C^H) &= \Theta_C(\bar{e}_{kH};\theta_C^H) + \Theta_C(\bar{e}_{kL};\theta_C^L) - \Theta_C(\bar{e}_{kH};\theta_C^L) - \Theta_C(\bar{e}_{kL};\theta_C^H) \\
&= \Theta_C(\bar{e}_{kH};\theta_C^H) - \Theta_C(\bar{e}_{kH};\theta_C^L) - (\Theta_C(\bar{e}_{kL};\theta_C^H) - \Theta_C(\bar{e}_{kL};\theta_C^L)) \\
&\leq 0 = s_{kH} - \Theta_C(\bar{e}_{kH};\theta_C^H).
\end{aligned}
$$

Further, notice that the objective function is increasing in s_{kL}, s_{kH}. Therefore, the sponsor's objective is minimized at $\bar{e}_{kl}$. The result now follows. □

CLAIM EC.3. *Constraint* (4) *is binding in the optimal solution.*

The proof of the above claim follows a similar argument as in Claim EC.1. Combining the above two results, we can rewrite the sponsor's problem as follows for a given k:

$$
\begin{aligned}
\min_{a_{kH},a_{kL},e_{kH},e_{kL}} \quad & \Pi_k^{SI} = q\Theta_I(a_{kH};\theta_I^k) + \bar{q}\Theta_I(a_{kL};\theta_I^k) + q\Theta_C(e_{kH};\theta_C^H) + \bar{q}(\Theta_C(e_{kH};\theta_C^H) + \Theta_C(e_{kL};\theta_C^L) - \Theta_C(e_{kH};\theta_C^L)) \\
& + N\bar{\delta}\left(q\frac{\bar{\delta} - \lambda(a_{kH}+e_{kH}) - \lambda_J a_{kH}e_{kH}}{\lambda_f} + \bar{q}\frac{\bar{\delta} - \lambda(a_{kL}+e_{kL}) - \lambda_J a_{kL}e_{kL}}{\lambda_f}\right), \\
\text{s.t.} \quad & a_{kH}, a_{kL}, e_{kH}, e_{kL} \in [0,1].
\end{aligned}
$$

One can easily verify that the Hessian matrix for the unconstrained optimization problem is positive semidefinite. Let the $\hat{a}_{kl}, \hat{e}_{kl}$ be the solutions of the first order conditions $\frac{\partial \Pi_k^{SI}}{\partial a_{kl}} = 0$ and $\frac{\partial \Pi_k^{SI}}{\partial e_{kl}} = 0$. That is, we have

$$
\begin{aligned}
\hat{a}_{kL} &= \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J + \lambda_f\theta_C^L)}{\lambda_f^2\theta_I^k\theta_C^L - \bar{\delta}^2N^2\lambda_J^2} = a_{kL}^*, & \hat{e}_{kL} &= \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J + \lambda_f\theta_I^k)}{\lambda_f^2\theta_I^k\theta_C^L - \bar{\delta}^2N^2\lambda_J^2} = e_{kL}^*, \\
\hat{a}_{kH} &= \frac{\lambda\bar{\delta}N(q\bar{\delta}N\lambda_J + \lambda_f(\theta_C^H - \bar{q}\theta_C^L))}{\lambda_f^2\theta_I^k(\theta_C^H - \bar{q}\theta_C^L) - q\bar{\delta}^2N^2\lambda_J^2}, & \hat{e}_{kH} &= \frac{q\lambda\bar{\delta}N(\bar{\delta}N\lambda_J + \lambda_f\theta_I^k)}{\lambda_f^2\theta_I^k(\theta_C^H - \bar{q}\theta_C^L) - q\bar{\delta}^2N^2\lambda_J^2}.
\end{aligned}
$$

Using Assumption 1 and Lemma EC.1, we can show that $0 < \hat{a}_{kH}, \hat{e}_{kH} < 1$ and the corresponding $\hat{f}_{kl} = \frac{\bar{\delta} - \lambda(\hat{a}_{kl}+\hat{e}_{kl}) - \lambda_J\hat{a}_{kl}\hat{e}_{kl}}{\lambda_f} > 0$. Hence, the result. □

REMARK EC.1. To simplify expressions, we define $\hat{\theta}_C^L = \theta_C^L$, and $\hat{\theta}_C^H = \theta_C^H + \frac{\bar{q}}{q}(\theta_C^H - \theta_C^L)$. Then, we can write the sponsor's optimal solution under the SI model as follows:

$$
a_{kl}^\circ = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J + \lambda_f\hat{\theta}_C^l)}{\lambda_f^2\theta_I^k\hat{\theta}_C^l - \bar{\delta}^2N^2\lambda_J^2}, \quad e_{kl}^\circ = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J + \lambda_f\theta_I^k)}{\lambda_f^2\theta_I^k\hat{\theta}_C^l - \bar{\delta}^2N^2\lambda_J^2}.
$$

EC.1.5. Proof of Corollaries 2-4 and Propositions 3-4

The proofs of Corollaries 2, 4, and Proposition 4 are straightforward and hence, omitted for brevity. The proof of Proposition 3 is similar to that for Proposition 2 and hence, omitted.

The result in Corollary 3 directly follows from Corollary 1 when the coordinator is low type since the providers' optimal effort levels are the same as those under the centralized optimal solution. When the coordinator is high type, we have

$$
a_{kH}^\circ - e_{kH}^\circ = \frac{\lambda\bar{\delta}N\lambda_f(\theta_C^H - \theta_I^k + \frac{1-q}{q}(\theta_C^H - \theta_C^L))}{(\lambda_f^2\theta_I^k\theta_C^H - \bar{\delta}^2N^2\lambda_J^2) + \frac{1-q}{q}\lambda_f^2\theta_I^k(\theta_C^H - \theta_C^L)}.
$$

Therefore, $a_{kH}^\circ \geq e_{kH}^\circ$ if and only if $\theta_C^H + \frac{1-q}{q}(\theta_C^H - \theta_C^L) \geq \theta_I^k$. □

EC.1.6. Proof of Proposition 5

Note that $(a_{kl}^{\Diamond}, e_{kl}^{\Diamond}, f_{kl}^{\Diamond}), k, l \in \{L, H\}$ provides a feasible solution under the SI model. Further, the compensation $r_{kl}^{\Diamond} \geq \Theta_{I}(a_{kl}^{\Diamond}; \theta_{I}^{k})$ for $k, l \in \{L, H\}$. Hence, we have

$$\Pi_2 = \mathbb{E}_{k,l}\left[\Theta_I(a_{kl}^{\circ}; \theta_I^k) + s_{kl}^{\circ} + \bar{\delta} N f_{kl}^{\circ}\right] \leq \mathbb{E}_{k,l}\left[\Theta_I(a_{kl}^{\Diamond}; \theta_I^k) + s_{kl}^{\Diamond} + \bar{\delta} N f_{kl}^{\Diamond}\right] \leq \mathbb{E}_{k,l}\left[r_{kl}^{\Diamond} + s_{kl}^{\Diamond} + \bar{\delta} N f_{kl}^{\Diamond}\right] = \Pi_3.$$

Similarly, $(a_{kl}^{\circ}, e_{kl}^{\circ}, f_{kl}^{\circ}), k, l \in \{L, H\}$ provides a feasible solution for the centralized model. Further, $s_{kl}^{\circ} \geq \Theta_{C}(e_{kl}^{\circ}; \theta_{C}^{l})$ for $k, l \in \{L, H\}$. Therefore,

$$\Pi_1 = \mathbb{E}_{k,l}\left[\Theta_I(a_{kl}^{*}; \theta_I^k) + \Theta_C(e_{kl}^{*}; \theta_C^l) + \bar{\delta} N f_{kl}^{*}\right] \leq \mathbb{E}_{k,l}\left[\Theta_I(a_{kl}^{\circ}; \theta_I^k) + \Theta_C(e_{kl}^{\circ}; \theta_C^l) + \bar{\delta} N f_{kl}^{\circ}\right]$$

$$\leq \mathbb{E}_{k,l}\left[\Theta_I(a_{kl}^{\circ}; \theta_I^k) + s_{kl}^{\circ} + \bar{\delta} N f_{kl}^{\circ}\right] = \Pi_2.$$

Thus, $\Pi_3 \geq \Pi_2 \geq \Pi_1$.

The result that $P_3 \leq P_2 \leq P_1$ follows from the arguments below:

- The compensation P_2 equals to the value of P_1 by replacing θ_C^H with $\theta_C^H + \frac{\bar{q}}{q}(\theta_C^H - \theta_C^L)$, i.e., $P_2 = P_1|_{\left(\theta_C^H = \theta_C^H + \frac{\bar{q}}{q}(\theta_C^H - \theta_C^L)\right)}$.
- $P_3 = P_2|_{\left(\theta_I^H = \theta_I^H + \frac{\bar{p}}{p}(\theta_I^H - \theta_I^L)\right)} = P_1|_{\left(\theta_I^H = \theta_I^H + \frac{\bar{p}}{p}(\theta_I^H - \theta_I^L), \theta_C^H = \theta_C^H + \frac{\bar{q}}{q}(\theta_C^H - \theta_C^L)\right)}$.
- It is straightforward to observe that P_1 is decreasing in θ_I^H, θ_C^H.

Finally, notice that $F_m = \Pi_m - P_m, i = 1, 2, 3$, the result that $F_3 \geq F_2 \geq F_1$ is now immediate. $\square$

EC.1.7. Proof of Proposition 6

Let $x = \dfrac{\theta_C^L \lambda_f}{\bar{\delta} N \lambda_J}, \hat{y} = \dfrac{\theta_C^H \lambda_f}{\bar{\delta} N \lambda_J}, z = \dfrac{\theta_I^L \lambda_f}{\bar{\delta} N \lambda_J}, \hat{w} = \dfrac{\theta_I^H \lambda_f}{\bar{\delta} N \lambda_J}$, where $\hat{y} \geq x > 2, \hat{w} \geq z > 2$ (from Assumption 1). Further, let $G(u, v) = -\frac{u+v+2}{uv-1}$. Then, we can write Π_1, Π_2, Π_3 as follows:

$$\begin{aligned}
\Pi_1 &= \frac{\bar{\delta}^2 N}{\lambda_f} + \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}\left(pqG(\hat{y}, \hat{w}) + p\bar{q}G(x, \hat{w}) + \bar{p}qG(\hat{y}, z) + \bar{p}\bar{q}G(x, z)\right), \\
\Pi_2 &= \frac{\bar{\delta}^2 N}{\lambda_f} + \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}\left(pqG(\frac{1}{q}\hat{y} - \frac{\bar{q}}{q}x, \hat{w}) + p\bar{q}G(x, \hat{w}) + \bar{p}qG(\frac{1}{q}\hat{y} - \frac{\bar{q}}{q}x, z) + \bar{p}\bar{q}G(x, z)\right), \\
\Pi_3 &= \frac{\bar{\delta}^2 N}{\lambda_f} + \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}\left(pqG(\frac{1}{q}\hat{y} - \frac{\bar{q}}{q}x, \frac{1}{p}\hat{w} - \frac{\bar{p}}{p}z) + p\bar{q}G(x, \frac{1}{p}\hat{w} - \frac{\bar{p}}{p}z) + \bar{p}qG(\frac{1}{q}\hat{y} - \frac{\bar{q}}{q}x, z) + \bar{p}\bar{q}G(x, z)\right).
\end{aligned}$$

We next prove the results for $[\Pi_2 - \Pi_1]$ and $[\Pi_3 - \Pi_1]$.

$$\Pi_2 - \Pi_1 = \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}\left(pq[G(\frac{1}{q}\hat{y} - \frac{\bar{q}}{q}x, \hat{w}) - G(\hat{y}, \hat{w})] + \bar{p}q[G(\frac{1}{q}\hat{y} - \frac{\bar{q}}{q}x, z) - G(\hat{y}, z)]\right).$$

Using first-order derivative, it is easy to verify that $\Pi_2 - \Pi_1$ is decreasing in x, z, $\hat{w}$,λ_f, and increasing in λ, λ_J, $\bar{\delta}$. Therefore, $\Pi_2 - \Pi_1$ is decreasing θ_C^L, θ_I^L, and θ_I^H. Further,

$$\frac{\partial(\Pi_2 - \Pi_1)}{\partial \hat{y}} = \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}(1-q)\left(pq(1+\hat{w})^2 \frac{\bar{q}(q(\hat{w}x-1)^2 - \hat{w}^2(\hat{y}-x)^2)}{(\hat{w}\hat{y}-1)^2(\hat{w}\hat{y}-1-\bar{q}(\hat{w}x-1))^2} + \bar{p}q(1+z)^2 \frac{\bar{q}(q(zx-1)^2 - z^2(\hat{y}-x)^2)}{(z\hat{y}-1)^2(z\hat{y}-1-\bar{q}(zx-1))^2}\right),$$

where the first term within the bracket is negative if and only if $\hat{y} \geq \frac{\hat{w}x - \sqrt{q}}{(1-\sqrt{q})\hat{w}}$, and the second term within the bracket is negative if and only if $\hat{y} \geq \frac{zx - \sqrt{q}}{(1-\sqrt{q})z}$. Note that $\frac{\hat{w}x - \sqrt{q}}{(1-\sqrt{q})\hat{w}} \geq \frac{zx - \sqrt{q}}{(1-\sqrt{q})z}$. Therefore,

$\Pi_2-\Pi_1$ is increasing in $\hat{y}$ when $\hat{y}\le\frac{zx-\sqrt{q}}{(1-\sqrt{q})z}$ and decreasing in $\hat{y}$ when $\hat{y}\ge\frac{\hat{w}x-\sqrt{q}}{(1-\sqrt{q})\hat{w}}$. That is, we have $\Pi_2-\Pi_1$ is increasing in θ_C^H when $\theta_C^H \le \frac{\lambda_f^2\theta_C^L\theta_I^H-\bar{\delta}^2N^2(1-\sqrt{q})}{\bar{\delta}N\lambda_f\lambda_J(1-\sqrt{q})\theta_I^H}$, and is decreasing in θ_C^H when $\theta_C^H\ge\frac{\lambda_f^2\theta_C^H\theta_I^H-\bar{\delta}^2N^2(1-\sqrt{q})}{\bar{\delta}N\lambda_f\lambda_J(1-\sqrt{q})\theta_I^H}$. Hence, the results on the performance gap $\Pi_2-\Pi_1$ follow.

The results under the OM Model follow from similar arguments as above, thus, detailed proofs are avoided for brevity. □

EC.1.8. Proof of Proposition 7

We first simplify the expressions for $\Pi_m, m=1,2,3$ by defining the following values: $x=\dfrac{\theta_C^L\lambda_f}{\bar{\delta}N\lambda_J}, y=\dfrac{\theta_C^H\lambda_f}{\bar{\delta}N\lambda_J}-x, z=\dfrac{\theta_I^L\lambda_f}{\bar{\delta}N\lambda_J}, w=\dfrac{\theta_I^H\lambda_f}{\bar{\delta}N\lambda_J}-z$, where $x,z>2$ and $y,w>0$ from Assumption 1. Next we define $G(u,v)=-\dfrac{u+v+2}{uv-1}$. Further, it is straightforward to show that $G(u,v)$ is submodular in its argument. Then, we can write Π_1,Π_2,Π_3 as follows:

$$\begin{aligned}
\Pi_1 &= \frac{\bar{\delta}^2N}{\lambda_f}+\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\left(pqG(x+y,z+w)+p\bar{q}G(x,z+w)+\bar{p}qG(x+y,z)+\bar{p}\bar{q}G(x,z)\right),\\
\Pi_2 &= \frac{\bar{\delta}^2N}{\lambda_f}+\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\left(pqG(x+\frac{1}{q}y,z+w)+p\bar{q}G(x,z+w)+\bar{p}qG(x+\frac{1}{q}y,z)+\bar{p}\bar{q}G(x,z)\right),\\
\Pi_3 &= \frac{\bar{\delta}^2N}{\lambda_f}+\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\left(pqG(x+\frac{1}{q}y,z+\frac{1}{p}w)+p\bar{q}G(x,z+\frac{1}{p}w)+\bar{p}qG(x+\frac{1}{q}y,z)+\bar{p}\bar{q}G(x,z)\right).
\end{aligned}$$

Next consider $(\Pi_2-\Pi_1)$. Using the above expressions, we have

$$(\Pi_2-\Pi_1)=\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}[pq(G(x+\frac{1}{q}y,z+w)-G(x+y,z+w))+\bar{p}q(G(x+\frac{1}{q}y,z)-G(x+y,z))]$$

Since G is submodular, we have $G(x+\frac{1}{q}y,z+w)-G(x+y,z+w)\le G(x+\frac{1}{q}y,z)-G(x+y,z)$. Hence,

$$\begin{aligned}
(\Pi_2-\Pi_1) &\le \frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}[pq(G(x+\frac{1}{q}y,z)-G(x+y,z))+\bar{p}q(G(x+\frac{1}{q}y,z)-G(x+y,z))]\\
&= \frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\Big[\frac{q\bar{q}y(1+z)^2}{q(xz-1)^2+y^2z^2+qyz(xz-1)}\Big]\\
&\le \frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{\frac{q\bar{q}y(1+z)^2}{(xz-1)^2}}{q+\frac{y^2z^2}{(xz-1)^2}}\le\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{\frac{q\bar{q}y(1+z)^2}{(xz-1)^2}}{q+\frac{y^2z^2}{x^2z^2}}=\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{\frac{q\bar{q}ye_{LL}^2\lambda_J^2}{\lambda^2}}{q+\frac{y^2}{x^2}}\\
&= \frac{\bar{\delta}N\lambda_J}{2\lambda_f}\frac{q\bar{q}ye_{LL}^{*\,2}}{q+\frac{y^2}{x^2}}=\frac{\theta_C^L}{2}\frac{q\bar{q}\frac{y}{x}e_{LL}^{*\,2}}{q+\frac{y^2}{x^2}}\le\frac{q\bar{q}\frac{y}{x}}{q+\frac{y^2}{x^2}}\Pi_1 \qquad (\text{since } \Pi_1>\frac{\theta_C^Le_{LL}^{*\,2}}{2} \text{ and } 0\le q\le 1)\\
&= \frac{q\bar{q}}{q\frac{x}{y}+\frac{y}{x}}\Pi_1\le\frac{q\bar{q}}{2\sqrt{q}}\Pi_1\le\frac{\sqrt{q}(1-q)}{2}\Pi_1\le .193\Pi_1.
\end{aligned}$$

Last inequality follows from the fact that $\frac{\sqrt{q}(1-q)}{2}$ achieves maximum at $q=\frac{1}{\sqrt{3}}$.

Hence, $\frac{\Pi_2}{\Pi_1}\le 1.193$.

Next consider $(\Pi_3 - \Pi_1)$. Note that

$$
\begin{aligned}
(\Pi_3 - \Pi_1) &= \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}\Big[pq(G(x+\frac{1}{q}y, z+\frac{1}{p}w) - G(x+y, z+w)) \\
&\qquad + p\bar{q}(G(x, z+\frac{1}{p}w) - G(x, z+w)) + \bar{p}q(G(x+\frac{1}{q}y, z) - G(x+y, z))\Big] \\
&= \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}\Big[pq(G(x+\frac{1}{q}y, z+\frac{1}{p}w) - G(x+y, z+\frac{1}{p}w) + G(x+y, z+\frac{1}{p}w) - G(x+y, z+w)) \\
&\qquad + p\bar{q}(G(x, z+\frac{1}{p}w) - G(x, z+w)) + \bar{p}q(G(x+\frac{1}{q}y, z) - G(x+y, z))\Big] \\
&\le \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}\Big[pq(G(x+\frac{1}{q}y, z) - G(x+y, z) + G(x, z+\frac{1}{p}w) - G(x, z+w)) \\
&\qquad + p\bar{q}(G(x, z+\frac{1}{p}w) - G(x, z+w)) + \bar{p}q(G(x+\frac{1}{q}y, z) - G(x+y, z))\Big] \\
&= \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f}\Big[q(G(x+\frac{1}{q}y, z) - G(x+y, z)) + p(G(x, z+\frac{1}{p}w) - G(x, z+w))\Big] \\
&\le .386\Pi_1.
\end{aligned}
$$

The first inequality follows from the submodularity of function G, and the last inequality follows from a similar argument as that for the upper bound on Π_2/Π_1. Hence, $\frac{\Pi_3}{\Pi_1} \le 1.386$ □

EC.2. Commonly Observed Compensation Contracts

This section formally states the sponsor's participant retention problems for the SI and the OM models under the FC, LC, and CLC contracts. We first discuss the sponsor's decisions when he adopts the FC contract.

EC.2.1. The FC Contract

Under the SI model, given the investigator's type k, the sponsor provides a fixed compensation s_k to the coordinator and specifies lower bound $\underline{e}_k$ on effort level. Thus, the sponsor's problem is

$$
\begin{aligned}
&\min_{a_k \in [0,1]; s_k, f_k \ge 0} && \mathbb{E}_l[\Theta_{\scriptscriptstyle I}(a_k; \theta_{\scriptscriptstyle I}^k) + s_k + f_k N \delta(a_k, e_{kl}^\circ, f_k)], && && \text{(EC.3)} \\
&\text{s.t.} && \delta(a_k, e_{kl}^\circ, f_k) \ge \bar{\delta}, && \forall l \in \{L, H\}, && \text{(EC.4)} \\
&IR: && s_k - \Theta_{\scriptscriptstyle C}(e_{kl}^\circ; \theta_{\scriptscriptstyle C}^l) \ge 0, && \forall l \in \{L, H\}, && \text{(EC.5)} \\
&IC: && e_{kl}^\circ = \arg\max_{e_{kl} \ge \underline{e}_k} (s_k - \Theta_{\scriptscriptstyle C}(e_{kl}; \theta_{\scriptscriptstyle C}^l)), && \forall l \in \{L, H\}. && \text{(EC.6)}
\end{aligned}
$$

Under the OM model, the sponsor provides fixed compensation r (resp., s) to the investigator (resp., coordinator) and specifies lower bound $\underline{a}$ (resp., $\underline{e}$). Note that a provider's decision of effort is independent of the other provider. Hence, we use a single subscript k (resp., l) to denote the investigator (resp., coordinator) effort given his type. The sponsor's problem is

$$
\min_{r,s,f \ge 0} \quad \mathbb{E}_{k,l}[r + s + f N \delta(a_k^\Diamond, e_l^\Diamond, f)], \tag{EC.7}
$$

$$\text{s.t.} \quad \delta(a_k^\lozenge, e_l^\lozenge, f) \geq \bar{\delta}, \qquad \forall k, l \in \{L, H\}, \tag{EC.8}$$

$$IR: \ r - \Theta_{\scriptscriptstyle I}(a_k^\lozenge; \theta_{\scriptscriptstyle I}^k) \geq 0, \qquad \forall k \in \{L, H\}, \tag{EC.9}$$

$$IC: \ a_k^\lozenge = \arg\max_{a_k \geq \underline{a}} (r - \Theta_{\scriptscriptstyle I}(a_k; \theta_{\scriptscriptstyle I}^k)), \quad \forall k \in \{L, H\}, \tag{EC.10}$$

$$IR: \ s - \Theta_{\scriptscriptstyle C}(e_l^\lozenge; \theta_{\scriptscriptstyle C}^l) \geq 0, \qquad \forall l \in \{L, H\}, \tag{EC.11}$$

$$IC: \ e_l^\lozenge = \arg\max_{e_l \geq \underline{e}} (s - \Theta_{\scriptscriptstyle C}(e_l; \theta_{\scriptscriptstyle C}^l)), \qquad \forall l \in \{L, H\}. \tag{EC.12}$$

EC.2.1.1. Proof of Proposition 8. First, consider the SI model. Under this contract, the coordinator always exerts the lower bound $\underline{e}_k$ irrespective of his type. Therefore, it's optimal for the sponsor to set $s_k = \Theta_{\scriptscriptstyle C}(\underline{e}_k; \theta_{\scriptscriptstyle C}^{\scriptscriptstyle H})$. Further, at optimality, the retention constraint (EC.4) is binding $\forall l \in \{L, H\}$ (the proof is similar to that for Claim EC.1). Therefore, $f_k = \frac{\bar{\delta} - \lambda(a_k + \underline{e}_k) - \lambda_J a_k \underline{e}_k}{\lambda_f}$. Then, the sponsor's problem reduces to the following:

$$\min_{a_k \in [0,1]} \quad \Theta_{\scriptscriptstyle I}(a_k; \theta_{\scriptscriptstyle I}^k) + \Theta_{\scriptscriptstyle C}(\underline{e}_k; \theta_{\scriptscriptstyle C}^{\scriptscriptstyle H}) + \frac{\bar{\delta} - \lambda(a_k + \underline{e}_k) - \lambda_J a_k \underline{e}_k}{\lambda_f} N \bar{\delta}.$$

Note that the objective function is convex in a_k. Using first order conditions and Assumption 1 we have, $a_k^\circ = \frac{\bar{\delta} N(\lambda + \underline{e}_k \lambda_J)}{\lambda_f \theta_I^k}$, and the corresponding $f_k^\circ = \frac{(\bar{\delta} - \underline{e}_k \lambda)\lambda_f \theta_I^k - \bar{\delta} N(\lambda + \lambda_J \underline{e}_k)^2}{\lambda_f^2 \theta_I^k}$, $k \in \{L, H\}$.

Under the OM model, following a similar argument as above we have $r^\lozenge = \frac{1}{2}\theta_{\scriptscriptstyle I}^{\scriptscriptstyle H} \underline{a}^2$, $s^\lozenge = \frac{1}{2}\theta_{\scriptscriptstyle C}^{\scriptscriptstyle H} \underline{e}^2$, $f^\lozenge = \frac{\bar{\delta} - \lambda(\underline{a} + \underline{e}) - \lambda_J \underline{a}\underline{e}}{\lambda_f}$. This completes the proof. $\square$

EC.2.2. Proof of Corollary 5

We first consider the performance under the SI model. Define

$$\mathcal{V}_{kl}(x) = \frac{1}{2}\theta_{\scriptscriptstyle C}^l x^2 + \frac{1}{2}\theta_{\scriptscriptstyle I}^k \left(\frac{\bar{\delta} N(\lambda + \lambda_J x)}{\lambda_f \theta_{\scriptscriptstyle I}^k}\right)^2 + \bar{\delta} N \left(\frac{(\bar{\delta} - \lambda x)\lambda_f \theta_{\scriptscriptstyle I}^k - \bar{\delta} N(\lambda + \lambda_J x)^2}{\lambda_f^2 \theta_{\scriptscriptstyle I}^k}\right).$$

Note that $\Pi_{2k}^F = \Theta_{\scriptscriptstyle I}(a_k^\circ; \theta_{\scriptscriptstyle I}^k) + \Theta_{\scriptscriptstyle C}(e_k^\circ; \theta_{\scriptscriptstyle C}^{\scriptscriptstyle H}) + f_k^\circ N \delta(a_k^\circ, e_k^\circ, f_k^\circ) = \mathcal{V}_{k\scriptscriptstyle H}(\underline{e}_k)$. Further, we can write $\Pi_k^{SI^\circ} = q\mathcal{V}_{k\scriptscriptstyle H}(e_{k\scriptscriptstyle H}^\circ) + \bar{q}\mathcal{V}_{k\scriptscriptstyle L}(e_{k\scriptscriptstyle L}^\circ) + \frac{\bar{q}}{2}(\theta_{\scriptscriptstyle C}^{\scriptscriptstyle H} - \theta_{\scriptscriptstyle C}^{\scriptscriptstyle L}){e_{k\scriptscriptstyle H}^\circ}^2$.

It is easy to verify that $\mathcal{V}_{kl}(x)$ is convex in x. Therefore, for $l \in \{L, H\}$,

$$\mathcal{V}_{kl}(e_{kl}^\circ) \geq (e_{kl}^\circ - \underline{e}_k) \cdot \mathcal{V}_{kl}'(\underline{e}_k) + \mathcal{V}_{kl}(\underline{e}_k).$$

Combining the above two inequalities, we have

$$\begin{aligned}
\Pi_{2k}^F &= \mathcal{V}_{k\scriptscriptstyle H}(\underline{e}_k) = q\mathcal{V}_{k\scriptscriptstyle H}(\underline{e}_k) + \bar{q}\left[\mathcal{V}_{k\scriptscriptstyle L}(\underline{e}_k) + \frac{1}{2}(\theta_{\scriptscriptstyle C}^{\scriptscriptstyle H} - \theta_{\scriptscriptstyle C}^{\scriptscriptstyle L})\underline{e}_k^2\right] \\
&\leq q\left[\mathcal{V}_{k\scriptscriptstyle H}(e_{k\scriptscriptstyle H}^\circ) - (e_{k\scriptscriptstyle H}^\circ - \underline{e}_k) \cdot \mathcal{V}_{k\scriptscriptstyle H}'(\underline{e}_k)\right] + \bar{q}\left[\mathcal{V}_{k\scriptscriptstyle L}(e_{k\scriptscriptstyle L}^\circ) - (e_{k\scriptscriptstyle L}^\circ - \underline{e}_k) \cdot \mathcal{V}_{k\scriptscriptstyle L}'(\underline{e}_k)\right] + \frac{\bar{q}}{2}(\theta_{\scriptscriptstyle C}^{\scriptscriptstyle H} - \theta_{\scriptscriptstyle C}^{\scriptscriptstyle L})\underline{e}_k^2 \\
&= \Pi_k^{SI^\circ} - q(e_{k\scriptscriptstyle H}^\circ - \underline{e}_k) \cdot \mathcal{V}_{k\scriptscriptstyle H}'(\underline{e}_k) - \bar{q}(e_{k\scriptscriptstyle L}^\circ - \underline{e}_k) \cdot \mathcal{V}_{k\scriptscriptstyle L}'(\underline{e}_k) + \frac{\bar{q}}{2}(\theta_{\scriptscriptstyle C}^{\scriptscriptstyle H} - \theta_{\scriptscriptstyle C}^{\scriptscriptstyle L})(\underline{e}_k^2 - {e_{k\scriptscriptstyle H}^\circ}^2) \\
&= \Pi_k^{SI^\circ} + q\left(\theta_{\scriptscriptstyle C}^{\scriptscriptstyle H} - \frac{\bar{\delta}^2 N^2 \lambda_J^2}{\lambda_f^2 \theta_I^k}\right)(\underline{e}_k - e_{k\scriptscriptstyle H}^\circ)(\underline{e}_k - e_{k\scriptscriptstyle H}^*) + \bar{q}\left(\theta_{\scriptscriptstyle C}^{\scriptscriptstyle L} - \frac{\bar{\delta}^2 N^2 \lambda_J^2}{\lambda_f^2 \theta_I^k}\right)(\underline{e}_k - e_{k\scriptscriptstyle L}^\circ)(\underline{e}_k - e_{k\scriptscriptstyle L}^*) + \frac{\bar{q}}{2}(\theta_{\scriptscriptstyle C}^{\scriptscriptstyle H} - \theta_{\scriptscriptstyle C}^{\scriptscriptstyle L})(\underline{e}_k^2 - {e_{k\scriptscriptstyle H}^\circ}^2).
\end{aligned}$$

Therefore,

$$\frac{\Pi^F_{2k}}{\Pi^{SI\circ}_k} \le \frac{\Pi^{SI\circ}_k + q\left(\theta^H_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(\underline{e}_k - e^\circ_{kH})(\underline{e}_k - e^*_{kH}) + \bar q\left(\theta^L_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(\underline{e}_k - e^\circ_{kL})(\underline{e}_k - e^*_{kL}) + \frac{\bar q}{2}(\theta^H_C - \theta^L_C)(\underline{e}_k^2 - {e^\circ_{kH}}^2)}{\Pi^{SI\circ}_k}$$

$$= 1 + \frac{q\left(\theta^H_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(\underline{e}_k - e^\circ_{kH})(\underline{e}_k - e^*_{kH}) + \bar q\left(\theta^L_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(\underline{e}_k - e^\circ_{kL})(\underline{e}_k - e^*_{kL}) + \frac{\bar q}{2}(\theta^H_C - \theta^L_C)(\underline{e}_k^2 - {e^\circ_{kH}}^2)}{\Pi^{SI\circ}_k}$$

$$\le 1 + \frac{q\left(\theta^H_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(\underline{e}_k - e^\circ_{kH})(\underline{e}_k - e^*_{kH}) + \bar q\left(\theta^L_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(\underline{e}_k - e^\circ_{kL})(\underline{e}_k - e^*_{kL}) + \frac{\bar q}{2}(\theta^H_C - \theta^L_C)(\underline{e}_k^2 - {e^\circ_{kH}}^2)}{\frac{1}{2}\theta^L_C {e^\circ_{LL}}^2}$$

$$= 1 + \frac{2}{\theta^L_C {e^\circ_{LL}}^2}\left[q\left(\theta^H_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(\underline{e}_k - e^\circ_{kH})(\underline{e}_k - e^*_{kH}) + \bar q\left(\theta^L_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(\underline{e}_k - e^\circ_{kL})(\underline{e}_k - e^*_{kL})\right] + \bar q\left(\frac{\theta^H_C}{\theta^L_C} - 1\right)\frac{\underline{e}_k^2 - {e^\circ_{kH}}^2}{{e^\circ_{LL}}^2}$$

$$= 1 + \frac{2}{\theta^L_C {e^\circ_{LL}}^2}\mathbb{E}_l\left[\left(\theta^l_C - \frac{\bar\delta^2 N^2\lambda_J^2}{\lambda_f^2\theta^k_I}\right)(e^\circ_{kl} - \underline{e}_k)(e^*_{kl} - \underline{e}_k)\right] + \bar q\left(\frac{\theta^H_C}{\theta^L_C} - 1\right)\frac{\underline{e}_k^2 - {e^\circ_{kH}}^2}{{e^\circ_{LL}}^2}.$$

Next, consider the result for the OM model. Let $\mathcal{U}_{kl}(x,y) = \frac{1}{2}\theta^l_C y^2 + \frac{1}{2}\theta^k_I x^2 + \bar\delta N\frac{\bar\delta - \lambda(x+y) - \lambda_J xy}{\lambda_f}$. Then $\Pi^F_3 = \mathcal{U}_{HH}(\underline{a},\underline{e})$, $\Pi^{OM\Diamond} = \mathbb{E}_{k,l}[\mathcal{U}_{kl}(a^\Diamond_{kl}, e^\Diamond_{kl}) + \frac{1}{2}(\theta^H_C - \theta^l_C){e^\Diamond_{kH}}^2 + \frac{1}{2}(\theta^H_I - \theta^k_I){a^\Diamond_{Hl}}^2]$. It's easy to verify that $\mathcal{U}_{kl}(x,y)$ is a jointly convex function. Therefore,

$$\mathcal{U}_{kl}(\underline{a},\underline{e}) - \mathcal{U}_{kl}(a^\Diamond_{kl}, e^\Diamond_{kl}) \le \nabla\mathcal{U}_{kl}(\underline{a},\underline{e})^\intercal(\underline{a} - a^\Diamond_{kl}, \underline{e} - e^\Diamond_{kl}),$$

where $\nabla\mathcal{U}_{kl}(\underline{a},\underline{e})$ is the gradient of $\mathcal{U}_{kl}$ at $(\underline{a},\underline{e})$. Thus,

$$\begin{aligned}
\Pi^F_3 &= \mathcal{U}_{HH}(\underline{a},\underline{e}) = \mathbb{E}_{k,l}\left[\mathcal{U}_{kl}(\underline{a},\underline{e}) + \frac{1}{2}(\theta^H_C - \theta^l_C)\underline{e}^2 + \frac{1}{2}(\theta^H_I - \theta^k_I)\underline{a}^2\right]\\
&= \mathbb{E}_{k,l}\left[\mathcal{U}_{kl}(\underline{a},\underline{e}) - \mathcal{U}_{kl}(a^\Diamond_{kl}, e^\Diamond_{kl})\right] + \mathbb{E}_{k,l}\left[\frac{1}{2}(\theta^H_C - \theta^l_C)(\underline{e}^2 - {e^\Diamond_{kH}}^2) + \frac{1}{2}(\theta^H_I - \theta^k_I)(\underline{a}^2 - {a^\Diamond_{Hl}}^2)\right] + \Pi^{OM\Diamond}\\
&\le \mathbb{E}_{k,l}\left[\nabla\mathcal{U}_{kl}(\underline{a},\underline{e})^\intercal(\underline{a} - a^\Diamond_{kl}, \underline{e} - e^\Diamond_{kl})\right] + \mathbb{E}_{k,l}\left[\frac{1}{2}(\theta^H_C - \theta^l_C)(\underline{e}^2 - {e^\Diamond_{kH}}^2) + \frac{1}{2}(\theta^H_I - \theta^k_I)(\underline{a}^2 - {a^\Diamond_{Hl}}^2)\right] + \Pi^{OM\Diamond}\\
&= \mathbb{E}_{k,l}\Big[(\underline{a} - a^\Diamond_{kl})(\underline{a}\theta^k_I - \frac{\bar\delta N(\lambda + \lambda_J\underline{e})}{\lambda_f}) + (\underline{e} - e^\Diamond_{kl})(\underline{e}\theta^l_C - \frac{\bar\delta N(\lambda + \lambda_J\underline{a})}{\lambda_f})\\
&\qquad + \frac{1}{2}(\theta^H_C - \theta^l_C)(\underline{e}^2 - {e^\Diamond_{kH}}^2) + \frac{1}{2}(\theta^H_I - \theta^k_I)(\underline{a}^2 - {a^\Diamond_{Hl}}^2)\Big] + \Pi^{OM\Diamond}.
\end{aligned}$$

Further,

$$\begin{aligned}
\frac{\Pi^F_3}{\Pi^{OM\Diamond}} &= 1 + \frac{\mathbb{E}_{k,l}\left[(\underline{a} - a^\Diamond_{kl})(\underline{a}\theta^k_I - \frac{\bar\delta N(\lambda+\lambda_J\underline{e})}{\lambda_f}) + (\underline{e} - e^\Diamond_{kl})(\underline{e}\theta^l_C - \frac{\bar\delta N(\lambda+\lambda_J\underline{a})}{\lambda_f}) + \frac{1}{2}(\theta^H_C - \theta^l_C)(\underline{e}^2 - {e^\Diamond_{kH}}^2) + \frac{1}{2}(\theta^H_I - \theta^k_I)(\underline{a}^2 - {a^\Diamond_{Hl}}^2)\right]}{\Pi^{OM\Diamond}}\\
&\le 1 + \frac{\mathbb{E}_{k,l}\left[(\underline{a} - a^\Diamond_{kl})(\underline{a}\theta^k_I - \frac{\bar\delta N(\lambda+\lambda_J\underline{e})}{\lambda_f}) + \frac{1}{2}(\theta^H_I - \theta^k_I)(\underline{a}^2 - {a^\Diamond_{Hl}}^2)\right]}{\frac{1}{2}\theta^L_C {a^\Diamond_{LL}}^2} + \frac{\mathbb{E}_{k,l}\left[(\underline{e} - e^\Diamond_{kl})(\underline{e}\theta^l_C - \frac{\bar\delta N(\lambda+\lambda_J\underline{a})}{\lambda_f}) + \frac{1}{2}(\theta^H_C - \theta^l_C)(\underline{e}^2 - {e^\Diamond_{kH}}^2)\right]}{\frac{1}{2}\theta^L_C {e^\Diamond_{LL}}^2}\\
&= 1 + \frac{2}{\theta^L_I {a^\Diamond_{LL}}^2}\mathbb{E}_{k,l}\left[(\underline{a} - a^\Diamond_{kl})(\underline{a}\theta^k_I - \frac{\bar\delta N(\lambda+\lambda_J\underline{e})}{\lambda_f}) + \frac{1}{2}(\theta^H_I - \theta^k_I)(\underline{a}^2 - {a^\Diamond_{Hl}}^2)\right]\\
&\quad + \frac{2}{\theta^L_C {e^\Diamond_{LL}}^2}\mathbb{E}_{k,l}\left[(\underline{e} - e^\Diamond_{kl})(\underline{e}\theta^l_C - \frac{\bar\delta N(\lambda+\lambda_J\underline{a})}{\lambda_f}) + \frac{1}{2}(\theta^H_C - \theta^l_C)(\underline{e}^2 - {e^\Diamond_{kH}}^2)\right].
\end{aligned}$$

$\square$

EC.2.3. The LC Contract

Under the SI model, the sponsor's problem is

$$\begin{aligned}
\min_{a_k\in[0,1];f_k,\beta_k\geq 0} \quad & \mathbb{E}_l[\Theta_{\scriptscriptstyle I}(a_k;\theta_{\scriptscriptstyle I}^k)+\beta_k e_{kl}^\circ + f_k N\delta(a_k,e_{kl}^\circ,f_k)] && && \text{(EC.13)}\\
\text{s.t.} \quad & \delta(a_k,e_{kl}^\circ,f_k)\geq\bar{\delta}, && \forall l\in\{L,H\}, && \text{(EC.14)}\\
IR: \quad & \beta_k e_{kl}^\circ - \Theta_{\scriptscriptstyle C}(e_{kl}^\circ;\theta_{\scriptscriptstyle C}^l)\geq 0, && \forall l\in\{L,H\}, && \text{(EC.15)}\\
IC: \quad & e_{kl}^\circ\in\arg\max_e\left\{\beta_k e-\Theta_{\scriptscriptstyle C}(e;\theta_{\scriptscriptstyle C}^l)\right\}, && \forall l\in\{L,H\}. && \text{(EC.16)}
\end{aligned}$$

Note that under the OM model with LC contract, a provider's decision of effort is independent of the other provider. Hence, we use a single subscript k (resp., l) to denote the investigator (resp., coordinator) effort given his type. Then, the sponsor's problem is:

$$\begin{aligned}
\min_{\nu,\beta,f\geq 0} \quad & \mathbb{E}_{k,l}[\nu a_k^\Diamond+\beta e_l^\Diamond+fN\delta(a_k^\Diamond,e_l^\Diamond,f)], && && \text{(EC.17)}\\
\text{s.t.} \quad & \delta(a_k^\Diamond,e_l^\Diamond,f)\geq\bar{\delta}, && \forall k,l\in\{L,H\}, && \text{(EC.18)}\\
IR: \quad & \nu a_k^\Diamond-\theta_{\scriptscriptstyle I}^k(a_k^\Diamond;\theta_{\scriptscriptstyle I}^k)\geq 0, && \forall k\in\{L,H\}, && \text{(EC.19)}\\
IC: \quad & a_k^\Diamond\in\arg\max_a\left\{\nu a-\Theta_{\scriptscriptstyle I}(a;\theta_{\scriptscriptstyle I}^k)\right\}, && \forall k\in\{L,H\}, && \text{(EC.20)}\\
IR: \quad & \beta e_l^\Diamond-\Theta_{\scriptscriptstyle C}(e_l^\Diamond;\theta_{\scriptscriptstyle C}^l)\geq 0, && \forall l\in\{L,H\}, && \text{(EC.21)}\\
IC: \quad & e_l^\Diamond\in\arg\max_e\left\{\beta e-\Theta_{\scriptscriptstyle C}(e;\theta_{\scriptscriptstyle C}^l)\right\}, && \forall l\in\{L,H\}. && \text{(EC.22)}
\end{aligned}$$

EC.2.3.1. Proof of Proposition 9. Under the SI model, it is straightforward to derive that for any given β_k, the coordinator's optimal effort $e_{kl}^\circ=\frac{\beta_k}{\theta_C^l}, l\in\{L,H\}$. At optimality, we can show that retention constraint (EC.14) is binding for $l=H$ and hence, satisfied for $l=L$. Thus, we have $f_k^\circ=\frac{\bar{\delta}-\lambda(a_k^\circ+e_{kH}^\circ)-\lambda_J a_k^\circ e_{kH}^\circ}{\lambda_f}$. Now it is easy to verify that the sponsor's problem reduces to Problem P_{LSI} below:

Problem P_{LSI}

$$\min_{a_k\in[0,1];\beta_k\geq 0} \quad \mathbb{E}_l[\Theta_{\scriptscriptstyle I}(a_k;\theta_{\scriptscriptstyle I}^k)+\beta_k e_{kl}^\circ+f_k^\circ N\delta(a_k,e_{kl}^\circ,f_k^\circ)].$$

Following similar arguments as above, under the OM model we have (i) for a given β (resp., ν), the coordinator's (resp., investigator's) optimal effort $e_l^\Diamond=\frac{\beta}{\theta_C^l}, l\in\{L,H\}$ (resp., $a_k^\Diamond=\frac{\nu}{\theta_I^k}, k\in\{L,H\}$), and (ii) at optimality retention constraint (EC.18) is binding for $k=l=H$. Thus, the sponsor's problem reduces to the Problem P_{LOM} below:

Problem P_{LOM}

$$\min_{\nu\geq 0,\beta\geq 0} \quad \mathbb{E}_{k,l}[\nu a_k^\Diamond+\beta e_l^\Diamond+f^\Diamond N\delta(a_k^\Diamond,e_l^\Diamond,f^\Diamond)].$$

□

EC.2.3.2. Proof of Corollary 6. When both providers have a single type, we have $\theta_I^L=\theta_I^H=\theta_I, \theta_C^L=\theta_C^H=\theta_C$. First, consider the SI model. As mentioned in the proof of Proposition 9, for a given β value, the coordinator's optimal effort $e^\circ=\frac{\beta}{\theta_C}$, and the corresponding compensation is $\frac{\beta^2}{\theta_C}$. Therefore, the sponsor's problem is

$$\min_{a\in[0,1];f,0\le\beta\le\theta_C} \quad [\Theta_I(a;\theta_I)+\frac{\beta^2}{\theta_C}+fN\delta(a,\frac{\beta}{\theta_C},f)],$$
$$\text{s.t.} \quad \delta(a,\frac{\beta}{\theta_C},f)\ge\bar{\delta}.$$

We can rewrite the objective function as $[\Theta_I(a;\theta_I)+\frac{1}{2}2\theta_C(\frac{\beta}{\theta_C})^2+fN\delta(a,\frac{\beta}{\theta_C},f)]$. This is equivalent to the centralized model when the coordinator's effort cost parameter is $2\theta_C$ and decision variables are a and $e=\frac{\beta}{\theta_C}$. Hence, from Proposition 1, we have
$a^\circ=\frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J+2\lambda_f\theta_C)}{\lambda_f^2 2\theta_C\theta_I-\bar{\delta}^2N^2\lambda_J^2}$, $\frac{\beta^\circ}{\theta_C} = \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J+\lambda_f\theta_I)}{2\lambda_f^2\theta_C\theta_I-\bar{\delta}^2N^2\lambda_J^2}$, and $f^\circ=\frac{\bar{\delta}-\lambda(a^\circ+\frac{\beta^\circ}{\theta_C})-\lambda_J a^\circ\frac{\beta^\circ}{\theta_C}}{\lambda_f}$.

Since the retention constraint is binding, the optimal objective value is:

$$\Pi_2^N = [\frac{1}{2}\theta_I a^{\circ 2}+\frac{1}{2}2\theta_C(\frac{\beta^\circ}{\theta_C})^2+f^\circ N\bar{\delta}].$$

Defining $x=\dfrac{\theta_C\lambda_f}{\bar{\delta}N\lambda_J}, z=\dfrac{\theta_I\lambda_f}{\bar{\delta}N\lambda_J}$, where $x,z>2$ from Assumption 1. Further, let $G(u,v)=-\frac{u+v+2}{uv-1}$, where $G(u,v)$ is submodular in its arguments. Then the corresponding cost can be expressed by

$$\Pi_2^N=\frac{\bar{\delta}^2N}{\lambda_f}+\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}G(2x,z),$$
$$\Pi^{SI\circ}=\frac{\bar{\delta}^2N}{\lambda_f}+\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}G(x,z).$$

Then

$$\begin{aligned}\frac{\Pi_2^N-\Pi^{SI\circ}}{\Pi^{SI\circ}} &= \frac{1}{\Pi^{SI\circ}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}(G(2x,z)-G(x,z)) = \frac{1}{\Pi^{SI\circ}}\frac{\theta_C\lambda^2}{2x\lambda_J^2\lambda_f}\frac{x(1+z)^2}{(xz-1)(2xz-1)} \\ &< \frac{1}{\frac{1}{2}\theta_C e^{\circ 2}}\frac{\theta_C\lambda^2}{2\lambda_J^2\lambda_f}\frac{(1+z)^2}{2(xz-1)^2}=\frac{1}{2} \qquad (\text{since } e^\circ=\frac{\lambda(z+1)}{\lambda_J(xz-1)} \text{ and } (2xz-1)>(xz-1)).\end{aligned}$$

Therefore, $\frac{\Pi_2^N}{\Pi^{SI\circ}}<\frac{3}{2}$. Similarly, under the OM model, we have

$$\Pi_3^N=\frac{\bar{\delta}^2N}{\lambda_f}+\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}G(2x,2z),$$
$$\Pi^{OM\diamond}=\frac{\bar{\delta}^2N}{\lambda_f}+\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}G(x,z).$$

Then we have,

$$\begin{aligned}\frac{\Pi_3^N-\Pi^{OM\diamond}}{\Pi^{OM\diamond}} &= \frac{1}{\Pi^{OM\diamond}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}(G(2x,2z)-G(x,z)) = \frac{1}{\Pi^{OM\diamond}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}(G(2x,2z)-G(2x,z)+G(2x,z)-G(x,z)) \\ &\le \frac{1}{\Pi^{OM\diamond}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}(G(x,2z)-G(x,z)+G(2x,z)-G(x,z)) = \frac{1}{\Pi^{OM\diamond}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}(\frac{x(1+z)^2+z(1+x)^2}{(xz-1)(2xz-1)})\end{aligned}$$

$$
\begin{aligned}
&< \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{x(1+z)^2+z(1+x)^2}{(xz-1)^2} = \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2(xe^{\Diamond^2}+za^{\Diamond^2})}{2\lambda^2}\\
&= \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2 xe^{\Diamond^2}}{2\lambda^2} + \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2 za^{\Diamond^2}}{2\lambda^2}\\
&\le \frac{1}{\frac{1}{2}\theta_C e^{\Diamond^2}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2 xe^{\Diamond^2}}{2\lambda^2} + \frac{1}{\frac{1}{2}\theta_I a^{\Diamond^2}}\frac{\bar{\delta}N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2 za^{\Diamond^2}}{2\lambda^2} = 1.
\end{aligned}
$$

Therefore, $\frac{\Pi_3^N}{\Pi^{OM^\Diamond}} < 2$. The result now follows. □

EC.2.4. The CLC Contract

Under the SI model, the sponsor's problem for given $k \in \{L, H\}$ is

$$
\begin{aligned}
\min_{a_k\in[0,1];f_k,\beta_k\ge 0}\quad & \mathbb{E}_l[\Theta_I(a_k;\theta_I^k) + \beta_k e_{kl}^\circ + f_k N\delta(a_k, e_{kl}^\circ, f_k)], && && \text{(EC.23)}\\
\text{s.t.}\quad & \delta(a_k, e_{kl}^\circ, f_k) \ge \bar{\delta}, && \forall l \in \{L,H\}, && \text{(EC.24)}\\
IR:\quad & \beta_k e_{kl}^\circ - \Theta_C(e_{kl}^\circ;\theta_C^l) \ge 0, && \forall l \in \{L,H\}, && \text{(EC.25)}\\
IC:\quad & e_{kl}^\circ \in \arg\max_{e\ge \underline{e}_k}\left\{\beta_k e - \Theta_C(e;\theta_C^l)\right\}, && \forall l \in \{L,H\}. && \text{(EC.26)}
\end{aligned}
$$

Under the OM model, the sponsor's problem is:

$$
\begin{aligned}
\min_{\nu,\beta,f\ge 0}\quad & \mathbb{E}_{k,l}[\nu a_k^\Diamond + \beta e_l^\Diamond + fN\delta(a_k^\Diamond, e_l^\Diamond, f)], && && \text{(EC.27)}\\
\text{s.t.}\quad & \delta(a_k^\Diamond, e_l^\Diamond, f) \ge \bar{\delta}, && \forall k,l \in \{L,H\}, && \text{(EC.28)}\\
IR:\quad & \nu a_k^\Diamond - \theta_I^k(a_k^\Diamond;\theta_I^k) \ge 0, && \forall k \in \{L,H\}, && \text{(EC.29)}\\
IC:\quad & a_k^\Diamond \in \arg\max_{a\ge \underline{a}}\left\{\nu a - \Theta_I(a;\theta_I^k)\right\}, && \forall k \in \{L,H\}, && \text{(EC.30)}\\
IR:\quad & \beta e_l^\Diamond - \Theta_C(e_l^\Diamond;\theta_C^l) \ge 0, && \forall l \in \{L,H\}, && \text{(EC.31)}\\
IC:\quad & e_l^\Diamond \in \arg\max_{e\ge \underline{e}}\left\{\beta e - \Theta_C(e;\theta_C^l)\right\}, && \forall l \in \{L,H\}. && \text{(EC.32)}
\end{aligned}
$$

EC.2.4.1. Proof of Proposition 10. Under the SI model, it is straightforward to derive that for any given β_k, the coordinator's optimal effort $e_{kl}^\circ = \max\left\{\underline{e}_k, \frac{\beta_k}{\theta_C^l}\right\}, l \in \{L, H\}$. From the IR constraint, we have $\beta_k \ge \frac{\theta_C^H \underline{e}_k}{2}$. At optimality, we can show that retention constraint (EC.24) is binding for $l = H$ and hence, satisfied for $l = L$. Thus, we have $f_k^\circ = \frac{\bar{\delta} - \lambda(a_k^\circ + e_{kH}^\circ) - \lambda_J a_k^\circ e_{kH}^\circ}{\lambda_f}$. Now it is easy to verify that the sponsor's problem reduces to PROBLEM P_{CLSI} below:

PROBLEM P_{CLSI}

$$
\min_{a_k\in[0,1];\beta_k\ge\frac{\theta_C^H \underline{e}_k}{2}} \quad \mathbb{E}_l[\Theta_I(a_k;\theta_I^k) + \beta_k e_{kl}^\circ + f_k^\circ N\delta(a_k, e_{kl}^\circ, f_k^\circ)].
$$

Following similar arguments as above, under the OM model we have (i) for a given β (resp., ν), the coordinator's (resp., investigator's) optimal effort $e_l^\Diamond = \max\left\{\underline{e}, \frac{\beta}{\theta_C^l}\right\}, l \in \{L, H\}$ (resp.,

$a_k^{\diamondsuit} = \max\left\{\underline{a}, \frac{\nu}{\theta_I^k}\right\}, k \in \{L,H\})$, (ii) $\beta \geq \frac{\theta_C^H \underline{e}}{2}$ and $\nu \geq \frac{\theta_I^H \underline{a}}{2}$ and (iii) at optimality retention constraint (EC.28) is binding for $k=l=H$. Hence, $f^{\diamondsuit} = \frac{\bar{\delta} - \lambda(a_H^{\diamondsuit} + e_H^{\diamondsuit}) - \lambda_J a_H^{\diamondsuit} e_H^{\diamondsuit}}{\lambda_f}$. Thus, the sponsor's problem reduces to the Problem P_{CLOM} below:

Problem P_{CLOM}

$$\min_{\nu \geq \frac{\theta_I^H \underline{a}}{2}, \beta \geq \frac{\theta_C^H \underline{e}}{2}} \quad \mathbb{E}_{k,l}[\nu a_k^{\diamondsuit} + \beta e_l^{\diamondsuit} + f^{\diamondsuit} N \delta(a_k^{\diamondsuit}, e_l^{\diamondsuit}, f^{\diamondsuit})].$$

□

EC.2.5. Proof of Proposition 11

Before proving the results in this proposition, we establish the following result:

Lemma EC.2. *$\Pi_{2k}^{\hat{N}}(\underline{e}_k) \geq \min\{\Pi_{2k}^F(\underline{e}_k), \Pi_{2k}^N\}$ under the SI model.*

Proof Under the SI model, given the investigator's type k, the sponsor's problem under the conditional linear contract, is

$$\min_{a_k, f_k, \beta_k \geq 0} \quad \mathbb{E}_l[\Theta_I(a_k;\theta_I^k) + \beta_k e_{kl}^{\circ} + f_k N \delta(a_k, e_{kl}^{\circ}, f_k)], \tag{EC.33}$$

$$\text{s.t.} \quad \delta(a_k, e_{kl}^{\circ}, f_k) \geq \bar{\delta}, \qquad \forall l \in \{L,H\}, \tag{EC.34}$$

$$IR: \ \beta_k \hat{e}_{kl}^{\circ} - \Theta_C(\hat{e}_{kl}^{\circ};\theta_C^l) \geq 0, \qquad \forall l \in \{L,H\}, \tag{EC.35}$$

$$IC: \ \hat{e}_{kl}^{\circ} \in \arg\max_{e \geq \underline{e}_k} \left\{\beta_k e - \Theta_C(e;\theta_C^l)\right\}, \qquad \forall l \in \{L,H\}. \tag{EC.36}$$

From IC constraint, we get $\hat{e}_{kl}^{\circ} = \max\{\underline{e}_k, \frac{\beta_k}{\theta_C^l}\}, k,l \in \{L,H\}$. For a given $\underline{e}_k$, let $(\hat{\beta}_k^{\circ}, \hat{a}_k^{\circ}, \hat{f}_k^{\circ})$ be the sponsor's optimal decision under the conditional linear contract. Further, given a value of β_k, from the proof of Proposition 9, the coordinator's optimal effort under the LC contract is $\frac{\beta_k}{\theta_C^l}, k,l \in \{L,H\}$. We next consider the following two possibilities: (i) $\underline{e}_k \leq \frac{\hat{\beta}_k^{\circ}}{\theta_C^H}$ and (ii) $\underline{e}_k > \frac{\hat{\beta}_k^{\circ}}{\theta_C^H}$.

- Suppose $\underline{e}_k \leq \frac{\hat{\beta}_k^{\circ}}{\theta_C^H}$. Then, we have $\hat{e}_{kl}^{\circ} = \frac{\hat{\beta}_k^{\circ}}{\theta_C^l}, k,l \in \{L,H\}$. Hence, the coordinator's decision under the LC contract with parameter $\beta_k = \hat{\beta}_k^{\circ}$ is the same as that under the CLC contract. Therefore, the sponsor can achieve the same retention rate with the same retention cost under the LC contract by choosing $\beta_k = \hat{\beta}_k^{\circ}, a_k = \hat{a}_k^{\circ}, f_k = \hat{f}_k^{\circ}$. Thus, $\hat{\beta}_k^{\circ}$ is feasible for the sponsor's problem under the LC contract implying $\Pi_{2k}^{\hat{N}}(\underline{e}_k) \geq \Pi_{2k}^N \geq \min\{\Pi_{2k}^F(\underline{e}_k), \Pi_{2k}^N\}$.
- Suppose $\underline{e}_k > \frac{\hat{\beta}_k^{\circ}}{\theta_C^H}$. This implies $\hat{\beta}_k^{\circ} < \theta_C^H \underline{e}_k$ and $e_{kH}^{\circ} = \underline{e}_k$. Then, consider a fixed contract with a lower bound on the effort as $\underline{e}_k$ and the fixed compensation as $\frac{1}{2}\theta_C^H \underline{e}_k^2$. Then, we can easily show that the sponsor can achieve the target retention rate by setting $a_k = \hat{a}_k^{\circ}, f_k = \hat{f}_k^{\circ}$ under this fixed contract. Thus,

$$
\begin{aligned}
\Pi_{2k}^{F}(\underline{e}_k) =& \frac{1}{2}\theta_C^H \underline{e}_k^2 + \Theta_I(\hat{a}_k^\circ) + \hat{f}_k^\circ N\delta(\underline{e}_k, \hat{a}_k^\circ, \hat{f}_k^\circ) \\
\leq& \hat{\beta}_k^\circ \underline{e}_k + \Theta_I(\hat{a}_k^\circ) + q\hat{f}_k^\circ N\delta(\underline{e}_k, \hat{a}_k^\circ, \hat{f}_k^\circ) + (1-q)\hat{f}_k^\circ N\delta\left(\max\left\{\underline{e}_k, \frac{\hat{\beta}_k^\circ}{\theta_C^L}\right\}, \hat{a}_k^\circ, \hat{f}_k^\circ\right) \\
\leq& q\hat{\beta}_k^\circ \underline{e}_k + (1-q)\hat{\beta}_k^\circ \max\left\{\underline{e}_k, \frac{\hat{\beta}_k^\circ}{\theta_C^L}\right\} + \Theta_I(\hat{a}_k^\circ) + q\hat{f}_k^\circ N\delta(\underline{e}_k, \hat{a}_k^\circ, \hat{f}_k^\circ) + (1-q)\hat{f}_k^\circ N\delta\left(\max\left\{\underline{e}_k, \frac{\hat{\beta}_k^\circ}{\theta_C^L}\right\}, \hat{a}_k^\circ, \hat{f}_k^\circ\right) \\
=& \Pi_{2k}^{\hat{N}}(\underline{e}_k).
\end{aligned}
$$

Thus, $\Pi_{2k}^{\hat{N}}(\underline{e}_k) \geq \Pi_{2k}^{F}(\underline{e}_k) \geq \min\{\Pi_{2k}^{F}(\underline{e}_k), \Pi_{2k}^{N}\}$.

Therefore, the relationship $\Pi_{2k}^{\hat{N}}(\underline{e}_k) \geq \min\{\Pi_{2k}^{F}(\underline{e}_k), \Pi_{2k}^{N}\}$ always hold. □

Note that using the results of Proposition 8, it is easy to prove that $\Pi_{2k}^{F}(\underline{e}_k)$ is convex in $\underline{e}_k$. Since Π_{2k}^{N} does not change with $\underline{e}_k$, $\Pi_{2k}^{F}(\underline{e}_k)$ and Π_{2k}^{N} may intersect at the most twice in range $[0,1]$. Observe that for $\underline{e}_k = 0$, $\Pi_{2k}^{N} \leq \Pi_{2k}^{F}(\underline{e}_k = 0)$. Hence, there always exists at least one intersection point, say, $e_1 \geq 0$, such that $\Pi_{2k}^{N} \leq \Pi_{2k}^{F}(\underline{e}_k)$ if $\underline{e}_k \in [0, e_1]$. For our proof, we further assume that there are two intersections and hence, there exists $e_2 \in [0,1]$ such that $e_2 > e_1$. Proof under the other possibility where $e_2 > 1$ follows from similar arguments. Note that $\Pi_{2k}^{F}(\underline{e}_k) < \Pi_{2k}^{N}$ if and only if $\underline{e}_k \in (e_1, e_2)$.

Next, we prove (a) $\Pi_{2k}^{\hat{N}}(\underline{e}_k) = \Pi_{2k}^{N} \leq \Pi_{2k}^{F}(\underline{e}_k)$ for $\underline{e}_k \in [0, e_1]$, (b) $\Pi_{2k}^{\hat{N}}(\underline{e}_k) \geq \Pi_{2k}^{F}(\underline{e}_k)$ for $\underline{e}_k \geq e_1$ and (c) when $\theta_C^H \leq 2\theta_C^L$, $\Pi_{2k}^{\hat{N}}(\underline{e}_k) = \Pi_{2k}^{F}(\underline{e}_k)$ $\forall \underline{e}_k \in [e_1, 1]$.

(a) $\boldsymbol{\underline{e}_k \in [0, e_1]}$: Let the sponsor's optimal solution under the LC contract be $(\beta_k^\circ, a_k^\circ, f_k^\circ)$, and the corresponding optimal effort for the l-type coordinator be e_{kl}°. Note that from the proof of Proposition 9, $e_{kl}^\circ = \frac{\beta_k^\circ}{\theta_C^l}$.

Consider the optimal expected retention cost with the LC contract:

$$
\begin{aligned}
\Pi_{2k}^{N} =& \Theta_I(a_k^\circ; \theta_I^k) + q\beta_k^\circ e_{kH}^\circ + (1-q)\beta_k^\circ e_{kL}^\circ + qf_k^\circ N\delta(a_k^\circ, e_{kH}^\circ, f_k^\circ) + (1-q)f_k^\circ N\delta(a_k^\circ, e_{kL}^\circ, f_k^\circ) \\
\geq& \Theta_I(a_k^\circ; \theta_I^k) + \beta_k^\circ e_{kH}^\circ + f_k^\circ N\delta(a_k^\circ, e_{kH}^\circ, f_k^\circ).
\end{aligned}
$$

Note that the last inequality is equal to the expected retention cost for an FC contract with a lower bound of effort as e_{kH}°, the sponsor's decisions $(s_k = \frac{\beta_k^{\circ 2}}{\theta_C^H}, a_k = a_k^\circ, f_k = f_k^\circ)$. The coordinator's decision given this contract is e_{kH}°. Further, it is straightforward to verify that the IR and IC constraints are satisfied for the coordinator and that the retention constraint is also satisfied. Hence, the above decisions form a feasible solution under the FC contract. Thus, we must have $\Pi_{2k}^{N} \geq \Pi_{2k}^{F}(e_{kH}^\circ)$. Recall that $\Pi_{2k}^{N} \geq \Pi_{2k}^{F}(\underline{e}_k)$ if and only if $\underline{e}_k \in [e_1, e_2]$. Therefore, we must have $e_{kH}^\circ \geq e_1$.

Now consider an CLC contract with sponsor's decisions $(\beta_k^\circ, a_k^\circ, f_k^\circ)$, and lower bound on effort $\underline{e}_k$. Recall $\hat{e}_{kl}^\circ = \max\{\frac{\beta_k^\circ}{\theta_C^l}, \underline{e}_k\}$. From the arguments above $\hat{e}_{kl}^\circ = \frac{\beta_k^\circ}{\theta_C^l} = e_{kl}^\circ$. Further, it is straightforward to verify that the IR and IC constraints are satisfied for the coordinator and that the retention constraint is also satisfied. Hence, the above decisions form a feasible solution under the CLC contract. Thus, we must have $\Pi_{2k}^{\hat{N}}(\underline{e}_k) \leq \Pi_{2k}^{N}$. Combining with Lemma EC.2, we have $\Pi_{2k}^{\hat{N}}(\underline{e}_k) = \Pi_{2k}^{N}$ when $\underline{e}_k \leq e_1$. Hence, the result $\Pi_{2k}^{\hat{N}}(\underline{e}_k) = \Pi_{2k}^{N} \leq \Pi_{2k}^{F}(\underline{e}_k)$ follows.

(b) $\boldsymbol{\underline{e}_k \geq e_1}$**:** We consider two possibilities: (i) $e_1 \leq \underline{e}_k \leq e_2$ and (ii) $\underline{e}_k \geq e_2$.

(i) $e_1 \leq \underline{e}_k \leq e_2$: In this region, we have $\min\{\Pi_{2k}^{F}(\underline{e}_k), \Pi_{2k}^{N}\} = \Pi_{2k}^{F}(\underline{e}_k)$. Therefore, from Lemma EC.2, $\Pi_{2k}^{\hat{N}}(\underline{e}_k) \geq \min\{\Pi_{2k}^{F}(\underline{e}_k), \Pi_{2k}^{N}\} = \Pi_{2k}^{F}(\underline{e}_k)$.

(ii) $\underline{e}_k \geq e_2$: Let $(\hat{\beta}_k^{\circ}, \hat{a}_k^{\circ}, \hat{f}_k^{\circ})$ be the sponsor's optimal decision under the CLC contract with lower bound $\underline{e}_k \in [e_2, 1]$. Then, the coordinator's optimal effort level $\hat{e}_{kl}^{\circ} = \max\left\{\frac{\hat{\beta}_k^{\circ}}{\theta_C^l}, \underline{e}_k\right\}$. Note that from the IR constraint of the high-type coordinator, we have $\hat{\beta}_k^{\circ}\hat{e}_{kH}^{\circ} - \frac{1}{2}\theta_C^H {\hat{e}_{kH}^{\circ}}^2 \geq 0$. The sponsor's optimal expected retention cost given the CLC contract is

$$\begin{aligned}
\Pi_{2k}^{\hat{N}}(\underline{e}_k) =& \Theta_I(\hat{a}_k^{\circ}) + q\hat{\beta}_k^{\circ}\hat{e}_{kH}^{\circ} + (1-q)\hat{\beta}_k^{\circ}\hat{e}_{kL}^{\circ} + q\hat{f}_k^{\circ}N\delta(\hat{a}_k^{\circ}, \hat{e}_{kH}^{\circ}, \hat{f}_k^{\circ}) + (1-q)\hat{f}_k^{\circ}N\delta\left(\hat{a}_k^{\circ}, \hat{e}_{kL}^{\circ}, \hat{f}_k^{\circ}\right) \\
\geq& \Theta_I(\hat{a}_k^{\circ}) + \hat{\beta}_k^{\circ}\hat{e}_{kH}^{\circ} + \hat{f}_k^{\circ}N\delta(\hat{a}_k^{\circ}, \hat{e}_{kH}^{\circ}, \hat{f}_k^{\circ}) \\
\geq& \Theta_I(\hat{a}_k^{\circ}) + \frac{1}{2}\theta_C^H {\hat{e}_{kH}^{\circ}}^2 + \hat{f}_k^{\circ}N\delta(\hat{a}_k^{\circ}, \hat{e}_{kH}^{\circ}, \hat{f}_k^{\circ}),
\end{aligned}$$

where the last line is the sponsor's expected retention cost under an FC contract with the lower bound on effort $\hat{e}_{kH}^{\circ}$, payment $s_k = \frac{1}{2}\theta_C^H {\hat{e}_{kH}^{\circ}}^2$, and $a_k = \hat{a}_k^{\circ}, f_k = \hat{f}_k^{\circ}$. It is straightforward to establish that all the constraints are satisfied, and the solution is feasible under the FC contract. Therefore, $\Pi_{2k}^{F}(\hat{e}_{kH}^{\circ}) \leq \Pi_{2k}^{\hat{N}}(\underline{e}_k)$. Further, $\hat{e}_{kH}^{\circ} = \max\left\{\frac{\hat{\beta}_k^{\circ}}{\theta_C^H}, \underline{e}_k\right\} \geq \underline{e}_k \geq e_2$. $\Pi_{2k}^{F}(.)$ is increasing in its argument in the interval $[e_2, 1]$. Hence, $\Pi_{2k}^{F}(\hat{e}_{kH}^{\circ}) \geq \Pi_{2k}^{F}(\underline{e}_k)$ implying $\Pi_{2k}^{F}(\underline{e}_k) \leq \Pi_{2k}^{\hat{N}}(\underline{e}_k)$.

(c) $\boldsymbol{\underline{e}_k \in [e_1, 1]}$ **and** $\boldsymbol{\theta_C^H \leq 2\theta_C^L}$**:** From (b), we have $\Pi_{2k}^{F}(\underline{e}_k) \leq \Pi_{2k}^{\hat{N}}(\underline{e}_k)$. Hence, to show the result here, it suffices to prove that $\Pi_{2k}^{\hat{N}}(\underline{e}_k) \leq \Pi_{2k}^{F}(\underline{e}_k)$, which we derive below.

Given $\underline{e}_k$, let $(\bar{s}_k^{\circ}, \bar{a}_k^{\circ}, \bar{f}_k^{\circ})$ be the sponsor's optimal solution under the FC contract. Recall that from Proposition 8, the optimal compensation under the FC contract $\bar{s}_k^{\circ} = \frac{1}{2}\theta_C^H \underline{e}_k^2$. Consider the following solution under the CLC contract with lower bound on effort $\underline{e}_k$, $\beta_k = \frac{1}{2}\theta_C^H \underline{e}_k, a_k = \bar{a}_k^{\circ}$, and $f_k = \bar{f}_k^{\circ}$. Then, the optimal effort level that maximizes a l-type coordinator's benefit is $\max\left\{\underline{e}_k, \frac{\beta_k}{\theta_C^l}\right\} = \max\left\{\underline{e}_k, \frac{\theta_C^H}{2\theta_C^l}\underline{e}_k\right\} = \underline{e}_k$, which is the same as that under the FC contract, and the IR constraints under the CLC contract are satisfied ($\beta_k\underline{e}_k - \frac{1}{2}\theta_C^l \underline{e}_k^2 = \frac{\underline{e}_k^2}{2}(\theta_C^H - \theta_C^l) \geq 0$). Further, when adopting the above CLC contract, the sponsor can achieve the same retention rate as the FC contract by choosing $a_k = \bar{a}_k^{\circ}, f_k = \bar{f}_k^{\circ}$. Thus, the retention constraint is satisfied under the CLC contract as well. Hence, the solution with lower bound on effort $\underline{e}_k$, $\beta_k = \frac{1}{2}\theta_C^H \underline{e}_k, a_k = \bar{a}_k^{\circ}$, and $f_k = \bar{f}_k^{\circ}$ is feasible under the CLC contract, with the expected retention cost equals to

$$\left[\frac{1}{2}\theta_C^H \underline{e}_k\right] \times \underline{e}_k + \Theta_I(\bar{a}_k^{\circ}) + \bar{f}_k^{\circ}N\delta(\underline{e}_k, \bar{a}_k^{\circ}, \bar{f}_k^{\circ}) = \bar{s}_k^{\circ} + \Theta_I(\bar{a}_k^{\circ}) + \bar{f}_k^{\circ}N\delta(\underline{e}_k, \bar{a}_k^{\circ}, \bar{f}_k^{\circ}) = \Pi_{2k}^{F}(\underline{e}_k).$$

Therefore, the sponsor's optimal expected retention cost under the CLC contract is $\Pi_{2k}^{\hat{N}}(\underline{e}_k) \leq \Pi_{2k}^{F}(\underline{e}_k)$.

The proof is now complete.

□

EC.2.6. Proof of Proposition 12

We prove the first statement of the proposition. The proof of the second statement is similar and hence, avoided for brevity. To this end, we establish that there exists $\alpha > 0$ such that $\Pi_3^{\hat{N}}(\underline{a}, \underline{e}) < \Pi_3^F(\underline{a}, \underline{e}), \forall 0 \leq \underline{a} \leq \alpha, 0 \leq \underline{e} \leq 1$. Then, we show that there exist $\epsilon > 0$, such that $\Pi_3^{\hat{N}}(\underline{a}, \underline{e}) < \Pi_3^N, \forall \underline{a} \in [0, \epsilon], \sigma_1 \leq \underline{e} \leq \sigma_2$.

Consider the FC contract with $\underline{a} = 0$ and $\underline{e} \in [0, 1]$. Let the optimal monetary payment to the participant be $\bar{f}^\Diamond$. Recall that from Proposition 8, the optimal compensation under the FC contract $\bar{r}^\Diamond = 0$, $\bar{s}^\Diamond = \frac{1}{2}\theta_C^H \underline{e}^2$. Consider the following solution under the CLC contract with the same lower bounds on effort levels as in the FC contract above, and $\nu = 0$, $\beta = \frac{1}{2}\theta_C^H \underline{e}$, and $f = \bar{f}^\Diamond$. Then, the optimal effort level for the investigator is 0, and the optimal effort level that maximizes l-type coordinator's benefit is $\max\left\{\underline{e}, \frac{\beta}{\theta_C^l}\right\} = \max\left\{\underline{e}, \frac{\theta_C^H}{2\theta_C^l}\underline{e}\right\} = \underline{e}$. That is, the providers' effort levels under the above CLC contract are the same as that under the FC contract. Also, IR constraints under the CLC contract are satisfied ($\beta\underline{e} - \frac{1}{2}\theta_C^l \underline{e}^2 = \frac{\underline{e}^2}{2}(\theta_C^H - \theta_C^l) \geq 0$) implying when the sponsor chooses $\bar{f}^\Diamond$ under the CLC contract, the same retention rate is achieved with the same expected retention cost as the FC contract. Thus, $\Pi_3^{\hat{N}}(\underline{a} = 0, \underline{e}) \leq \Pi_3^F(\underline{a} = 0, \underline{e})$. We next show that $\Pi_3^{\hat{N}}(\underline{a} = 0, \underline{e}) < \Pi_3^F(\underline{a} = 0, \underline{e})$. Let C_1 be the sponsor's expected retention cost with the CLC contract above. Then

$$C_1 = \beta\underline{e} + (\lambda\underline{e} + \lambda_f \bar{f}^\Diamond) N \bar{f}^\Diamond = \Pi_3^F(\underline{a} = 0, \underline{e}).$$

Consider an alternative solution under the CLC contract where $\nu > 0$, $\beta = \frac{1}{2}\theta_C^H \underline{e}$, $f = \bar{f}^\Diamond - \frac{\lambda + \lambda_J \underline{e}}{\lambda_f \theta_I^H}\nu$. Then, the optimal effort level for the k-type investigator is $\hat{a}_k^\Diamond = \frac{\nu}{\theta_I^k}$. It is straightforward to verify that all the constraints are satisfied, and the solution is feasible under the CLC contract. Further, under the alternative solution, the sponsor's expected retention cost, which we express by $C_2(\nu)$, is given by

$$\begin{aligned} C_2(\nu) &= \mathbb{E}_{k,l}\left[\nu\frac{\nu}{\theta_I^k} + \beta\underline{e} + \left(\lambda\underline{e} + \lambda\hat{a}_k^\Diamond + \lambda_J \hat{a}_k^\Diamond \underline{e} + \lambda_f f\right)\left(\bar{f}^\Diamond - \frac{\lambda + \lambda_J \underline{e}}{\lambda_f \theta_I^H}\nu\right) N\right]. \\ &= \mathbb{E}_{k,l}\left[\nu\frac{\nu}{\theta_I^k} + \beta\underline{e} + \left(\lambda\underline{e} + \lambda_f \bar{f}^\Diamond - (\lambda + \lambda_J \underline{e})\left(\frac{\nu}{\theta_I^H} - \frac{\nu}{\theta_I^k}\right)\right)\left(\bar{f}^\Diamond - \frac{\lambda + \lambda_J \underline{e}}{\lambda_f \theta_I^H}\nu\right) N\right]. \end{aligned}$$

Let $\tau = \frac{\theta_I^L}{2}\frac{\lambda + \lambda_J \underline{e}}{\lambda_f \theta_I^H}\left(\lambda\underline{e} + \lambda_f \bar{f}^\Diamond \frac{2\theta_I^L - \theta_I^H}{\theta_I^H}\right) N$. Then when $\nu < \tau$,

$$\begin{aligned} \frac{\partial C_2(\nu)}{\partial \nu} =& \mathbb{E}_{k,l}\left[\frac{2\nu}{\theta_I^k} - \frac{\lambda + \lambda_J \underline{e}}{\lambda_f \theta_I^H}\left(\lambda\underline{e} + \lambda_f \bar{f}^\Diamond - (\lambda + \lambda_J \underline{e})\left(\frac{\nu}{\theta_I^H} - \frac{\nu}{\theta_I^k}\right)\right) N - (\lambda + \lambda_J \underline{e})\left(\frac{1}{\theta_I^H} - \frac{1}{\theta_I^k}\right)\left(\bar{f}^\Diamond - \frac{\lambda + \lambda_J \underline{e}}{\lambda_f \theta_I^H}\nu\right) N\right] \\ =& \mathbb{E}_{k,l}\left[-\frac{\lambda + \lambda_J \underline{e}}{\lambda_f \theta_I^H}\left(\lambda\underline{e} + \lambda_f \bar{f}^\Diamond \frac{2\theta_I^k - \theta_I^H}{\theta_I^k}\right) N + \frac{2}{\theta_I^k}\left(1 - \frac{N(\lambda + \lambda_J \underline{e})^2(\theta_I^H - \theta_I^k)}{\lambda_f {\theta_I^H}^2}\right)\nu\right] \\ \leq& \mathbb{E}_{k,l}\left[-\frac{\lambda + \lambda_J \underline{e}}{\lambda_f \theta_I^H}\left(\lambda\underline{e} + \lambda_f \bar{f}^\Diamond \frac{2\theta_I^L - \theta_I^H}{\theta_I^H}\right) N + \frac{2}{\theta_I^L}\nu\right] \\ <& 0. \end{aligned}$$

Further, notice that $C_2(0) - C_1 = 0$ implying $C_2(\nu) < C_1, \forall 0 < \nu < \tau$. Thus, the sponsor's optimal expected retention cost $\Pi_3^{\hat{N}}(\underline{a} = 0, \underline{e}) < \Pi_3^F(\underline{a} = 0, \underline{e}), \forall \underline{e} \in [0,1]$. Since $[\Pi_3^{\hat{N}}(\underline{a}, \underline{e}) - \Pi_3^F(\underline{a}, \underline{e})]$ is continuous in $(\underline{a}, \underline{e})$ in the compact set $[0,1] \times [0,1]$. Therefore, there exists $\alpha > 0$, such that $[\Pi_3^{\hat{N}}(\underline{a}, \underline{e}) - \Pi_3^F(\underline{a}, \underline{e})] < 0, \forall \underline{a} \in [0, \alpha], \underline{e} \in [0,1]$.

Next, we compare the sponsor's expected retention costs under the CLC and LC contracts when $\underline{a} = 0$. Recall that under the LC contract, the k-type investigator's optimal effort is $\frac{\nu^\diamond}{\theta_I^k}$ and the l-type coordinator's optimal effort is $\frac{\beta^\diamond}{\theta_C^l}$. The sponsor's expected retention cost is

$$\begin{aligned}\Pi_3^N =& \mathbb{E}_k\left[\nu^\diamond \times \frac{\nu^\diamond}{\theta_I^k}\right] + q\beta^\diamond \times \frac{\beta^\diamond}{\theta_C^H} + (1-q)\beta^\diamond \times \frac{\beta^\diamond}{\theta_C^L} + \mathbb{E}_k\left[q\delta(\frac{\nu^\diamond}{\theta_I^k}, \frac{\beta^\diamond}{\theta_C^H}, f^\diamond) + (1-q)\delta(\frac{\nu^\diamond}{\theta_I^k}, \frac{\beta^\diamond}{\theta_C^L}, f^\diamond)\right] N f^\diamond, \\ >& \mathbb{E}_k\left[\nu^\diamond \times \frac{\nu^\diamond}{\theta_I^k}\right] + q\frac{\beta^\diamond}{2} \times \frac{\beta^\diamond}{\theta_C^H} + (1-q)\frac{\beta^\diamond}{2} \times \max\left\{\frac{\beta^\diamond}{2\theta_C^L}, \frac{\beta^\diamond}{\theta_C^H}\right\} + \mathbb{E}_k\left[q\delta(\frac{\nu^\diamond}{\theta_I^k}, \frac{\beta^\diamond}{\theta_C^H}, f^\diamond) + (1-q)\delta(\frac{\nu^\diamond}{\theta_I^k}, \max\left\{\frac{\beta^\diamond}{2\theta_C^L}, \frac{\beta^\diamond}{\theta_C^H}\right\}, f^\diamond)\right] N f^\diamond.\end{aligned}$$

Now, consider the following feasible value of $\beta = \frac{\beta^\diamond}{2}$ for the CLC contract with $\underline{a} = 0, \nu = \nu^\diamond$, $\underline{e} = \frac{\beta^\diamond}{\theta_C^H}$. Then, $\Pi_3^{\hat{N}}(\underline{a} = 0, \underline{e} = \frac{\beta^\diamond}{\theta_C^H})$ is the same as the LHS expression in the last line above. Therefore, $\Pi_3^{\hat{N}}(\underline{a} = 0, \underline{e} = \frac{\beta^\diamond}{\theta_C^H}) < \Pi_3^N$. Since $\Pi_3^{\hat{N}}(\cdot,\cdot)$ is a continuous function. Therefore, there exist $\epsilon, 0 < \epsilon < \max\left\{\frac{\beta^\diamond}{\theta_C^H}, 1 - \frac{\beta^\diamond}{\theta_C^H}\right\}$, such that $\Pi_3^{\hat{N}}(\underline{a}, \underline{e}) < \Pi_3^N, \forall \underline{a} \in [0, \epsilon], \underline{e} \in [\frac{\beta^\diamond}{\theta_C^H} - \epsilon, \frac{\beta^\diamond}{\theta_C^H} + \epsilon]$.

Now, let $\alpha_0 = \min\{\alpha, \epsilon\}$, $\sigma_1 = \frac{\beta^\diamond}{\theta_C^H} - \epsilon, \sigma_2 = \frac{\beta^\diamond}{\theta_C^H} + \epsilon$. Then, we have

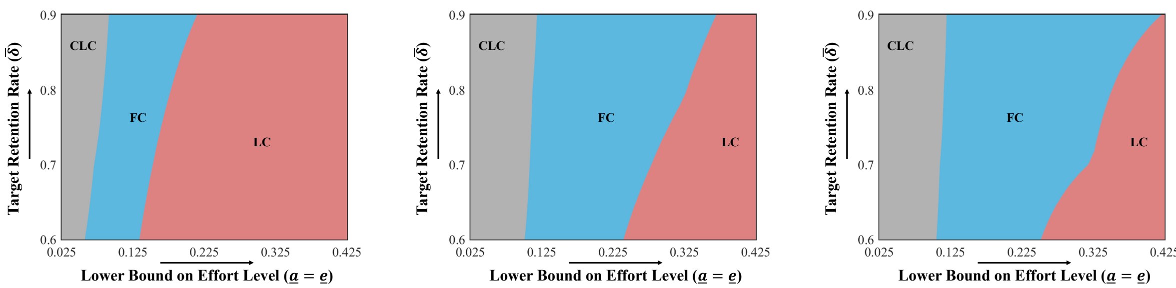
$$\Pi_3^{\hat{N}}(\underline{a}, \underline{e}) < \Pi_3^F(\underline{a}, \underline{e}), \quad \Pi_3^{\hat{N}}(\underline{a}, \underline{e}) < \Pi_3^N, \qquad \forall 0 \le \underline{a} \le \alpha_0, \sigma_1 \le \underline{e} \le \sigma_2.$$

The result now follows.

EC.3. Computational Study Results for the OM Model

EC.3.1. Understanding the Relative Performance of the FC, LC, and CLC Contracts

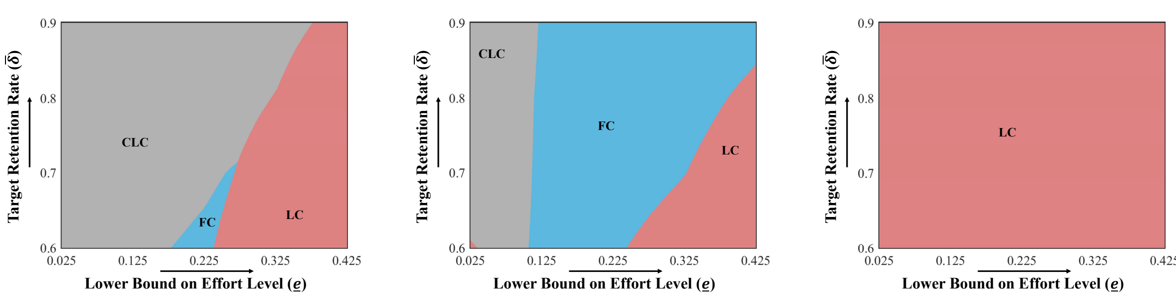
Figure EC.1 illustrates the impact of the lower bound on effort, the target retention rate, and the effectiveness of effort on the best-performing contract among the FC, LC, and CLC contracts under the OM model when $\underline{a} = \underline{e}$. We observe that the cost of the CLC contract is lower on average for lower $\underline{e}$, while the FC and LC contracts perform better on average for medium and high values of $\underline{e}$, respectively. The performance of the three contracts relative to $\bar{\delta}$ and λ is the same as that under the SI model.



(a) $\lambda = 0.4$ (b) $\lambda = 0.7$ (c) $\lambda = 1.0$

Figure EC.1 Comparing the Performances of the FC, LC, and CLC Contracts under the OM Model

We further extend our numerical study to incorporate the settings where $\underline{a} \neq \underline{e}$. Figure EC.2 illustrates the impact of the lower bound on effort and the target retention rate on the best-performing contract among the FC, LC, and CLC contracts under the OM model for $\lambda = 0.7$. We observe that for low and medium $\underline{a}$ values, the cost of the CLC contract is lower on average for lower $\underline{e}$, while the FC and LC contracts perform better on average for medium and high values of $\underline{e}$, which is consistent with our findings above. When $\underline{a}$ is high, the cost of the LC contract is lower on average. This suggests that when either of the lower bound values $\underline{a}, \underline{e}$ is high, the LC contract is preferred on average.



(a) $\underline{a} = 0.025$ (b) $\underline{a} = 0.225$ (c) $\underline{a} = 0.425$

Figure EC.2 Comparing the Performances of the FC, LC, and CLC Contracts for $\lambda = 0.7$ under the OM Model

EC.3.2. When to Adopt the Optimal Compensation Contracts

Among the instances where $\mathcal{R}_3 \geq 1.10$, 58% has effectiveness of effort parameter $\lambda \geq 0.85$. For instances with $\lambda \geq 0.85$, we illustrate the impact of $\lambda_f, p, q, \underline{e}$ on the ratio $\mathcal{R}_3$ in Figure EC.3. Similar to the observations in Section 5.2, given the high λ values, the ratio $\mathcal{R}_3$ is typically greater than 1.10 when at least two of the following conditions are satisfied: (1) the lower bounds on effort levels $(\underline{e}, \underline{a})$ are either high or low, (2) the effectiveness of monetary payment (λ_f) is low, (3) the probability of having a high-type coordinator (q) is low and (4) the probability of having a high-type investigator (p) is low.

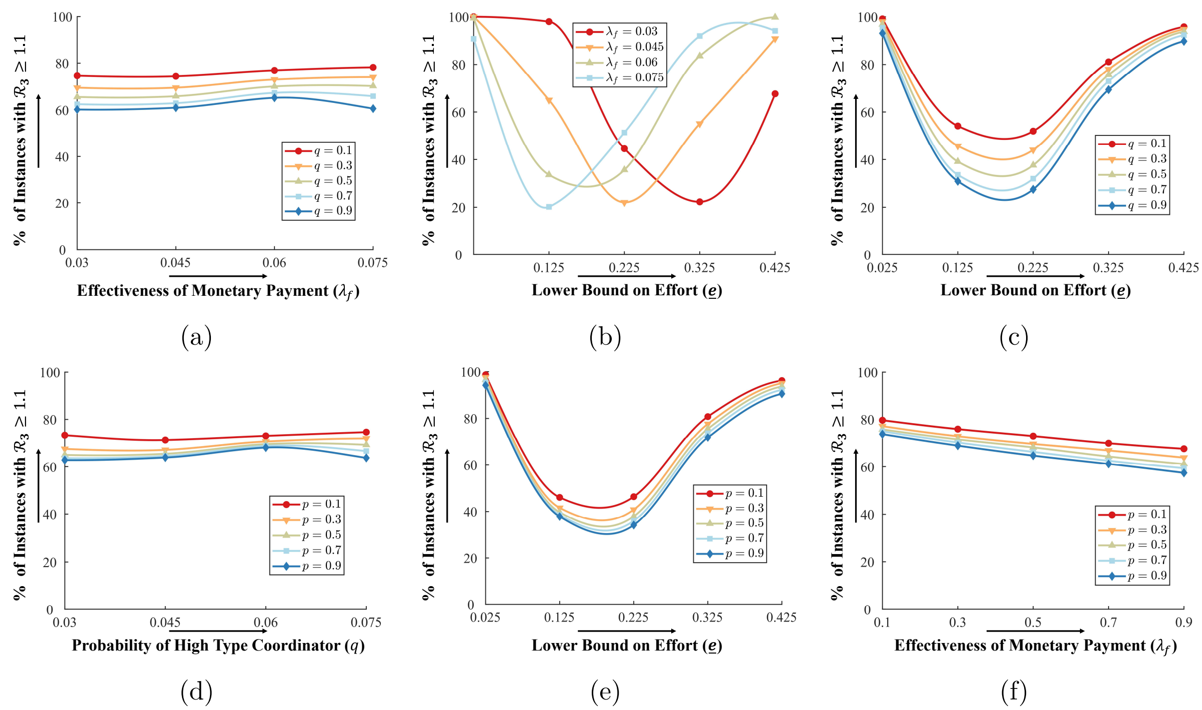


Figure EC.3 **Impact of $\underline{e}, \lambda_f$, p, and q on the Ratio $\mathcal{R}_3$**