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

Xueze Song
The University of Alabama, xueze.song@ua.edu
Mili Mehrotra
University of Illinois Urbana-Champaign, milim@illinois.edu
Tharanga Rajapakshe
University of Florida, tharanga.rajapakshe@warrington.ufl.edu

EC.1. Regulations and Industry Guidelines for Clinical Studies

Below, we discuss the regulations and industry guidelines for clinical studies, which ensure that the effort is typically observable and contractable and that outcome-based compensation contracts are discouraged.

1. **Observability and Contractability of Effort:** Clinical studies must comply with federal regulations and adhere to industry guidelines. Below, we discuss the regulations that mandate effort documentation from the providers and monitoring by the sponsor. These regulations enable the sponsor to track the providers' efforts, making them observable and contractable.
   - *Effort Documentation by Providers:*
     — There are two key federal regulations that require detailed documentation of the activities performed by the providers during the clinical studies: General responsibilities of investigators (2004) (§21 CFR 312.60) and Investigator recordkeeping and record retention (2004) (§21 CFR 312.62). The former regulation requires an investigator to be "responsible for ensuring that an investigation is conducted according to the signed investigator statement, the investigational plan, and applicable regulations." The latter regulation requires an investigator to maintain adequate records including disposition of the drug (dates, quantity, and use by subjects) and case histories (such as signed and dated consent forms and medical record, progress notes of the physician, the individual's hospital chart(s), and the nurses' notes, etc.). The documentation of case histories also enables the identification of the coordinator's effort in explaining the study to the participants, obtaining informed consent, collecting data, etc.

     — The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) develops an internationally recognized standard for Good Clinical Practice (GCP). The ICH GCP guidelines also state that the providers' activities should be documented and retained.

> *"The investigator/institution should have access to and the ability to maintain and retain the essential records generated by the investigator/institution before, during and after the trial... Whether a specific clinical trial record generated before, during and after the trial is essential and needs to be retained should be based on the following criteria ...*
>
> *(d) Documents the conduct of relevant trial procedures...*
>
> *(j) Is, where necessary, documentation that demonstrates signatures/initials of staff undertaking trial-specific activities; for example, completing data acquisition tools;"* (ICH 2023)

Such guidelines further validate that the providers' efforts are documented.

- *Monitoring by the Sponsor:*
  - —Federal regulation on General responsibilities of sponsors (2004) (§21 CFR 312.50) specifies a sponsor's responsibility for monitoring the providers' investigation during the study to ensure that the investigation(s) is conducted following the general investigational plan.
  - —The ICH GCP also emphasizes the sponsor's responsibility for monitoring the providers' investigation during the study.

    > *"The sponsor should ensure that trial processes are conducted in compliance with the trial protocol and related documents as well as with applicable regulatory requirements and ethical standards ...The selection and oversight of investigators and service providers are fundamental features of the oversight process. Oversight by the sponsor includes quality assurance and quality control processes relating to the trial-related activities of investigators and service providers."* (ICH 2023)
  - —Pharmaceutical companies (sponsors) regularly conduct monitoring or audits on the study sites to ensure the providers' investigation conforms to their effort reporting (Johnson & Johnson 2020, Merck 2019). Further, the sponsors pay interim visits to monitor the progress of a clinical trial, maintain trial integrity, and maintain a communication link with the research team (Love et al. 2022, Pfeiffer and Wells 2017).
- *Effort Recording and Monitoring Tools*: To facilitate effort recording and verification, we observe several tools available in practice, such as the Research Effort Tracking Application from the University of Michigan (James et al. 2011), ASCO Clinical Trial Workload Assessment Tool (Good et al. 2016). Commercial Clinical Trial Management System are also available to track the providers' effort (e.g., IQVIA 2021, Veeva 2016). There is a growing trend of integrating blockchain technology in such systems to ensure transparency and immutability of effort records (Arjun et al. 2024, Subbiah 2023).

In conclusion, the above regulations, guidelines, and availability of tracking tools imply that a sponsor can observe and verify providers' efforts, allowing the sponsor to offer compensation contracts based on providers' efforts.

2. **Non-Applicability of Outcome-Based Compensation Contracts:** The U.S. Food and Drug Administration (FDA) discourages contracting on outcomes to avoid bias and conflict of interest for service providers. For example, Financial disclosure by clinical investigators (1998) (§21 CFR Part 54) regulation identifies the financial interest of the clinical investigator in the outcome of the study due to the payment structure as a potential source of bias and the FDA will take any necessary action (including disregarding the clinical study) in case it finds the integrity of the data affected by the financial interests of any clinical investigator.

   In light of this regulation, many medical associations, Institutional Review Boards (IRBs), and healthcare companies have advised against compensation based on outcomes in clinical studies. For example:

   > *"Compensation must be: ... Explicitly identified in the investigator contract and linked to their performing specific and necessary protocol-required services (e.g., medical procedures, collection of data)."* (Pfizer 2021)
   >
   > *"...the rate of payment for professional services should remain constant irrespective of the [Principal Investigator]'s success enrolling or completing study subjects compared to their target."* (Rose 2021)
   >
   > *"When setting up a reimbursement model for a study's [Principal Investigator], be sure to adhere to the following rules: All compensation should be linked to the [Principal Investigator] performing specific and necessary services — medical procedures, the collection of data, reviews of study reports, etc. ...Financial compensation to a [Principal Investigator] should not: Be tied to any particular study outcome..."* (Ingram 2021)

EC.2. 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.2.1. Proof of Lemma EC.1

Note that

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

The proof is now completed. □

EC.2.2. 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 $\delta(\gamma a_{kl}^*, \gamma e_{kl}^*, \gamma f_{kl}^*) \geq \bar{\delta}$. Further, the objective function is increasing in a_{kl}, e_{kl}, and f_{kl}. Therefore, $\Pi_{kl}^{CM}(\gamma a_{kl}^*, \gamma e_{kl}^*, \gamma f_{kl}^*) \leq \Pi_{kl}^{CM}(a_{kl}^*, e_{kl}^*, f_{kl}^*)$ 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

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

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

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 $0 < \bar{a}_{kl} < 1, 0 < \bar{e}_{kl} < 1$ and constraint EC.2 is satisfied.

Using inequalities from Lemma EC.1 and Assumption 1, we can show that $0 < \bar{a}_{kl} < 1, 0 < \bar{e}_{kl} < 1$. 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.2.3. 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

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

From Assumption 1, $\lambda_f(\lambda_f^3\theta_C^l\theta_I^k - \bar{\delta}^2 N^2\lambda_J^2)^3$ is positive. Denote the numerator in the above expression on the right hand side as $\mathcal{L}_{kl}$. Then $\frac{\partial f^*_{kl}}{\partial \bar{\delta}} = \frac{\mathcal{L}_{kl}}{\lambda_f(\lambda_f^3\theta_C^l\theta_I^k - \bar{\delta}^2 N^2\lambda_J^2)^3}$. 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.2.4. 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 $l=H$.

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

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:

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

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

$$\hat{a}_{kL} = \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J + \lambda_f\theta_C^L)}{\lambda_f^2\theta_I^k\theta_C^L - \bar{\delta}^2N^2\lambda_J^2} = a_{kL}^*, \qquad \hat{e}_{kL} = \frac{\lambda\bar{\delta}N(\bar{\delta}N\lambda_J + \lambda_f\theta_I^k)}{\lambda_f^2\theta_I^k\theta_C^L - \bar{\delta}^2N^2\lambda_J^2} = e_{kL}^*,$$

$$\hat{a}_{kH} = \frac{\lambda\bar{\delta}N(q\bar{\delta}N\lambda_J + \lambda_f(\theta_C^H - \bar{q}\theta_C^L))}{\lambda_f^2\theta_I^k(\theta_C^H - \bar{q}\theta_C^L) - q\bar{\delta}^2N^2\lambda_J^2}, \quad \hat{e}_{kH} = \frac{q\lambda\bar{\delta}N(\bar{\delta}N\lambda_J + \lambda_f\theta_I^k)}{\lambda_f^2\theta_I^k(\theta_C^H - \bar{q}\theta_C^L) - q\bar{\delta}^2N^2\lambda_J^2}.$$

Using Assumption 1 and Lemma EC.1, we can show that $0 < \hat{a}_{kH} < 1, 0 < \hat{e}_{kH} < 1$ 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.2.5. Proof of Corollaries 2-3 and Propositions 3-4

The proofs of Corollaries 2, 3, 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.

☐

EC.2.6. 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

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

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,

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

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

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. ☐

EC.2.7. 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 $\hat{y} \leq \frac{zx-\sqrt{q}}{(1-\sqrt{q})z}$ and decreasing in $\hat{y}$ when $\hat{y} \geq \frac{\hat{w}x-\sqrt{q}}{(1-\sqrt{q})\hat{w}}$. That is, we have $\Pi_2 - \Pi_1$ is increasing in θ_C^H when $\theta_C^H \leq \frac{\lambda_f^2 \theta_C^L \theta_I^H - \bar{\delta}^2 N^2 (1-\sqrt{q})}{\bar{\delta} N \lambda_f \lambda_J (1-\sqrt{q})\theta_I^H}$, and is decreasing in θ_C^H when $\theta_C^H \geq \frac{\lambda_f^2 \theta_C^H \theta_I^H - \bar{\delta}^2 N^2 (1-\sqrt{q})}{\bar{\delta} N \lambda_f \lambda_J (1-\sqrt{q})\theta_I^H}$. 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.2.8. Proof of Proposition 7

We first simplify the expressions for $\Pi_m, m = 1,2,3$, by defining the following values: $x = \dfrac{\theta_C^L \lambda_f}{\bar{\delta} N \lambda_J}$, $y = \frac{\theta_C^H \lambda_f}{\bar{\delta} N \lambda_J} - x$, $z = \frac{\theta_I^L \lambda_f}{\bar{\delta} N \lambda_J}$, $w = \frac{\theta_I^H \lambda_f}{\bar{\delta} N \lambda_J} - z$, 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:

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

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,

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

The 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.3. 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.3.1. 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

$$\begin{aligned}
\min_{a_k\in[0,1];s_k,f_k\geq 0} \quad & \mathbb{E}_l[\Theta_I(a_k;\theta_I^k)+s_k+f_kN\delta(a_k,e_{kl}^{\circ},f_k)], && && \text{(EC.3)}\\
\text{s.t.} \quad & \delta(a_k,e_{kl}^{\circ},f_k)\geq\bar{\delta}, && \forall l\in\{L,H\}, && \text{(EC.4)}\\
IR: \quad & s_k-\Theta_C(e_{kl}^{\circ};\theta_C^l)\geq 0, && \forall l\in\{L,H\}, && \text{(EC.5)}\\
IC: \quad & e_{kl}^{\circ}=\arg\max_{e_{kl}\geq \underline{e}_k}(s_k-\Theta_C(e_{kl};\theta_C^l)), && \forall l\in\{L,H\}. && \text{(EC.6)}
\end{aligned}$$

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

$$\begin{aligned}
\min_{r,s,f\geq 0} \quad & \mathbb{E}_{k,l}[r+s+fN\delta(a_k^{\Diamond},e_l^{\Diamond},f)], && && \text{(EC.7)}\\
\text{s.t.} \quad & \delta(a_k^{\Diamond},e_l^{\Diamond},f)\geq\bar{\delta}, && \forall k,l\in\{L,H\}, && \text{(EC.8)}\\
IR: \quad & r-\Theta_I(a_k^{\Diamond};\theta_I^k)\geq 0, && \forall k\in\{L,H\}, && \text{(EC.9)}\\
IC: \quad & a_k^{\Diamond}=\arg\max_{a_k\geq\underline{a}}(r-\Theta_I(a_k;\theta_I^k)), && \forall k\in\{L,H\}, && \text{(EC.10)}\\
IR: \quad & s-\Theta_C(e_l^{\Diamond};\theta_C^l)\geq 0, && \forall l\in\{L,H\}, && \text{(EC.11)}\\
IC: \quad & e_l^{\Diamond}=\arg\max_{e_l\geq\underline{e}}(s-\Theta_C(e_l;\theta_C^l)), && \forall l\in\{L,H\}. && \text{(EC.12)}
\end{aligned}$$

EC.3.1.1. 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 $s_k=\Theta_C(\underline{e}_k;\theta_C^H)$. 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:

$$\min_{a_k\in[0,1]} \quad \Theta_I(a_k;\theta_I^k)+\Theta_C(\underline{e}_k;\theta_C^H)+\frac{\bar{\delta}-\lambda(a_k+\underline{e}_k)-\lambda_J a_k\underline{e}_k}{\lambda_f}N\bar{\delta}.$$

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 $r^{\Diamond}=\frac{1}{2}\theta_I^H\underline{a}^2$, $s^{\Diamond}=\frac{1}{2}\theta_C^H\underline{e}^2$, $f^{\Diamond}=\frac{\bar{\delta}-\lambda(\underline{a}+\underline{e})-\lambda_J\underline{a}\underline{e}}{\lambda_f}$. This completes the proof. □

EC.3.2. Proof of Corollary 4

We first consider the performance under the SI model. Define

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

Note that $\Pi_{2k}^F=\Theta_I(a_k^\circ;\theta_I^k)+\Theta_C(e_k^\circ;\theta_C^H)+f_k^\circ N\delta(a_k^\circ,e_k^\circ,f_k^\circ)=\mathcal{V}_{kH}(\underline{e}_k)$. Further, we can write $\Pi_k^{SI\circ}=q\mathcal{V}_{kH}(e_{kH}^\circ)+\bar{q}\mathcal{V}_{kL}(e_{kL}^\circ)+\frac{\bar{q}}{2}(\theta_C^H-\theta_C^L){e_{kH}^\circ}^2$.

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

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

Combining the above two inequalities, we have

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

Therefore,

$$\begin{aligned}
\frac{\Pi_{2k}^F}{\Pi_k^{SI\circ}}&\leq\frac{\Pi_k^{SI\circ}+q\left(\theta_C^H-\frac{\bar{\delta}^2N^2\lambda_J^2}{\lambda_f^2\theta_I^k}\right)(\underline{e}_k-e_{kH}^\circ)(\underline{e}_k-e_{kH}^*)+\bar{q}\left(\theta_C^L-\frac{\bar{\delta}^2N^2\lambda_J^2}{\lambda_f^2\theta_I^k}\right)(\underline{e}_k-e_{kL}^\circ)(\underline{e}_k-e_{kL}^*)+\frac{\bar{q}}{2}(\theta_C^H-\theta_C^L)(\underline{e}_k^2-{e_{kH}^\circ}^2)}{\Pi_k^{SI\circ}}\\
&=1+\frac{q\left(\theta_C^H-\frac{\bar{\delta}^2N^2\lambda_J^2}{\lambda_f^2\theta_I^k}\right)(\underline{e}_k-e_{kH}^\circ)(\underline{e}_k-e_{kH}^*)+\bar{q}\left(\theta_C^L-\frac{\bar{\delta}^2N^2\lambda_J^2}{\lambda_f^2\theta_I^k}\right)(\underline{e}_k-e_{kL}^\circ)(\underline{e}_k-e_{kL}^*)+\frac{\bar{q}}{2}(\theta_C^H-\theta_C^L)(\underline{e}_k^2-{e_{kH}^\circ}^2)}{\Pi_k^{SI\circ}}\\
&\leq1+\frac{q\left(\theta_C^H-\frac{\bar{\delta}^2N^2\lambda_J^2}{\lambda_f^2\theta_I^k}\right)(\underline{e}_k-e_{kH}^\circ)(\underline{e}_k-e_{kH}^*)+\bar{q}\left(\theta_C^L-\frac{\bar{\delta}^2N^2\lambda_J^2}{\lambda_f^2\theta_I^k}\right)(\underline{e}_k-e_{kL}^\circ)(\underline{e}_k-e_{kL}^*)+\frac{\bar{q}}{2}(\theta_C^H-\theta_C^L)(\underline{e}_k^2-{e_{kH}^\circ}^2)}{\frac{1}{2}\theta_C^L{e_{LL}^\circ}^2}\\
&=1+\frac{2}{\theta_C^L{e_{LL}^\circ}^2}\left[q\left(\theta_C^H-\frac{\bar{\delta}^2N^2\lambda_J^2}{\lambda_f^2\theta_I^k}\right)(\underline{e}_k-e_{kH}^\circ)(\underline{e}_k-e_{kH}^*)+\bar{q}\left(\theta_C^L-\frac{\bar{\delta}^2N^2\lambda_J^2}{\lambda_f^2\theta_I^k}\right)(\underline{e}_k-e_{kL}^\circ)(\underline{e}_k-e_{kL}^*)\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}\\
&=1+\frac{2}{\theta_C^L{e_{LL}^\circ}^2}\mathbb{E}_l\left[\left(\theta_C^l-\frac{\bar{\delta}^2N^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}.
\end{aligned}$$

Next, consider the result for the OM model. Let $\mathcal{U}_{kl}(x,y)=\frac{1}{2}\theta_C^l y^2+\frac{1}{2}\theta_I^k x^2+\bar{\delta}N\frac{\bar{\delta}-\lambda(x+y)-\lambda_J xy}{\lambda_f}$. Then $\Pi_3^F=\mathcal{U}_{HH}(\underline{a},\underline{e})$, $\Pi^{OM\Diamond}=\mathbb{E}_{k,l}[\mathcal{U}_{kl}(a_{kl}^\Diamond,e_{kl}^\Diamond)+\frac{1}{2}(\theta_C^H-\theta_C^l){e_{kH}^\Diamond}^2+\frac{1}{2}(\theta_I^H-\theta_I^k){a_{Hl}^\Diamond}^2]$. It's easy to verify that $\mathcal{U}_{kl}(x,y)$ is a jointly convex function. Therefore,

$$\mathcal{U}_{kl}(\underline{a},\underline{e})-\mathcal{U}_{kl}(a_{kl}^\Diamond,e_{kl}^\Diamond)\leq\nabla\mathcal{U}_{kl}(\underline{a},\underline{e})^\intercal(\underline{a}-a_{kl}^\Diamond,\underline{e}-e_{kl}^\Diamond),$$

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

$$\begin{aligned}
\Pi_3^F&=\mathcal{U}_{HH}(\underline{a},\underline{e})=\mathbb{E}_{k,l}\left[\mathcal{U}_{kl}(\underline{a},\underline{e})+\frac{1}{2}(\theta_C^H-\theta_C^l)\underline{e}^2+\frac{1}{2}(\theta_I^H-\theta_I^k)\underline{a}^2\right]\\
&=\mathbb{E}_{k,l}\left[\mathcal{U}_{kl}(\underline{a},\underline{e})-\mathcal{U}_{kl}(a_{kl}^\Diamond,e_{kl}^\Diamond)\right]+\mathbb{E}_{k,l}\left[\frac{1}{2}(\theta_C^H-\theta_C^l)(\underline{e}^2-{e_{kH}^\Diamond}^2)+\frac{1}{2}(\theta_I^H-\theta_I^k)(\underline{a}^2-{a_{Hl}^\Diamond}^2)\right]+\Pi^{OM\Diamond}
\end{aligned}$$

$$
\begin{aligned}
&\leq \mathbb{E}_{k,l}\left[\nabla \mathcal{U}_{kl}(\underline{a},\underline{e})^{\mathsf{T}}(\underline{a}-a^{\Diamond}_{kl},\underline{e}-e^{\Diamond}_{kl})\right]+\mathbb{E}_{k,l}\left[\frac{1}{2}(\theta^{H}_{C}-\theta^{l}_{C})(\underline{e}^{2}-{e^{\Diamond}_{kH}}^{2})+\frac{1}{2}(\theta^{H}_{I}-\theta^{k}_{I})(\underline{a}^{2}-{a^{\Diamond}_{Hl}}^{2})\right]+\Pi^{OM\Diamond}\\
&=\mathbb{E}_{k,l}\Big[(\underline{a}-a^{\Diamond}_{kl})(\underline{a}\theta^{k}_{I}-\frac{\bar{\delta}N(\lambda+\lambda_{J}\underline{e})}{\lambda_{f}})+(\underline{e}-e^{\Diamond}_{kl})(\underline{e}\theta^{l}_{C}-\frac{\bar{\delta}N(\lambda+\lambda_{J}\underline{a})}{\lambda_{f}})\\
&\qquad+\frac{1}{2}(\theta^{H}_{C}-\theta^{l}_{C})(\underline{e}^{2}-{e^{\Diamond}_{kH}}^{2})+\frac{1}{2}(\theta^{H}_{I}-\theta^{k}_{I})(\underline{a}^{2}-{a^{\Diamond}_{Hl}}^{2})\Big]+\Pi^{OM\Diamond}.
\end{aligned}
$$

Further,

$$
\begin{aligned}
\frac{\Pi^{F}_{3}}{\Pi^{OM\Diamond}}&=1+\frac{\mathbb{E}_{k,l}\Big[(\underline{a}-a^{\Diamond}_{kl})(\underline{a}\theta^{k}_{I}-\frac{\bar{\delta}N(\lambda+\lambda_{J}\underline{e})}{\lambda_{f}})+(\underline{e}-e^{\Diamond}_{kl})(\underline{e}\theta^{l}_{C}-\frac{\bar{\delta}N(\lambda+\lambda_{J}\underline{a})}{\lambda_{f}})+\frac{1}{2}(\theta^{H}_{C}-\theta^{l}_{C})(\underline{e}^{2}-{e^{\Diamond}_{kH}}^{2})+\frac{1}{2}(\theta^{H}_{I}-\theta^{k}_{I})(\underline{a}^{2}-{a^{\Diamond}_{Hl}}^{2})\Big]}{\Pi^{OM\Diamond}}\\
&\leq 1+\frac{\mathbb{E}_{k,l}\Big[(\underline{a}-a^{\Diamond}_{kl})(\underline{a}\theta^{k}_{I}-\frac{\bar{\delta}N(\lambda+\lambda_{J}\underline{e})}{\lambda_{f}})+\frac{1}{2}(\theta^{H}_{I}-\theta^{k}_{I})(\underline{a}^{2}-{a^{\Diamond}_{Hl}}^{2})\Big]}{\frac{1}{2}\theta^{L}_{C}{a^{\Diamond}_{LL}}^{2}}+\frac{\mathbb{E}_{k,l}\Big[(\underline{e}-e^{\Diamond}_{kl})(\underline{e}\theta^{l}_{C}-\frac{\bar{\delta}N(\lambda+\lambda_{J}\underline{a})}{\lambda_{f}})+\frac{1}{2}(\theta^{H}_{C}-\theta^{l}_{C})(\underline{e}^{2}-{e^{\Diamond}_{kH}}^{2})\Big]}{\frac{1}{2}\theta^{L}_{C}{e^{\Diamond}_{LL}}^{2}}\\
&=1+\frac{2}{\theta^{L}_{I}{a^{\Diamond}_{LL}}^{2}}\mathbb{E}_{k,l}\Big[(\underline{a}-a^{\Diamond}_{kl})(\underline{a}\theta^{k}_{I}-\frac{\bar{\delta}N(\lambda+\lambda_{J}\underline{e})}{\lambda_{f}})+\frac{1}{2}(\theta^{H}_{I}-\theta^{k}_{I})(\underline{a}^{2}-{a^{\Diamond}_{Hl}}^{2})\Big]\\
&\qquad+\frac{2}{\theta^{L}_{C}{e^{\Diamond}_{LL}}^{2}}\mathbb{E}_{k,l}\Big[(\underline{e}-e^{\Diamond}_{kl})(\underline{e}\theta^{l}_{C}-\frac{\bar{\delta}N(\lambda+\lambda_{J}\underline{a})}{\lambda_{f}})+\frac{1}{2}(\theta^{H}_{C}-\theta^{l}_{C})(\underline{e}^{2}-{e^{\Diamond}_{kH}}^{2})\Big].
\end{aligned}
$$

□

EC.3.3. The LC Contract

Under the SI model, the sponsor's problem is

$$
\begin{aligned}
\min_{a_{k}\in[0,1];f_{k},\beta_{k}\geq 0}\quad & \mathbb{E}_{l}[\Theta_{I}(a_{k};\theta^{k}_{I})+\beta_{k}e^{\circ}_{kl}+f_{k}N\delta(a_{k},e^{\circ}_{kl},f_{k})] && && \text{(EC.13)}\\
\text{s.t.}\quad & \delta(a_{k},e^{\circ}_{kl},f_{k})\geq\bar{\delta}, && \forall l\in\{L,H\}, && \text{(EC.14)}\\
IR:\ & \beta_{k}e^{\circ}_{kl}-\Theta_{C}(e^{\circ}_{kl};\theta^{l}_{C})\geq 0, && \forall l\in\{L,H\}, && \text{(EC.15)}\\
IC:\ & e^{\circ}_{kl}\in\arg\max_{e}\left\{\beta_{k}e-\Theta_{C}(e;\theta^{l}_{C})\right\}, && \forall l\in\{L,H\}. && \text{(EC.16)}
\end{aligned}
$$

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:

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

EC.3.3.1. Proof of Proposition 9. Under the SI model, it is straightforward to derive that for any given β_k, the coordinator's optimal effort $e^{\circ}_{kl} = \frac{\beta_k}{\theta^l_C}, l \in \{L, H\}$. At optimality, we can show that retention constraint (EC.14) is binding for $l = H$ and hence, satisfied for $l = L$. Thus, we have $f^{\circ}_k = \frac{\bar{\delta} - \lambda(a^{\circ}_k + e^{\circ}_{kH}) - \lambda_J a^{\circ}_k e^{\circ}_{kH}}{\lambda_f}$. Now it is easy to verify that the sponsor's problem reduces to Problem P$_{LSI}$ below:

Problem P$_{LSI}$

$$\min_{a_k \in [0,1]; \beta_k \geq 0} \quad \mathbb{E}_l[\Theta_{I}(a_k; \theta^k_{I}) + \beta_k e^{\circ}_{kl} + f^{\circ}_k N \delta(a_k, e^{\circ}_{kl}, f^{\circ}_k)].$$

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 $e^{\diamondsuit}_l = \frac{\beta}{\theta^l_C}, l \in \{L, H\}$ (resp., $a^{\diamondsuit}_k = \frac{\nu}{\theta^k_I}, k \in \{L, H\}$), 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$_{LOM}$

$$\min_{\nu \geq 0, \beta \geq 0} \quad \mathbb{E}_{k,l}[\nu a^{\diamondsuit}_k + \beta e^{\diamondsuit}_l + f^{\diamondsuit} N \delta(a^{\diamondsuit}_k, e^{\diamondsuit}_l, f^{\diamondsuit})].$$

□

EC.3.3.2. Proof of Corollary 5. When both providers have a single type, we have $\theta^{L}_{I} = \theta^{H}_{I} = \theta_{I}, \theta^{L}_{C} = \theta^{H}_{C} = \theta_{C}$. 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

$$\begin{aligned} \min_{a \in [0,1]; f, 0 \leq \beta \leq \theta_C} \quad & [\Theta_{I}(a; \theta_{I}) + \frac{\beta^2}{\theta_{C}} + f N \delta(a, \frac{\beta}{\theta_{C}}, f)], \\ \text{s.t.} \quad & \delta(a, \frac{\beta}{\theta_{C}}, f) \geq \bar{\delta}. \end{aligned}$$

We can rewrite the objective function as $[\Theta_{I}(a; \theta_{I}) + \frac{1}{2} 2\theta_{C} (\frac{\beta}{\theta_C})^2 + f N \delta(a, \frac{\beta}{\theta_C}, f)]$. 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

$a^{\circ} = \frac{\lambda \bar{\delta} N(\bar{\delta} N \lambda_J + 2\lambda_f \theta_C)}{\lambda_f^2 2\theta_C \theta_I - \bar{\delta}^2 N^2 \lambda_J^2}$, $\frac{\beta^{\circ}}{\theta_C} = \frac{\lambda \bar{\delta} N(\bar{\delta} N \lambda_J + \lambda_f \theta_I)}{2\lambda_f^2 \theta_C \theta_I - \bar{\delta}^2 N^2 \lambda_J^2}$, 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:

$$\Pi^N_2 = [\frac{1}{2}\theta_{I} a^{\circ 2} + \frac{1}{2} 2\theta_{C} (\frac{\beta^{\circ}}{\theta_{C}})^2 + f^{\circ} N \bar{\delta}].$$

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

$$\begin{aligned} \Pi^N_2 &= \frac{\bar{\delta}^2 N}{\lambda_f} + \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f} G(2x, z), \\ \Pi^{SI\circ} &= \frac{\bar{\delta}^2 N}{\lambda_f} + \frac{\bar{\delta} N \lambda^2}{2\lambda_J \lambda_f} G(x, z). \end{aligned}$$

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

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

Then we have,

$$\begin{aligned}\frac{\Pi_3^N - \Pi^{OM^\Diamond}}{\Pi^{OM^\Diamond}} &= \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}(G(2x,2z)-G(x,z)) = \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}(G(2x,2z)-G(2x,z)+G(2x,z)-G(x,z))\\ &\leq \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}(G(x,2z)-G(x,z)+G(2x,z)-G(x,z)) = \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}(\frac{x(1+z)^2+z(1+x)^2}{(xz-1)(2xz-1)})\\ &< \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}\frac{x(1+z)^2+z(1+x)^2}{(xz-1)^2} = \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2(xe^{\Diamond 2}+za^{\Diamond 2})}{2\lambda^2}\\ &= \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2 xe^{\Diamond 2}}{2\lambda^2} + \frac{1}{\Pi^{OM^\Diamond}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2 za^{\Diamond 2}}{2\lambda^2}\\ &\leq \frac{1}{\frac{1}{2}\theta_C e^{\Diamond 2}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2 xe^{\Diamond 2}}{2\lambda^2} + \frac{1}{\frac{1}{2}\theta_I a^{\Diamond 2}}\frac{\bar{\delta} N\lambda^2}{2\lambda_J\lambda_f}\frac{\lambda_J^2 za^{\Diamond 2}}{2\lambda^2} = 1.\end{aligned}$$

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

EC.3.4. The CLC Contract

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

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

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

$$\begin{aligned}\min_{\nu,\beta,f\geq 0} \quad & \mathbb{E}_{k,l}[\nu a_k^\Diamond + \beta e_l^\Diamond + fN\delta(a_k^\Diamond, e_l^\Diamond, f)], && && \text{(EC.27)}\\ \text{s.t.} \quad & \delta(a_k^\Diamond, e_l^\Diamond, f) \geq \bar{\delta}, && \forall k,l \in \{L,H\}, && \text{(EC.28)}\\ IR: \ & \nu a_k^\Diamond - \theta_I^k(a_k^\Diamond;\theta_I^k) \geq 0, && \forall k \in \{L,H\}, && \text{(EC.29)}\end{aligned}$$

$$IC: a_k^{\Diamond} \in \arg\max_{a \geq \underline{a}} \left\{ \nu a - \Theta_{I}(a;\theta_{I}^{k}) \right\}, \qquad \forall k \in \{L,H\}, \tag{EC.30}$$

$$IR: \beta e_l^{\Diamond} - \Theta_{C}(e_l^{\Diamond};\theta_{C}^{l}) \geq 0, \qquad \forall l \in \{L,H\}, \tag{EC.31}$$

$$IC: e_l^{\Diamond} \in \arg\max_{e \geq \underline{e}} \left\{ \beta e - \Theta_{C}(e;\theta_{C}^{l}) \right\}, \qquad \forall l \in \{L,H\}. \tag{EC.32}$$

EC.3.4.1. 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 $\beta_k \geq \frac{\theta_C^H \underline{e}_k}{2}$. 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$_{CLSI}$

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

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., $a_k^{\Diamond} = \max\left\{\underline{a}, \frac{\nu}{\theta_I^k}\right\}, k \in \{L,H\}$), (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, $f^{\Diamond} = \frac{\bar{\delta} - \lambda(a_H^{\Diamond} + e_H^{\Diamond}) - \lambda_J a_H^{\Diamond} e_H^{\Diamond}}{\lambda_f}$. Thus, the sponsor's problem reduces to the Problem P$_{CLOM}$ below:

Problem P$_{CLOM}$

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

□

EC.3.5. 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 $\underline{e}_k$ and $\Pi_{2k}^F(\underline{e}_k)$; 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 $(s_k = \frac{{\beta^\circ_k}^2}{\theta^H_C}, a_k = a^\circ_k, f_k = f^\circ_k)$. 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 $\Pi^N_{2k} \geq \Pi^F_{2k}(e^\circ_{kH})$. Recall that $\Pi^N_{2k} \geq \Pi^F_{2k}(\underline{e}_k)$ if and only if $\underline{e}_k \in [e_1, e_2]$. Therefore, we must have $e^\circ_{kH} \geq e_1$.

Now consider an CLC contract with sponsor's decisions $(\beta^\circ_k, a^\circ_k, f^\circ_k)$, and lower bound on effort $\underline{e}_k$. Recall $\hat{e}^\circ_{kl} = \max\{\frac{\beta^\circ_k}{\theta^l_C}, \underline{e}_k\}$. From the arguments above $\hat{e}^\circ_{kl} = \frac{\beta^\circ_k}{\theta^l_C} = e^\circ_{kl}$. 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 $\Pi^{\hat{N}}_{2k}(\underline{e}_k) \leq \Pi^N_{2k}$. Combining with Lemma EC.2, we have $\Pi^{\hat{N}}_{2k}(\underline{e}_k) = \Pi^N_{2k}$ when $\underline{e}_k \leq e_1$. Hence, the result $\Pi^{\hat{N}}_{2k}(\underline{e}_k) = \Pi^N_{2k} \leq \Pi^F_{2k}(\underline{e}_k)$ 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 $\min\{\Pi^F_{2k}(\underline{e}_k), \Pi^N_{2k}\} = \Pi^F_{2k}(\underline{e}_k)$. Therefore, from Lemma EC.2, $\Pi^{\hat{N}}_{2k}(\underline{e}_k) \geq \min\{\Pi^F_{2k}(\underline{e}_k), \Pi^N_{2k}\} = \Pi^F_{2k}(\underline{e}_k)$.

(ii) $\underline{e}_k \geq e_2$: Let $(\hat{\beta}^\circ_k, \hat{a}^\circ_k, \hat{f}^\circ_k)$ 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 $\hat{e}^\circ_{kl} = \max\left\{\frac{\hat{\beta}^\circ_k}{\theta^l_C}, \underline{e}_k\right\}$. Note that from the IR constraint of the high-type coordinator, we have $\hat{\beta}^\circ_k \hat{e}^\circ_{kH} - \frac{1}{2}\theta^H_C {\hat{e}^\circ_{kH}}^2 \geq 0$. The sponsor's optimal expected retention cost given the CLC contract is

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

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

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

Given $\underline{e}_k$, let $(\bar{s}^\circ_k, \bar{a}^\circ_k, \bar{f}^\circ_k)$ be the sponsor's optimal solution under the FC contract. Recall that from Proposition 8, the optimal compensation under the FC contract $\bar{s}^\circ_k = \frac{1}{2}\theta^H_C \underline{e}_k^2$. 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.3.6. 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 those 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 $\nu > 0$, $\beta = \frac{1}{2}\theta_C^H \underline{e}$, $f = \bar{f}^\Diamond - \frac{\lambda + \lambda_f \underline{e}}{\lambda_f \theta_I^H}\nu$. 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

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

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

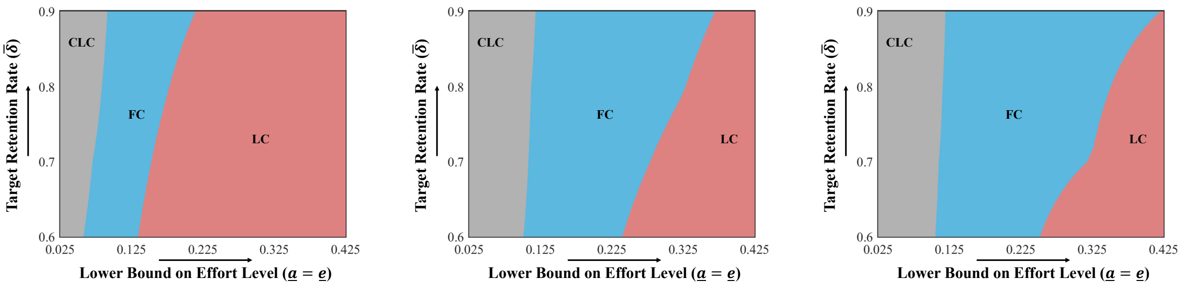
$$\Pi_3^{\hat{N}}(\underline{a},\underline{e}) < \Pi_3^F(\underline{a},\underline{e}), \quad \Pi_3^{\hat{N}}(\underline{a},\underline{e}) < \Pi_3^N, \qquad \forall 0\leq\underline{a}\leq\alpha_0, \sigma_1\leq\underline{e}\leq\sigma_2.$$

The result now follows.

EC.4. Computational Study Results for the OM Model

EC.4.1. Understanding the Relative Performance of the FC, LC, and CLC Contracts

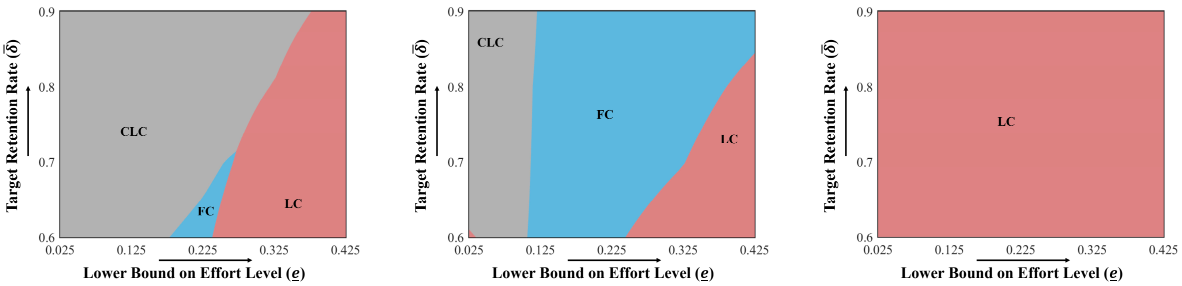
Figure EC.1 illustrates the impact of the lower bound on effort, the target retention rate, and the effectiveness of effort on the best-performing contract among the FC, LC, and CLC contracts under the OM model when $\underline{a} = \underline{e}$. We observe that the cost of the CLC contract is lower 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.



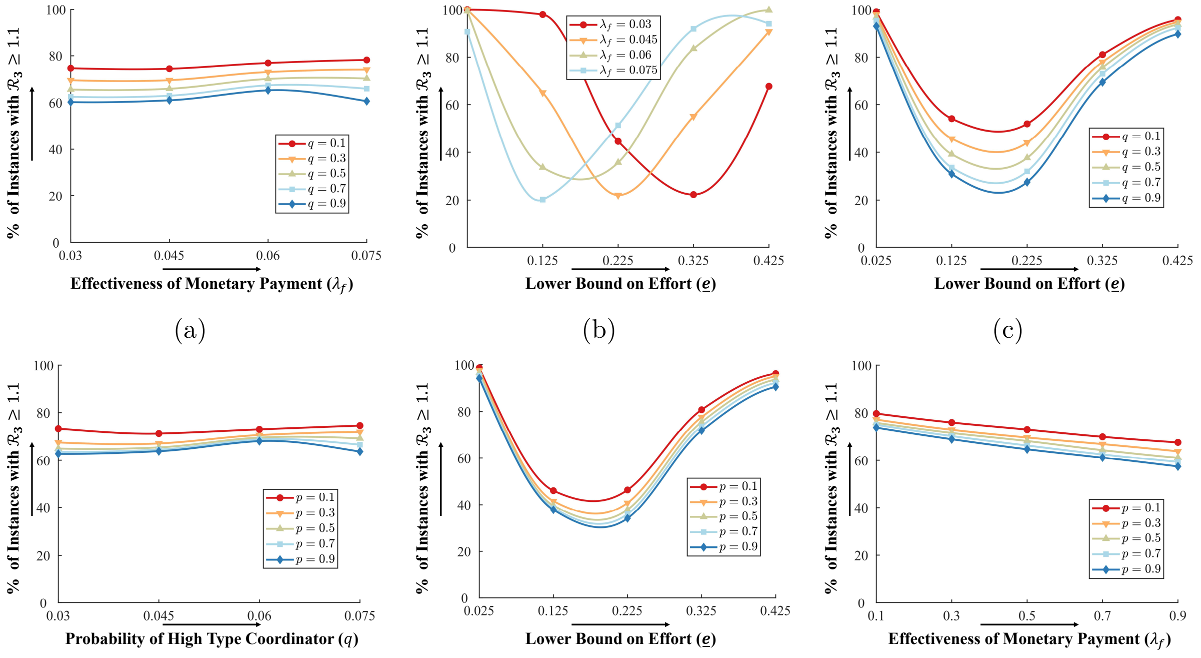
(a) $\underline{a} = 0.025$ (b) $\underline{a} = 0.225$ (c) $\underline{a} = 0.425$

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

EC.4.2. 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.



(a) (b) (c)

(d) (e) (f)

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

EC.5. Model Extensions

To test the robustness of our insights, we consider the following three extensions of our models in the main paper: (i) uncertainty in the retention rate, (ii) continuous distribution of effort cost parameters, and (iii) moral hazard issue.

EC.5.1. 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.

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,

$$Prob\left(\tilde{\delta}(a,e,f) \geq \bar{\delta}\right) \geq \zeta. \tag{EC.37}$$

Solving the sponsor's problem with retention constraint (EC.37) 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.

EC.5.1.1. Computational Analysis. We consider the test bed described in Section 5.1 and assume that ϕ follows a multivariate distribution with mean $\mu = (\lambda, \lambda_J, \lambda_f, 0)$ and the covariance matrix $\Sigma = \text{diag}(\alpha^2\lambda^2, \alpha^2\lambda_J^2, \alpha^2\lambda_f^2, \alpha^2)$, where values of $\lambda, \lambda_J, \lambda_f$ are as specified in Table 2 (Section 5.1). 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.

α	Centralized Model	SI Model				OM Model			
		Optimal Contract	FC contract	LC contract	CLC contract	Optimal Contract	FC contract	LC contract	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 EC.1 Percentage Increase in Expected Retention Cost for Different Levels of Uncertainty

Table EC.1 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 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

$\underline{a} = \underline{e}$	FC Contract		CLC Contract	
	SI Model	OM Model	SI Model	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 EC.2 Percentage Increase in Expected Retention Cost Due to Uncertainty for Different Lower Bounds on Effort

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 EC.2). For a given value(s) of the lower bound(s), the value 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 (EC.37) 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.

In summary, uncertainties in effectiveness parameters and due to unobserved factors do not significantly impact the cost performance of the FC and the CLC contracts. In addition, these two contracts are more robust to uncertainties for the higher values of the lower bound on the effort. 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 of those insights here.

EC.5.2. A Continuous Distribution of Effort Cost Parameters

In this section, we extend our models to consider a continuum of investigator (resp., coordinator's) types in the range $[\underline{\theta}_I, \bar{\theta}_I]$ (resp., $[\underline{\theta}_C, \bar{\theta}_C]$). Recall from Section 3, the type of a provider represents his effort cost parameter. Thus, we assume that there is a monotone ordering of types for each provider such that an ordering of types implies an ordering of effort cost parameters. We assume that the investigator and the coordinator's types (θ_I and θ_C, respectively) are independent random draws from their respective type intervals with cumulative distribution functions G and H, respectively. Further, as before, whether the sponsor knows the types of providers or only the distribution of types depends on the model (CM, SI, or OM) considered. In this section, we use $\phi_I = \frac{g}{G}$ and $\phi_C = \frac{f}{F}$ to denote the reverse hazard rate function of θ_I and θ_C, respectively, where g and f are the corresponding probability density functions. We further make the following two assumptions to focus on practical settings and tractability:

Assumption EC.1. *We assume*

1. $\min\{\underline{\theta}_I, \underline{\theta}_C\} > \frac{\max\{\bar{\delta}N(\lambda_J+\lambda), 2N(\bar{\delta}\lambda_J+\lambda^2)\}}{\lambda_f}$.
2. $\frac{\partial \phi_I}{\partial \theta_I} \leq 0$, $\frac{\partial \phi_C}{\partial \theta_C} \leq 0$.

The first statement of the assumption is similar to Assumption 1 in Section 3.1. The second part of the assumption is a regularity condition for the effort cost parameter distributions. This is a common assumption in the literature (see, e.g., Bolton 2005, Laffont and Martimort 2002). Many common distributions, such as the exponential, Weibull, and log-normal distributions, satisfy this assumption (Block et al. 1998). Next, we consider the sponsor's optimal decisions under the CM, SI, and OM models, respectively.

EC.5.2.1. The Centralized (CM) Model. We first discuss the optimal solution under the CM model. Recall that under this model, the sponsor knows the providers' types. Thus, given the effort cost parameters θ_I and θ_C, the expressions for the optimal decisions are the same as the ones derived in Proposition 1. Consequently, the results in Corollary 1 also hold under this extension.

EC.5.2.2. Decentralized Models. Next, we identify the sponsor's optimal contracts under the two decentralized models.

The Sponsor-Investigator (SI) Model: Under this model, the sponsor knows the investigator's type but does not know that of the coordinator. Using the revelation principle, we assume that the sponsor offers a direct mechanism to the coordinator, which consists of a set $[\underline{\theta}_C, \bar{\theta}_C]$ of type reports, a compensation with $T : [\underline{\theta}_C, \bar{\theta}_C] \to \mathbb{R}$, and effort level $e : [\underline{\theta}_C, \bar{\theta}_C] \to [0,1]$. Given the value of θ_I, the sponsor's problem is

$$
\begin{aligned}
\min_{a,e,T,f} \quad & \Pi^{SI}(\theta_I) = \mathbb{E}_{\theta_C}[\Theta_I(a(\theta_C);\theta_I) + T(\theta_C) + f(\theta_C)N\delta(a(\theta_C),e(\theta_C),f(\theta_C))], && && \text{(EC.38)}\\
\text{s.t.} \quad & \delta(a(\theta_C),e(\theta_C),f(\theta_C)) \geq \bar{\delta}, && \forall \theta_C \in [\underline{\theta}_C, \bar{\theta}_C], && \text{(EC.39)}\\
& T(\theta_C) - \Theta_C(e(\theta_C);\theta_C) \geq T(\theta'_C) - \Theta_C(e(\theta'_C);\theta_C), && \forall \theta_C, \theta'_C \in [\underline{\theta}_C, \bar{\theta}_C], && \text{(EC.40)}\\
& T(\theta_C) - \Theta_C(e(\theta_C);\theta_C) \geq 0, && \forall \theta_C \in [\underline{\theta}_C, \bar{\theta}_C]. && \text{(EC.41)}
\end{aligned}
$$

Let $\hat{\theta}_C = \theta_C + \frac{H(\theta_C)}{h(\theta_C)}$. The following proposition characterizes the sponsor's optimal solution under the SI model.

Proposition EC.1. *Given the investigator's effort cost parameter θ_I, the optimal solution to the SI model is as follows:*

1. $a^\circ(\theta_C) = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J + \lambda_f\hat{\theta}_C)}{\lambda_f^2\theta_I\hat{\theta}_C - \bar{\delta}^2N^2\lambda_J^2}, \quad e^\circ(\theta_C) = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J + \lambda_f\theta_I)}{\lambda_f^2\theta_I\hat{\theta}_C - \bar{\delta}^2N^2\lambda_J^2},$
2. $T(\theta_C) = \int_{\theta_C}^{\bar{\theta}_C} \frac{\partial\Theta_C(e^\circ(\theta);\theta)}{\partial\theta} d\theta + \Theta_C(e^\circ(\theta_C);\theta_C),$
3. $f^\circ(\theta_C) = \frac{\bar{\delta} - \lambda(a^\circ(\theta_C) + e^\circ(\theta_C)) - \lambda_J a^\circ(\theta_C)e^\circ(\theta_C)}{\lambda_f},$

where $\theta_C \in [\underline{\theta}_C, \bar{\theta}_C]$ is the coordinator's reported type.

As discussed in Section 3.2, to ensure the coordinator truthfully reveals his type, the sponsor provides information rent to all types of coordinators other than the one with the highest cost parameter, and the coordinator's optimal effort is lower as compared to the CM model unless $\theta_C = \underline{\theta}_C$. Our results in Corollary 2 with respect to $\bar{\delta}$ from the main paper also extend to the continuous type space.

The Outsourcing (OM) Model: Under the OM model, the sponsor does not observe the investigator's and the coordinator's type. As before, the sponsor offers a direct mechanism to the investigator and the coordinator, which consists of sets $[\underline{\theta}_I, \bar{\theta}_I]$, $[\underline{\theta}_C, \bar{\theta}_C]$ of type reports for the investigator and the coordinator, respectively, compensations with $T_I, T_C : [\underline{\theta}_I, \bar{\theta}_I] \times [\underline{\theta}_C, \bar{\theta}_C] \to \mathbb{R}$, and effort levels $a, e : [\underline{\theta}_I, \bar{\theta}_I] \times [\underline{\theta}_C, \bar{\theta}_C] \to [0,1]$. The sponsor's problem is then

$$
\begin{aligned}
\min_{T_I, T_C, a, e, f} \quad & \Pi^{OM} = \mathbb{E}_{\theta_I,\theta_C}[T_I(\theta_I,\theta_C) + T_C(\theta_I,\theta_C) + f(\theta_I,\theta_C)N\delta(a(\theta_I,\theta_C), e(\theta_I,\theta_C), f(\theta_I,\theta_C))],\\
& \text{s.t.}
\end{aligned}
$$

$$\begin{aligned}
&\delta(a(\theta_I,\theta_C),e(\theta_I,\theta_C),f(\theta_I,\theta_C)) \geq \bar{\delta}, && \forall \theta_I \in [\underline{\theta}_I,\bar{\theta}_I], \theta_C \in [\underline{\theta}_C,\bar{\theta}_C],\\
&T_C(\theta_I,\theta_C) - \Theta_C(e(\theta_I,\theta_C);\theta_C) \geq T_C(\theta_I,\theta_C') - \Theta_C(e(\theta_I,\theta_C');\theta_C), && \forall \theta_I \in [\underline{\theta}_I,\bar{\theta}_I], \forall \theta_C,\theta_C' \in [\underline{\theta}_C,\bar{\theta}_C],\\
&T_C(\theta_I,\theta_C) - \Theta_C(e(\theta_I,\theta_C);\theta_C) \geq 0, && \forall \theta_I \in [\underline{\theta}_I,\bar{\theta}_I], \theta_C \in [\underline{\theta}_C,\bar{\theta}_C],\\
&T_I(\theta_I,\theta_C) - \Theta_C(e(\theta_I,\theta_C);\theta_I) \geq T_I(\theta_I',\theta_C) - \Theta_C(e(\theta_I',\theta_C);\theta_I), && \forall \theta_I,\theta_I' \in [\underline{\theta}_I,\bar{\theta}_I], \theta_C \in [\underline{\theta}_C,\bar{\theta}_C],\\
&T_I(\theta_I,\theta_C) - \Theta_I(a(\theta_I,\theta_C);\theta_I) \geq 0, && \forall \theta_I \in [\underline{\theta}_I,\bar{\theta}_I], \theta_C \in [\underline{\theta}_C,\bar{\theta}_C].
\end{aligned}$$

We define $\hat{\theta}_I = \theta_I + \frac{G(\theta_I)}{g(\theta_I)}$. The sponsor's optimal decisions are as follows:

Proposition EC.2. *The optimal solution to the OM model is as follows:*

1. $a^{\Diamond}(\theta_I,\theta_C) = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J+\lambda_f\hat{\theta}_C)}{\lambda_f^2\hat{\theta}_C\hat{\theta}_C-\bar{\delta}^2N^2\lambda_J^2}$, $\quad e^{\Diamond}(\theta_I,\theta_C) = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J+\lambda_f\hat{\theta}_I)}{\lambda_f^2\hat{\theta}_I\hat{\theta}_C-\bar{\delta}^2N^2\lambda_J^2}$,
2. $T_I(\theta_I,\theta_C) = \int_{\theta_I}^{\bar{\theta}_I} \frac{\partial\Theta_I(a^{\Diamond}(\theta,\theta_C);\theta)}{\partial\theta} d\theta + \Theta_I(a^{\Diamond}(\theta_I,\theta_C);\theta_I)$, $\quad T_C(\theta_I,\theta_C) = \int_{\theta_C}^{\bar{\theta}_C} \frac{\partial\Theta_C(e^{\Diamond}(\theta_I,\theta);\theta)}{\partial\theta} d\theta + \Theta_C(e^{\Diamond}(\theta_I,\theta_C);\theta_C)$,
3. $f^{\Diamond}(\theta_I,\theta_C) = \frac{\bar{\delta}-\lambda(a^{\Diamond}(\theta_I,\theta_C)+e^{\Diamond}(\theta_I,\theta_C))-\lambda_J a^{\Diamond}(\theta_I,\theta_C)e^{\Diamond}(\theta_I,\theta_C)}{\lambda_f}$,

where $\theta_I \in [\underline{\theta}_I,\bar{\theta}_I]$ *is the investigator's reported type and* $\theta_C \in [\underline{\theta}_C,\bar{\theta}_C]$ *is the coordinator's reported type.*

Under the OM model, to ensure truthful revealing from both the investigator and the coordinator, the sponsor provides information rent to both, and our discussion from Section 3.3 still holds. Furthermore, it is easy to observe that our results in Propositions 4 and 5 on comparing three models, and the impact of $\lambda, \lambda_J, \lambda_f$, and $\bar{\delta}$ on the cost difference among the models (Proposition 6) also extend to the continuous type space.

In the following section, we discuss the sponsor's optimal decisions under compensation contracts in practice.

EC.5.2.3. Compensation Contracts in Practice.

Let $\underline{a}$ (resp., $\underline{e}$) be the lower bound on effort. The following proposition establishes the properties of the optimal decisions under the FC, LC, and CLC contracts. The results are similar to those presented in Section 4.

Proposition EC.3. *When the providers' effort cost parameters are drawn from continuous distributions, the sponsor's optimal decisions under the SI and OM models are as follows:*

1. *When the sponsor adopts the FC contract,*
    - *Under the SI model, given the investigator's effort cost parameter* θ_I*, the sponsor's optimal decisions are:* $s^{\circ}(\theta_I) = \frac{1}{2}\bar{\theta}_C\underline{e}^2$, $a^{\circ}(\theta_I) = \frac{\bar{\delta}N(\lambda+\lambda_J\underline{e})}{\lambda_f\theta_I}$, $f^{\circ}(\theta_I) = \frac{(\bar{\delta}-\lambda\underline{e})\lambda_f\theta_I-\bar{\delta}N(\lambda+\lambda_J\underline{e})^2}{\lambda_f^2\theta_I}$*. The coordinator's optimal effort is:* $e^{\circ} = \underline{e}$ *regardless of his type.*
    - *Under the OM model, the sponsor's optimal decisions are:* $r^{\Diamond} = \frac{1}{2}\bar{\theta}_I\underline{a}^2$, $s^{\Diamond} = \frac{1}{2}\bar{\theta}_C\underline{e}^2$, *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* $a^{\Diamond} = \underline{a}$, $e^{\Diamond} = \underline{e}$.
2. *When the sponsor adopts the LC contract,*

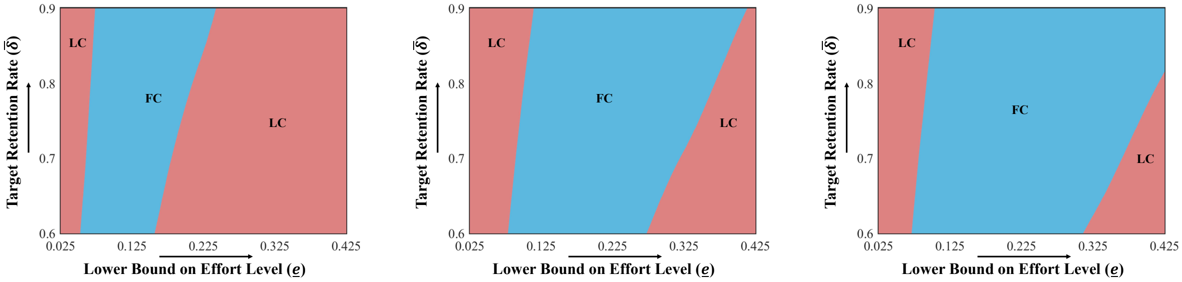
- *Under the SI model, given the investigator's effort cost parameter* θ_I*, the sponsor's optimal decisions are:* $f^\circ(\theta_I) = \frac{\bar{\delta}-\lambda(a^\circ(\theta_I)+e^\circ(\theta_I,\bar{\theta}_C))-\lambda_J a^\circ(\theta_I)e^\circ(\theta_I,\bar{\theta}_C)}{\lambda_f}$*. The optimal effort of the coordinator with cost parameter* θ_C *is* $e^\circ(\theta_I,\theta_C) = \frac{\beta^\circ(\theta_I)}{\theta_C}$*.*
- *Under the OM model, the sponsor's optimal monetary payment to the participant satisfies* $f^\diamondsuit = \frac{\bar{\delta}-\lambda(a^\diamondsuit(\bar{\theta}_I,\bar{\theta}_C)+e^\diamondsuit(\bar{\theta}_I,\bar{\theta}_C))-\lambda_J a^\diamondsuit(\bar{\theta}_I,\bar{\theta}_C)e^\diamondsuit(\bar{\theta}_I,\bar{\theta}_C)}{\lambda_f}$*. The optimal effort of the investigator with cost parameter* θ_I *is* $a^\diamondsuit(\theta_I) = \frac{\nu^\diamondsuit}{\theta_I}$*, and the optimal effort of the coordinator with cost parameter* θ_C *is* $e^\diamondsuit(\theta_C) = \frac{\beta^\diamondsuit}{\theta_C}$*.*

3. *When the sponsor adopts the CLC contract,*
    - *Under the SI model, given the investigator's effort cost parameter* θ_I*, the sponsor's optimal decisions are:* $f^\circ(\theta_I) = \frac{\bar{\delta}-\lambda(a^\circ(\theta_I)+e^\circ(\theta_I,\bar{\theta}_C))-\lambda_J a^\circ(\theta_I)e^\circ(\theta_I,\bar{\theta}_C)}{\lambda_f}$*. The optimal effort of the coordinator with cost parameter* θ_C *is* $e^\circ(\theta_I,\theta_C) = \max\left\{\underline{e}, \frac{\beta^\circ(\theta_I)}{\theta_C}\right\}$*.*
    - *Under the OM model, the sponsor's optimal monetary payment to the participant satisfies* $f^\diamondsuit = \frac{\bar{\delta}-\lambda(a^\diamondsuit(\bar{\theta}_I,\bar{\theta}_C)+e^\diamondsuit(\bar{\theta}_I,\bar{\theta}_C))-\lambda_J a^\diamondsuit(\bar{\theta}_I,\bar{\theta}_C)e^\diamondsuit(\bar{\theta}_I,\bar{\theta}_C)}{\lambda_f}$*. The optimal effort of the investigator with cost parameter* θ_I *is* $a^\diamondsuit(\theta_I) = \max\left\{\underline{a}, \frac{\nu^\diamondsuit}{\theta_I}\right\}$*, and the optimal effort of the coordinator with cost parameter* θ_C *is* $e^\diamondsuit(\theta_C) = \max\left\{\underline{e}, \frac{\beta^\diamondsuit}{\theta_C}\right\}$*.*

We rely on the computational study to numerically identify the optimal values of $\beta^\circ(\theta_I), a^\circ(\theta_I)$ under the SI model, and $\beta^\diamondsuit, \nu^\diamondsuit$ under the OM model. In the following section, we discuss our findings from the computational study.

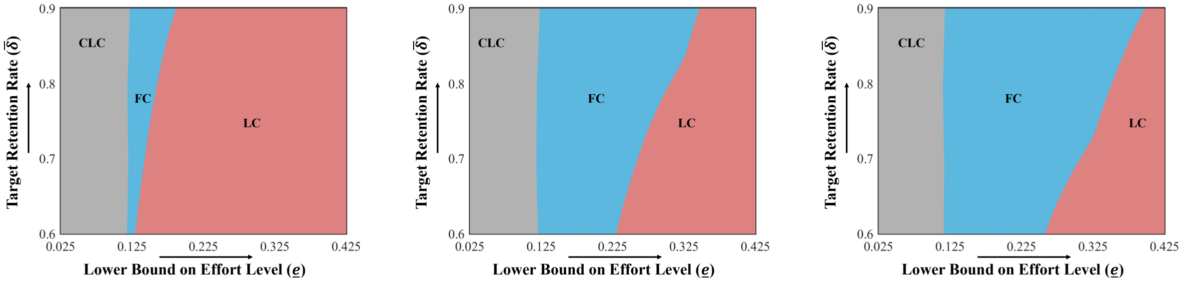
EC.5.2.4. Computational Study. In the computational study, our goal is to understand the performance of the three contracts observed in practice relative to each other and to the optimal contract. We assume that the providers' effort cost parameters follow a uniform distribution, i.e., $\theta_C \sim U[\underline{\theta}_C, \bar{\theta}_C], \theta_I \sim U[\underline{\theta}_I, \bar{\theta}_I]$. The parameter values are set similarly to that in Section 5. In particular, we set $\underline{\theta}_I = 6000$ and select $\frac{\bar{\theta}_I}{\underline{\theta}_I} \in \{1.05, 1.15, 1.25, 1.35\}$. We further consider $\frac{\underline{\theta}_C}{\underline{\theta}_I} \in \{0.4, 0.6, 0.8, 1.0\}$, and choose the values of $\bar{\theta}_C$ such that $\frac{\bar{\theta}_C - \underline{\theta}_C}{\bar{\theta}_I - \underline{\theta}_I} \in \{0.5, 1.5, 2.5, 3.5\}$. The rest of the parameters are the same as in Section 5.1. The computational study (Figures EC.4 and EC.5) shows that our findings in Section 5 are robust to the continuous type space for the effort-cost parameters.

Relative Performance of the FC, LC, and CLC Contracts:



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

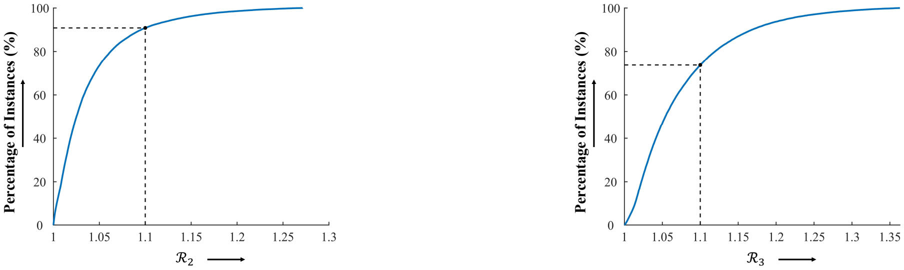
Figure EC.4 Comparing the Performances of the FC, LC, and CLC Contracts



(d) OM Model, $\lambda = 0.4$ (e) OM Model, $\lambda = 0.7$ (f) OM Model, $\lambda = 1.0$

Figure EC.4 Comparing the Performances of the FC, LC, and CLC Contracts (Continued)

Performance Relative to the Optimal Contract:



(a) SI Model (b) OM Model

Figure EC.5 Cumulative Distribution of $\mathcal{R}_2$ and $\mathcal{R}_3$

The ratios $\mathcal{R}_2 = \frac{\min_{t\in\{F,N,\hat{N}\}} \Pi_2^t}{\Pi^{SI\diamond}}$, $\mathcal{R}_3 = \frac{\min_{t\in\{F,N,\hat{N}\}} \Pi_3^t}{\Pi^{OM\diamond}}$. $\mathbb{E}\mathcal{R}_2 = 1.04$, $\mathbb{E}\mathcal{R}_3 = 1.08$; $\text{Prob}(\mathcal{R}_2 < 1.10) = 91\%$, $\text{Prob}(\mathcal{R}_3 < 1.10) = 74\%$.

EC.5.3. Moral Hazard Issue

Although the providers' effort is observable and contractable in the practical context of clinical studies, for theoretical completeness of the analysis, we extend our analysis to consider a setting where effort is unobservable (moral hazard issue) but the providers' type (effort cost parameter) is known. In this extension, similar to Appendix EC.5.1, we consider that unobserved factors may influence the retention rate. Hence, the achieved retention rate is random and defined as $\tilde{\delta}(a,e,f) = \lambda a + \lambda e + \lambda_J ae + \lambda_f f + \epsilon$, where ϵ is a random variable with support on $[-\kappa, \kappa]$ and represents the uncertainty resulting from unobserved factors. Let Ω (resp., ω) be the c.d.f. (resp., p.d.f.) of ϵ. We assume that the distribution of ϵ is common knowledge.

Under the CM model, for any given θ_I and θ_C, the sponsor's decision problem then becomes:

$$\begin{aligned} \min_{a,e,f} \quad & \frac{1}{2}\theta_I a^2 + \frac{1}{2}\theta_C e^2 + \tilde{\delta}(a,e,f)Nf, \\ \text{s.t.} \quad & Prob\left(\tilde{\delta}(a,e,f) \geq \bar{\delta}\right) \geq \zeta. \end{aligned} \tag{EC.42}$$

As discussed in Appendix EC.5.1, Constraint (EC.42) is equivalent to $\lambda a + \lambda e + \lambda_J ae + \lambda_f f \geq \bar{\delta} - \Omega^{-1}(1-\zeta)$, which is the same as Constraint (2) in Section 3.1 by replacing $\bar{\delta}$ with $[\bar{\delta} - \Omega^{-1}(1-\zeta)]$, and hence, we have the following result:

PROPOSITION EC.4. *Given the providers' effort cost parameter θ_I and θ_C, the sponsor's optimal decisions are: $a^* = \frac{\lambda(\bar{\delta}-\Omega^{-1}(1-\zeta))N((\bar{\delta}-\Omega^{-1}(1-\zeta))N\lambda_J+\lambda_f\theta_C)}{\lambda_f^2\theta_C\theta_I-(\bar{\delta}-\Omega^{-1}(1-\zeta))^2N^2\lambda_J^2}$, $e^* = \frac{\lambda(\bar{\delta}-\Omega^{-1}(1-\zeta))N((\bar{\delta}-\Omega^{-1}(1-\zeta))N\lambda_J+\lambda_f\theta_I)}{\lambda_f^2\theta_C\theta_I-(\bar{\delta}-\Omega^{-1}(1-\zeta))^2N^2\lambda_J^2}$, and $f^* = \frac{(\bar{\delta}-\Omega^{-1}(1-\zeta))-\lambda(a^*+e^*)-\lambda_J a^* e^*}{\lambda_f}$.*

Under the SI and OM models, the uncertainty in observed factors and the possibility that the sponsor may not observe the providers' *true* efforts result in a moral hazard issue faced by the sponsor. A common way to deal with moral hazard in the literature and practice (Simester and Zhang 2010, Siemsen et al. 2007) is to contract on outcomes. In our setting, the outcome is the achieved retention rate. However, as mentioned earlier, the sponsor cannot offer outcome-based compensation due to clinical study regulations. Thus, similar to Zhang et al. (2018) and Dey et al. (2010), we next analyze the SI and OM models under the setting where the sponsor contracts and offers compensation to the providers based on the *reported* effort levels. The analyses show that our main insights from the original model on (1) the relative performances of the three contracts in practice, and (2) their performances relative to the optimal contracts still hold qualitatively in this extension.

EC.5.3.1. The SI Model. Under the SI model, the sponsor offers the coordinator a compensation contract $T:[0,1]\to\mathbb{R}$ that is based on the reported effort $\hat{e}$, where $T(\hat{e})$ is non-decreasing in the reported effort $\hat{e}$. The sequence of events is as follows: first, the sponsor announces the compensation contract T and the payment to the participant f and chooses the investigator's effort (a). Second, the coordinator chooses the effort level e. Third, the achieved retention rate is observed by the sponsor and the coordinator (i.e., the uncertainty is resolved). Finally, the coordinator reports his effort $\hat{e}$ and receives compensation $T(\hat{e})$. Next, we present the coordinator's optimal decision on the reported effort level.

LEMMA EC.3. *Given the value of a, f, and the observed retention rate (say $\hat{\delta}$), the coordinator's reported effort level in the optimal solution is $\frac{\hat{\delta}+\kappa-\lambda a-\lambda_f f}{\lambda+\lambda_J a}$.*

Given the random variable $\epsilon \in [-\kappa,\kappa]$, it is easy to verify the feasible range of $\hat{e}$ is $\left[\frac{\hat{\delta}-\kappa-\lambda a-\lambda_f f}{\lambda+\lambda_J a}, \frac{\hat{\delta}+\kappa-\lambda a-\lambda_f f}{\lambda+\lambda_J a}\right]$; reporting any effort outside the range will signal the coordinator is misreporting. Since the compensation is based on the reported effort and is non-decreasing in $\hat{e}$, it is optimal for the coordinator to report $\frac{\hat{\delta}+\kappa-\lambda a-\lambda_f f}{\lambda+\lambda_J a}$.

Note that for a realized value of ϵ, we can write $\frac{\hat{\delta}+\kappa-\lambda a-\lambda_f f}{\lambda+\lambda_J a}$ as $e + \frac{\epsilon+\kappa}{\lambda+\lambda_J a}$, where e is the true effort level. Thus, $\hat{e}(e) = e + \frac{\epsilon+\kappa}{\lambda+\lambda_J a}$. Since at the time of deciding true effort, uncertainty is not resolved, the sponsor's decision problem is:

$$\min_{T,a,f} \quad \mathbb{E}_\epsilon\left[\frac{1}{2}\theta_I a^2 + T\left(\hat{e}(e^\triangle)\right) + \tilde{\delta}(a, e^\triangle, f)Nf\right], \tag{EC.43}$$

$$\text{s.t.} \quad Prob\left(\tilde{\delta}(a, e^{\triangle}, f) \geq \bar{\delta}\right) \geq \zeta, \tag{EC.44}$$

$$e^{\triangle} = \arg\max_{e\in[0,1]} \mathbb{E}_{\epsilon}\left[T\left(\hat{e}(e)\right) - \frac{1}{2}\theta_C e^2\right], \tag{EC.45}$$

$$\mathbb{E}_{\epsilon}\left[T\left(\hat{e}(e^{\triangle})\right) - \frac{1}{2}\theta_C {e^{\triangle}}^2\right] \geq 0. \tag{EC.46}$$

Constraint (EC.44) is the retention chance constraint. Constraint (EC.45) ensures the coordinator's choice of effort maximizes his expected benefit, and Constraint (EC.46) is the ex-ante individual rationality (IR) constraint. The next proposition describes the optimal contract.

PROPOSITION EC.5. *Under the SI model with moral hazard, a two-part tariff contract* ($T(\hat{e}) = \beta\hat{e} + w$) *achieves the first best solution, where* $\beta = \theta_C e^*$ *and* $w = -\left[\frac{\kappa\theta_C e^*}{\lambda+\lambda_J a^*} + \frac{1}{2}\theta_C {e^*}^2\right]$, *and* e^* *and* a^* *are the optimal effort levels under the CM model.*

Next, we discuss how moral hazard affects the sponsor's optimal decisions under the three contracts adopted in practice.

Fixed Compensation (FC) Contract: Under the FC contract, the sponsor offers the coordinator a fixed compensation s and requires reported effort level $\hat{e}$ to be at least $\underline{e}$. Then, we have the following result.

LEMMA EC.4. *The coordinator's optimal true effort is* $\underline{e}$.

Lemma EC.4 shows that the coordinator always exerts the minimum effort level $\underline{e}$. This is straightforward because the coordinator must exert at least effort $\underline{e}$ to ensure the reported effort level $\left[e + \frac{\epsilon+\kappa}{\lambda+\lambda_J a}\right]$ is at least $\underline{e}$ regardless of the realization of ϵ, and there is no incentive for the coordinator to exert effort higher than $\underline{e}$ as the compensation is fixed. The following result characterizes the sponsor's optimal decisions under the FC contract.

PROPOSITION EC.6. *When the sponsor adopts the FC contract, the sponsor's optimal decisions are:* $s^\circ = \frac{1}{2}\theta_C \underline{e}^2$, $a^\circ = \frac{(\bar{\delta}-\Omega^{-1}(1-\zeta))N(\lambda+\lambda_J\underline{e})}{\lambda_f\theta_I}$, $f^\circ = \frac{(\bar{\delta}-\Omega^{-1}(1-\zeta)-\lambda\underline{e})\lambda_f\theta_I-(\bar{\delta}-\Omega^{-1}(1-\zeta))N(\lambda+\lambda_J\underline{e})^2}{\lambda_f^2\theta_I}$.

Linear Compensation (LC) Contract: Under the LC contract, the compensation is linear to the reported effort, i.e., $T(\hat{e}) = \beta\hat{e}$. We have the following result.

LEMMA EC.5. *Under the LC contract, given the value of* β, *the coordinator's optimal true effort level is* $\frac{\beta}{\theta_C}$.

Conditional Linear Compensation Contract: Here, the sponsor offers a linear compensation contract $T(\hat{e}) = \beta\hat{e}$, while requiring the reported effort $\hat{e} \geq \underline{e}$. The following result characterizes the coordinator's optimal decision.

LEMMA EC.6. *Under the CLC contract, given the value of* β, *the coordinator's optimal true effort level is* $\max\left\{\underline{e}, \frac{\beta}{\theta_C}\right\}$.

To summarize, when the sponsor adopts the FC contract under the moral hazard setting, the optimal solution of the sponsor is the same as when the true effort is observable and $\theta_C^H = \theta_C^L$ (see Section 4.1, Proposition 8). However, the sponsor's optimal solution under the LC and CLC contract may differ from the situation where the true effort is observable. In Appendix EC.5.3.3, we compare the performances of these contracts numerically. Next, we discuss the sponsor's optimal decisions under the OM model.

EC.5.3.2. The OM Model. In the OM model, the sponsor offers the compensation contract T_I and T_C to the investigator and the coordinator, respectively. The sequence of events is as follows: The sponsor first announces contracts T_I and T_C, and the payment to the participant f. Following that, the investigator and the coordinator choose their effort levels a and e simultaneously. Then, the retention rate $\hat{\delta}$ is observed. After that, the investigator and the coordinator report their effort levels $\hat{a}$ and $\hat{e}$ at the same time. Finally, the sponsor compensates the investigator and the coordinator for these reported effort levels. We consider the following contracts for the providers:

$$T_I(\hat{a}, \hat{e}; f, \hat{\delta}) = \nu\hat{a} + w_I - \nu\hat{a}\mathbb{1}\left\{(\delta_{\min}(\hat{a}, \hat{e}, f) > \hat{\delta}) \wedge (\hat{e} \leq \hat{a})\right\},$$
$$T_C(\hat{a}, \hat{e}; f, \hat{\delta}) = \beta\hat{e} + w_C - \beta\hat{e}\mathbb{1}\left\{(\delta_{\min}(\hat{a}, \hat{e}, f) > \hat{\delta}) \wedge (\hat{a} \leq \hat{e})\right\}.$$

where $\delta_{\min}(\hat{a}, \hat{e}, f) = (\lambda\hat{a} + \lambda\hat{e} + \lambda_J\hat{a}\hat{e} + \lambda_f f) - \kappa$ is the minimum retention rate that can be achieved for any reported values of $\hat{e}$ and $\hat{a}$, given the uncertainty in unobserved factors $\epsilon \in [-\kappa, \kappa]$, and payment f.

When $\delta_{\min}(\hat{a}, \hat{e}, f) \leq \hat{\delta}$ the sponsor cannot infer that the providers are cheating. Hence, the above contracts reduce to two-part tariff contracts, where the compensations $(\nu\hat{a} + w_I, \beta\hat{e} + w_C)$ are based on the reported effort. Otherwise, the sponsor can infer that either one or both providers are not truthfully reporting their effort levels. However, the sponsor cannot distinguish whether the investigator or the coordinator is misreporting, and hence, the contract is such that the provider who reports a higher effort level is penalized. The penalty is represented by the third term in the contracts specified above. We next prove that the sponsor can achieve the centralized outcome by optimally designing the value of β, ν, w_I, w_C.

PROPOSITION EC.7. *Under the optimal contract, there is a unique subgame perfect equilibrium that attains the first-best outcome.*

Next, we discuss the sponsor's problems under the contracts observed in practice. To analyze the FC, LC, and CLC contracts under moral hazard, we make the following assumption: the sponsor offers compensations to both providers only if $\delta_{\min}(\hat{a}, \hat{e}, f) \leq \hat{\delta}$. We next examine the providers' reporting decisions.

Providers' Effort Reporting Decision: Given the observed retention rate $\hat{\delta}$, let $\mathcal{N}$ denote the set of possible equilibria representing the providers' effort reporting decisions. The following result shows the existence of infinitely many equilibria under the FC, LC, or CLC contract.

LEMMA EC.7. *For the FC, LC, and CLC contracts, there are infinitely many equilibria in the providers' reporting game. Further,*

$$\mathcal{N} \supseteq \{(\hat{a}, \hat{e}) \in [0,1]^2 : (\lambda \hat{a} + \lambda \hat{e} + \lambda_J \hat{a}\hat{e} + \lambda_f f) - \kappa = \hat{\delta}\}.$$

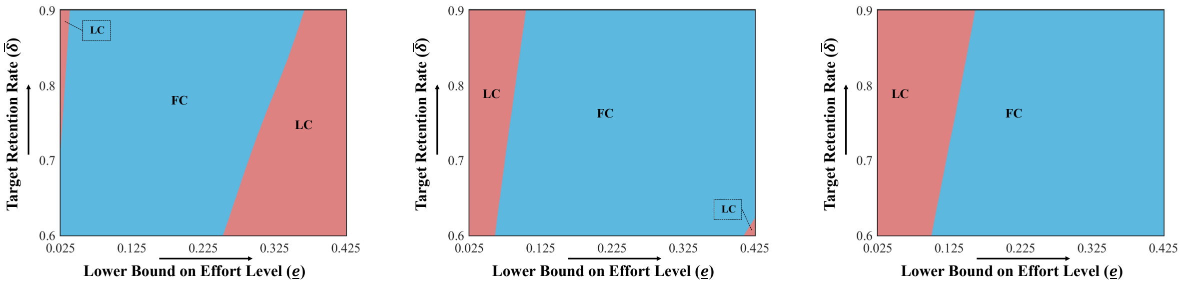
Given the result above, obtaining closed-form solutions for the three contracts is challenging. Hence, we resort to computational analysis to identify the equilibrium effort levels and the sponsor's optimal decisions and to analyze their performances. To focus on a unique equilibrium in our numerical analysis, we assume the equilibrium solution satisfies

$$\frac{\hat{a}}{a} = \frac{\hat{e}}{e}, (\hat{a}, \hat{e}) \in \mathcal{N}.$$

That is, we focus on an equilibrium solution where both providers misreport their effort levels by the same factor.

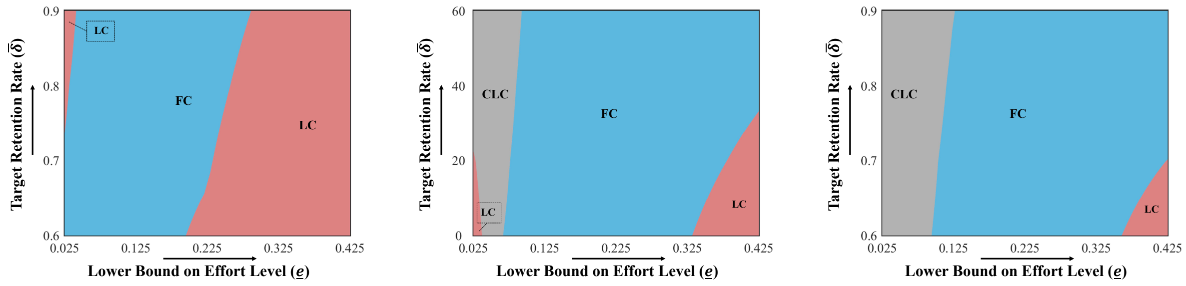
EC.5.3.3. Computational Study. We now conduct a computational study to compare the sponsor's expected retention cost under different contracts when they are contingent on the reported effort levels. We assume ϵ follows a uniform distribution on $[-\kappa, \kappa]$, where $\kappa \in \{0.025, 0.050, 0.075\}$. Note the providers' effort cost parameters are observable to the sponsor, i.e., there is a single type of investigator and coordinator. We set $\theta_I = \$6000$, which is the same as the cost parameter of the low-type investigator in the base model from the manuscript. We further consider $\frac{\theta_C}{\theta_I} \in \{0.4, 0.6, 0.8, 1.0\}$, which corresponds to the ratio of the low-type coordinator's cost parameter to that of the low-type investigator in the base model. The rest of the parameters are the same as in Section 5 of the manuscript. The analyses suggest that our insights derived in Section 5 still hold when the true effort is not observable.

Relative Performance of the FC, LC, and CLC Contracts:



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

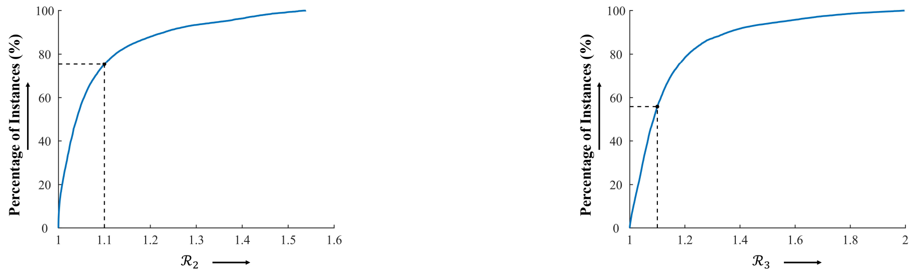
Figure EC.6 Comparing the Performances of the FC, LC, and CLC Contracts



(d) OM Model, $\lambda = 0.4$ (e) OM Model, $\lambda = 0.7$ (f) OM Model, $\lambda = 1.0$

Figure EC.6 Comparing the Performances of the FC, LC, and CLC Contracts (Continued)

Relative Performance of the FC, LC, and CLC Contracts:



(a) SI Model (b) OM Model

Figure EC.7 Cumulative Distribution of $\mathcal{R}_2$ and $\mathcal{R}_3$

The ratios $\mathcal{R}_2 = \frac{\min_{t\in\{F,N,\hat{N}\}} \Pi_2^t}{\Pi^{SI\circ}}$ and $\mathcal{R}_3 = \frac{\min_{t\in\{F,N,\hat{N}\}} \Pi_3^t}{\Pi^{OM\diamond}}$. $\mathbb{E}\mathcal{R}_2 = 1.08$, $\mathbb{E}\mathcal{R}_3 = 1.15$; Prob$(\mathcal{R}_2 < 1.10) = 75\%$, Prob$(\mathcal{R}_3 < 1.10) = 56\%$.

EC.6. Proof of Technical Results in Appendix EC.5.2

EC.6.1. Proof of Proposition EC.1

Similar to the proof of Proposition 2, we can show that the retention constraint is binding, which implies the monetary payment $f(\theta_C) = \frac{\bar{\delta} - \lambda(a(\theta_C) + e(\theta_C)) - \lambda_J a(\theta_C) e(\theta_C)}{\lambda_f}, \forall \theta_C$.

We next derive the sponsor's optimal decisions using the approach described in Bolton (2005) and Mirrlees (1971). Let $\theta^\circ(\theta_C)$ be the optimal reported type given the true type θ_C. Let $U_C(\theta_C) = [T(\theta^\circ(\theta_C)) - \Theta_C(e(\theta^\circ(\theta_C)); \theta_C)]$. Then, using the Envelope Theorem, we have

$$\frac{dU_C(\theta_C)}{d\theta_C} = \frac{\partial U_C(\theta_C)}{\partial \theta_C} = -\frac{\partial(\Theta_C(e(\theta^\circ(\theta_C)); \theta_C))}{\partial \theta_C}.$$

Further note, as the IC constraint implies, $\theta^\circ(\theta_C) = \theta_C$. Also, similar to the proof of Proposition 2, we can show that the IR constraint is only binding for $\bar{\theta}_C$. Thus, $U_C(\bar{\theta}_C) = 0$. Therefore,

$$U_C(\theta_C) = \int_{\theta_C}^{\bar{\theta}_C} \frac{\partial \Theta_C(e(\theta); \theta)}{\partial \theta} d\theta,$$

and the optimal compensation $T(\theta_C) = \int_{\theta_C}^{\bar{\theta}_C} \frac{\partial \Theta_C(e(\theta);\theta)}{\partial \theta} d\theta + \Theta_C(e(\theta_C); \theta_C)$.

Therefore, the sponsor's problem can be rewritten as

$$
\begin{aligned}
\min_{a,e} \quad & \mathbb{E}_{\theta_C}[\Theta_I(a(\theta_C);\theta_I) + T(\theta_C) + f(\theta_C)N\bar{\delta}]. \\
\text{s.t.} \quad & f(\theta_C) = \frac{\bar{\delta} - \lambda(a(\theta_C) + e(\theta_C)) - \lambda_J a(\theta_C)e(\theta_C)}{\lambda_f}, && \text{(EC.47)} \\
& T(\theta_C) = \int_{\theta_C}^{\bar{\theta}_C} \frac{\partial \Theta_C(e(\theta);\theta)}{\partial \theta} d\theta + \Theta_C(e(\theta_C);\theta_C), && \text{(EC.48)} \\
& \frac{\partial e(\theta_C)}{\partial \theta} \leq 0, \quad \forall \theta_C \in [\underline{\theta}_C, \bar{\theta}_C], && \text{(EC.49)}
\end{aligned}
$$

where Constraints (EC.48) and (EC.49) together ensure the original IC constraint (EC.40) is satisfied (Bolton 2005). Replacing $f(\theta_C), T(\theta_C)$ in the objective function, the sponsor's objective then becomes

$$
\min_{a,e} \quad \mathbb{E}_{\theta_C}\left[\Theta_I(a(\theta_C);\theta_I) + \int_{\theta_C}^{\bar{\theta}_C} \frac{\partial \Theta_C(e(\theta);\theta)}{\partial \theta} d\theta + \Theta_C(e(\theta_C);\theta_C) + \frac{\bar{\delta} - \lambda(a(\theta_C) + e(\theta_C)) - \lambda_J a(\theta_C)e(\theta_C)}{\lambda_f} N\bar{\delta}\right].
$$

Using integration by parts for the second term in the above expression, we can simplify the above objective as

$$
\begin{aligned}
\min_{a,e} \int_{\underline{\theta}_C}^{\bar{\theta}_C} \Big[\Theta_I(a(\theta_C);\theta_I) + & \frac{H(\theta_C)}{h(\theta_C)} \frac{\partial \Theta_C(e(\theta_C);\theta_C)}{\partial \theta_C} + \Theta_C(e(\theta_C);\theta_C) \\
& + \frac{\bar{\delta} - \lambda(a(\theta_C) + e(\theta_C)) - \lambda_J a(\theta_C)e(\theta_C)}{\lambda_f} N\bar{\delta}\Big] h(\theta_C) d\theta_C. \qquad \text{(EC.50)}
\end{aligned}
$$

Applying the Euler-Lagrange equations, the optimal $e^{\circ}(\theta_C)$ and $a^{\circ}(\theta_C)$ satisfy:

$$
\begin{aligned}
& \frac{H(\theta_C)}{h(\theta_C)} \frac{\partial^2 \Theta_C(e;\theta_C)}{\partial \theta_C \partial e} + \frac{\partial[\Theta_C(e;\theta_C)]}{\partial e} + \frac{\bar{\delta}N}{\lambda_f} \frac{\partial[\bar{\delta} - \lambda(a+e) - \lambda_J ae]}{\partial e} = 0, \\
& \frac{\partial \Theta_I(a;\theta_I)}{\partial a} + \frac{\bar{\delta}N}{\lambda_f} \frac{\partial[\bar{\delta} - \lambda(a+e) - \lambda_J ae]}{\partial a} = 0.
\end{aligned}
$$

Solving the above two equations, we have

$$
a^{\circ} = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J + \lambda_f \hat{\theta}_C)}{\lambda_f^2 \theta_I \hat{\theta}_C - \bar{\delta}^2 N^2 \lambda_J^2}, \quad e^{\circ} = \frac{\bar{\delta}N\lambda(\bar{\delta}N\lambda_J + \lambda_f \theta_I)}{\lambda_f^2 \theta_I \hat{\theta}_C - \bar{\delta}^2 N^2 \lambda_J^2}.
$$

Further, it can be verified that the Hessian matrix of (EC.50) with respect to (a, e) is positive semi-definite for any $\theta_C \in [\underline{\theta}_C, \bar{\theta}_C]$ under Assumption EC.1. Therefore, the functions a°, e° is the optimal solution to (EC.50).

Finally, note that $e^{\circ}(\theta_C)$ is non-increasing in θ_C, and hence, Constraint (EC.49) is satisfied. The result now follows. □

EC.6.2. Proof of Proposition EC.2

Similar to the proof of Proposition EC.1, we can find the optimal compensation contract T_I and T_C by solving the sponsor's problem:

$$
\begin{aligned}
T_I(\theta_I,\theta_C) &= \int_{\theta_I}^{\bar{\theta}_I} \frac{\partial \Theta_I(a(\theta,\theta_C);\theta)}{\partial \theta} d\theta + \Theta_I(a(\theta_I,\theta_C);\theta_I),\\
T_C(\theta_I,\theta_C) &= \int_{\theta_C}^{\bar{\theta}_C} \frac{\partial \Theta_C(e(\theta_I,\theta);\theta)}{\partial \theta} d\theta + \Theta_C(e(\theta_I,\theta_C);\theta_C).
\end{aligned}
$$

Replacing the terms in the sponsor's objective function and simplifying it, we get

$$
\begin{aligned}
\min_{a,e} \quad \int_{\underline{\theta}_I}^{\bar{\theta}_I} \int_{\underline{\theta}_C}^{\bar{\theta}_C} \Big[& \frac{G(\theta_I)}{g(\theta_I)} \frac{\partial \Theta_I(a(\theta_I,\theta_C);\theta_I)}{\partial \theta_I} + \Theta_I(a(\theta_I,\theta_C);\theta_I) + \frac{H(\theta_C)}{h(\theta_C)} \frac{\partial \Theta_C(e(\theta_I,\theta_C);\theta_C)}{\partial \theta_C} + \Theta_C(e(\theta_I,\theta_C);\theta_C)\\
& + \frac{\bar{\delta} - \lambda(a(\theta_I,\theta_C) + e(\theta_I,\theta_C)) - \lambda_J a(\theta_I,\theta_C) e(\theta_I,\theta_C)}{\lambda_f} N\bar{\delta} \Big] g(\theta_I) h(\theta_C) d\theta_C d\theta_I,
\end{aligned}
$$

It can be verified that the Hessian matrix of the above function with respect to (a,e) is positive semi-definite for any $\theta_I \in [\underline{\theta}_I, \bar{\theta}_I], \theta_C \in [\underline{\theta}_C, \bar{\theta}_C]$ under Assumption EC.1. Applying the Euler-Lagrange equations, we have

$$
\begin{aligned}
\frac{G(\theta_I)}{g(\theta_I)} \frac{\partial^2 \Theta_I(a;\theta_I)}{\partial \theta_I \partial a} + \frac{\partial \Theta_I(a;\theta_I)}{\partial a} + \frac{\bar{\delta} N}{\lambda_f} \frac{\partial[\bar{\delta} - \lambda(a+e) - \lambda_J ae]}{\partial a} &= 0,\\
\frac{H(\theta_C)}{h(\theta_C)} \frac{\partial^2 \Theta_C(e;\theta_C)}{\partial \theta_C \partial e} + \frac{\partial \Theta_C(e;\theta_C)}{\partial e} + \frac{\bar{\delta} N}{\lambda_f} \frac{\partial[\bar{\delta} - \lambda(a+e) - \lambda_J ae]}{\partial e} &= 0.
\end{aligned}
$$

Solving the above equations, we have

$$
a^{\Diamond} = \frac{\bar{\delta} N \lambda(\bar{\delta} N \lambda_J + \lambda_f \hat{\theta}_C)}{\lambda_f^2 \hat{\theta}_I \hat{\theta}_C - \bar{\delta}^2 N^2 \lambda_J^2}, \quad e^{\Diamond} = \frac{\bar{\delta} N \lambda(\bar{\delta} N \lambda_J + \lambda_f \hat{\theta}_I)}{\lambda_f^2 \hat{\theta}_I \hat{\theta}_C - \bar{\delta}^2 N^2 \lambda_J^2},
$$

where $\hat{\theta}_I = \theta_I + \frac{G(\theta_I)}{g(\theta_I)}$, $\hat{\theta}_C = \theta_C + \frac{H(\theta_C)}{h(\theta_C)}$. ☐

EC.6.3. Proof of Proposition EC.3

The proof is similar to that for the proofs of Propositions 8-10, and hence omitted. ☐

EC.7. Proof of Technical Results in Appendix EC.5.3

Before we prove the results in Appendix EC.5.3, we first prove the following technical result. For any given a, let $\mathscr{e}(a)$ denote the optimal solution to

$$
\max_{e \in [0,1]} \quad \beta \mathbb{E}\left[\frac{\sqrt{\lambda^2 + \lambda_J(\lambda(a+e) + \lambda_J ae + \kappa + \epsilon)} - \lambda}{\lambda_J} \right] + w_C - \frac{1}{2}\theta_C e^2, \tag{EC.51}
$$

where $a \in [0,1], \beta > 0$. Then we have the following result:

LEMMA EC.8. *The function $\mathscr{e}(a)$ is increasing concave in a, and $\mathscr{e}(a=0) > 0$.*

Proof It is straightforward to verify that for any given $a \in [0,1]$, the objective function (EC.51) is concave in e. Therefore, there exists a unique solution $\not{e}(a)$ to (EC.51). Let $t(a,e) = \beta\mathbb{E}\left[\frac{\lambda+\lambda_J a}{2\sqrt{\lambda^2+\lambda_J(\lambda(a+e)+\lambda_J ae+\kappa+\epsilon)}}\right] - \theta_C e$ be the first-order derivative of (EC.51) with respect to e given a. Then, by the definition of $\not{e}(a)$, $t(a,\not{e}(a))=0$. Further, it is straightforward to verify that

$$\begin{aligned}
\frac{\partial t(a,e)}{\partial a} &= \mathbb{E}\left[\frac{\beta\lambda_J\left(2(\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+e\lambda_J)\right)}{4\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+e\lambda_J)\right)^{3/2}}\right] > 0,\\
\frac{\partial t(a,e)}{\partial e} &= -\mathbb{E}\left[\frac{\beta\lambda_J(\lambda+a\lambda_J)^2}{4\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+e\lambda_J)\right)^{3/2}}\right] - \theta < 0,\\
\frac{\partial^2 t(a,e)}{\partial a^2} &= -\mathbb{E}\left[\frac{\beta\lambda_J^2(\lambda+e\lambda_J)\left(4(\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+e\lambda_J)\right)}{8\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+e\lambda_J)\right)^{5/2}}\right] < 0,\\
\frac{\partial^2 t(a,e)}{\partial e^2} &= \mathbb{E}\left[\frac{3\beta\lambda_J^2(\lambda+a\lambda_J)^3}{8\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+e\lambda_J)\right)^{5/2}}\right] > 0,\\
\frac{\partial^2 t(a,e)}{\partial e\partial a} &= -\mathbb{E}\left[\frac{\beta\lambda_J^2(\lambda+a\lambda_J)\left(4(\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+e\lambda_J)\right)}{8\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+e\lambda_J)\right)^{5/2}}\right] < 0.
\end{aligned}$$

Note that $t(a,e=0)>0$, and $\frac{\partial t(a,e)}{\partial e}<0, \forall a$. Therefore, $\not{e}(a)>0$. In particular, when $a=0$, we have $\not{e}(0)>0$.

Finally, we show $\not{e}(a)$ is concave in a. For simplicity, define $t_a(a,e) = \frac{\partial t(a,e)}{\partial a}, t_e(a,e) = \frac{\partial t(a,e)}{\partial e}$. Similarly, let $t_{aa}(a,e) = \frac{\partial^2 t(a,e)}{\partial a^2}, t_{ae}(a,e) = \frac{\partial^2 t(a,e)}{\partial a\partial e}, t_{ee}(a,e) = \frac{\partial^2 t(a,e)}{\partial e^2}$. Since $t(a,\not{e}(a))=0$, taking the derivative of $t(a,\not{e}(a))$ with respect to a gives

$$t_a(a,\not{e}(a)) + t_e(a,\not{e}(a))\frac{\partial \not{e}(a)}{\partial a} = 0.$$

Therefore,

$$\frac{\partial \not{e}(a)}{\partial a} = -\frac{t_a(a,\not{e}(a))}{t_e(a,\not{e}(a))} > 0.$$

Further,

$$\begin{aligned}
\frac{\partial^2 \not{e}(a)}{\partial a^2} &= -\frac{\frac{\partial}{\partial a}\left(t_a(a,\not{e}(a))\right)t_e(a,\not{e}(a)) - \frac{\partial}{\partial a}\left(t_e(a,\not{e}(a))\right)t_a(a,\not{e}(a))}{\left(t_e(a,\not{e}(a))\right)^2}\\
&= -\frac{t_{aa}(a,\not{e}(a))t_e^2(a,\not{e}(a)) - 2t_a(a,\not{e}(a))t_e(a,\not{e}(a))t_{ae}(a,\not{e}(a)) + t_{ee}(a,\not{e}(a))t_a^2(a,\not{e}(a))}{\left(t_e(a,\not{e}(a))\right)^3}.
\end{aligned}$$

Note that

$$\begin{aligned}
&t_{aa}(a,\not{e}(a))t_e^2(a,\not{e}(a)) - 2t_a(a,\not{e}(a))t_e(a,\not{e}(a))t_{ae}(a,\not{e}(a)) + t_{ee}(a,\not{e}(a))t_a^2(a,\not{e}(a))\\
=&-\beta\lambda_J^2\mathbb{E}\Bigg[\frac{\beta^2(\epsilon+\kappa)\lambda_J^3(\lambda+a\lambda_J)^3 + 4\beta\lambda_J(\lambda+a\lambda_J)\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+\not{e}\lambda_J)\right)^{3/2}\left(4(\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+\not{e}\lambda_J)\right)\theta_C}{32\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+\not{e}\lambda_J)\right)^{9/2}}\\
&\qquad + \frac{4(\lambda+\not{e}\lambda_J)\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+\not{e}\lambda_J)\right)^2\left(4(\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+\not{e}\lambda_J)\right)\theta_C^2}{32\left((\epsilon+\kappa)\lambda_J+(\lambda+a\lambda_J)(\lambda+\not{e}\lambda_J)\right)^{9/2}}\Bigg] < 0.
\end{aligned}$$

Therefore, $\frac{\partial^2 e(a)}{\partial a^2} < 0$.

Thus, the result is proved. $\square$

We next provide the proofs to the technical results in Appendix EC.5.3.

EC.7.1. Proof of Proposition EC.5

Let (a^*, e^*, f^*) be the optimal solution under the CM model as defined in Proposition EC.4 and Π^* be the corresponding optimal objective value. Next, we show that the sponsor can achieve the centralized outcome under the SI model by adopting a two-part tariff contract $T(\hat{e}) = \beta\hat{e} + w$, where $\hat{e} = e + \frac{\epsilon+\kappa}{\lambda+\lambda_J a}$ is the reported effort and e is the true effort. Then, given a, the coordinator's expected benefit is

$$\mathbb{E}_\epsilon\left[T\left(e + \frac{\epsilon+\kappa}{\lambda+\lambda_J a}\right) - \frac{1}{2}\theta_C e^2\right] = \mathbb{E}_\epsilon\left[\beta\left(e + \frac{\epsilon+\kappa}{\lambda+\lambda_J a}\right) + w - \frac{1}{2}\theta_C e^2\right] = \beta e + w + \frac{\beta\kappa}{\lambda+\lambda_J a} - \frac{1}{2}\theta_C e^2,$$

where $\mathbb{E}(\epsilon) = 0$. For a given pair $(\beta,\, w)$, the effort level that maximizes the coordinator's expected benefit is $\frac{\beta}{\theta_C}$.

Consider the following values $\beta = \theta_C e^*, w = -\left[\frac{\kappa\theta_C e^*}{\lambda+\lambda_J a^*} + \frac{1}{2}\theta_C {e^*}^2\right], a = a^*$, and $f = f^*$. It is easy to verify that these values satisfy the IR constraint in the sponsor's problem. Further, from the optimality of (a^*, e^*, f^*) under the CM model, the retention constraint is also satisfied. The sponsor's expected retention cost under the SI model is

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{2}\theta_I {a^*}^2 + T\left(e^* + \frac{\epsilon+\kappa}{\lambda+\lambda_J a^*}\right) + \tilde{\delta}(a^*, e^*, f^*) N f^*\right] \\
=&\mathbb{E}\left[\frac{1}{2}\theta_I {a^*}^2 + \frac{1}{2}\theta_C {e^*}^2 + \tilde{\delta}(a^*, e^*, f^*) N f^*\right] = \Pi^*.
\end{aligned}$$

Thus, by providing the contract $T(\hat{e})$, the sponsor can achieve the optimal retention cost Π^* under the SI model. $\square$

The proofs of Lemma EC.4, Lemma EC.5, and Lemma EC.6 follow a similar structure as the proofs of the corresponding results in Section 4, and are omitted for brevity.

EC.7.2. Proof of Proposition EC.6

In Appendix EC.5.1 in the manuscript, we have established that when ϵ is uncertain, the sponsor's problem can be solved by replacing $\bar{\delta}$ with $\bar{\delta} - \Phi^{-1}(1-\zeta)$ in the retention constraint. Therefore, the above result is a direct extension of Proposition 8 in the manuscript. $\square$

EC.7.3. Proof of Proposition EC.7

Consider the following contracts as described in Appendix EC.5.3.2:

$$\begin{aligned}
T_I(\hat{a}, \hat{e}; f, \hat{\delta}) &= \nu\hat{a} + w_I - \nu\hat{a}\mathbb{1}\left\{((\lambda\hat{a} + \lambda\hat{e} + \lambda_J\hat{a}\hat{e} + \lambda_f f) - \kappa > \hat{\delta}) \wedge (\hat{e} \le \hat{a})\right\}, \\
T_C(\hat{a}, \hat{e}; f, \hat{\delta}) &= \beta\hat{e} + w_C - \beta\hat{e}\mathbb{1}\left\{((\lambda\hat{a} + \lambda\hat{e} + \lambda_J\hat{a}\hat{e} + \lambda_f f) - \kappa > \hat{\delta}) \wedge (\hat{a} \le \hat{e})\right\}.
\end{aligned}$$

We next prove that by optimally designing the value of β, ν, w_I, w_C, the sponsor can achieve the centralized outcome. Let (a^*, e^*, f^*) be the optimal solution under the CM model. Further, let Π^* be the sponsor's optimal expected retention cost under the CM model. Consider the solution where the sponsor chooses monetary payment f^*, and

$$\beta = \theta_C e^* \Bigg/ \mathbb{E}\left[\frac{\lambda + \lambda_J a^*}{2\sqrt{\lambda^2 + \lambda_J(\lambda(a^* + e^*) + \lambda_J a^* e^* + \kappa + \epsilon)}}\right], \tag{EC.52}$$

$$\nu = \theta_I a^* \Bigg/ \mathbb{E}\left[\frac{\lambda + \lambda_J e^*}{2\sqrt{\lambda^2 + \lambda_J(\lambda(a^* + e^*) + \lambda_J a^* e^* + \kappa + \epsilon)}}\right], \tag{EC.53}$$

$$w_I = \frac{1}{2}\theta_I a^{*2} - \nu \mathbb{E}\left[\frac{\sqrt{\lambda^2 + \lambda_J(\lambda(a^* + e^*) + \lambda_J a^* e^* + \kappa + \epsilon)} - \lambda}{\lambda_J}\right], \tag{EC.54}$$

$$w_C = \frac{1}{2}\theta_C e^{*2} - \nu \mathbb{E}\left[\frac{\sqrt{\lambda^2 + \lambda_J(\lambda(a^* + e^*) + \lambda_J a^* e^* + \kappa + \epsilon)} - \lambda}{\lambda_J}\right]. \tag{EC.55}$$

We first analyze the providers' game on reporting effort, then the providers' game on choosing the true effort. We have the following results.

CLAIM EC.4. *There exists a unique equilibrium* $(\hat{a}^\diamond, \hat{e}^\diamond)$ *to the providers' effort reporting game. Further,* $\hat{a}^\diamond = \hat{e}^\diamond$*.*

CLAIM EC.5. *There exists a unique subgame perfect equilibrium* (a^*, e^*) *to the providers' effort choice. Further, the sponsor's expected retention cost is* Π^**.*

Proof of Claim EC.4:

We now consider the providers' decision on the reported effort. Let $\hat{\mathscr{e}}(\hat{a})$ (resp., $\hat{\mathscr{a}}(\hat{e})$) denote the coordinator's (resp., investigator's) best response to the investigator's (resp., coordinator's) reported effort $\hat{a}$ (resp., $\hat{e}$). For the coordinator, the best response maximizes the expected benefit:

$$\max_{\hat{e} \in [0,1]} \quad T_C(\hat{a}, \hat{e}; f, \hat{\delta}) - \frac{1}{2}\theta_C e^2, \tag{EC.56}$$

Observe from the contract that the payoff increases in the reported effort unless $\lambda(\hat{a} + \hat{e}) + \lambda_J \hat{a}\hat{e} + \lambda_f f - \kappa > \hat{\delta}$ and $\hat{e} \geq \hat{a}$. Thus, $\hat{\mathscr{e}}(\hat{a}) = \max\{\hat{a} - \xi, \frac{\hat{\delta} + \kappa - \lambda\hat{a} - \lambda_f f}{\lambda + \lambda_J \hat{a}}\}$, where ξ is an infinitesimal positive number. Similarly, we can argue that the best response function for the investigator $\hat{\mathscr{a}}(\hat{e}) = \max\{\hat{e} - \xi, \frac{\hat{\delta} + \kappa - \lambda\hat{e} - \lambda_f f}{\lambda + \lambda_J \hat{e}}\}$. Thus, we have the following possibilities:

1. $\hat{\mathscr{e}}(\hat{a}) = \hat{a} - \xi, \hat{\mathscr{a}}(\hat{e}) = \hat{e} - \xi$. For any equilibrium, it must satisfy $\hat{e} = \hat{\mathscr{e}}(\hat{a}) = \hat{\mathscr{e}}(\hat{\mathscr{a}}(\hat{e})) = \hat{e} - 2\xi$, which forms a contradiction. Therefore, there is no feasible solution.
2. $\hat{\mathscr{e}}(\hat{a}) = \hat{a} - \xi, \hat{\mathscr{a}}(\hat{e}) = \frac{\hat{\delta} + \kappa - \lambda\hat{e} - \lambda_f f}{\lambda + \lambda_J \hat{e}}$. Under this case, we have $\hat{a} - \xi > \frac{\hat{\delta} + \kappa - \lambda\hat{a} - \lambda_f f}{\lambda + \lambda_J \hat{a}}$, or, $\hat{\delta} + \kappa - \lambda\hat{a} - \lambda_f f < (\hat{a} - \xi)(\lambda + \lambda_J \hat{a})$. Therefore, for any equilibrium, we have $\hat{a} = \hat{\mathscr{a}}(\hat{\mathscr{e}}(\hat{a})) = \frac{\hat{\delta} + \kappa - \lambda(\hat{a} - \xi) - \lambda_f f}{\lambda + \lambda_J(\hat{a} - \xi)} = \frac{\hat{\delta} + \kappa - \lambda\hat{a} - \lambda_f f}{\lambda + \lambda_J(\hat{a} - \xi)} + \frac{\lambda\xi}{\lambda + \lambda_J(\hat{a} - \xi)} < \frac{(\hat{a} - \xi)(\lambda + \lambda_J \hat{a})}{\lambda + \lambda_J(\hat{a} - \xi)} + \frac{\lambda\xi}{\lambda + \lambda_J(\hat{a} - \xi)} = \frac{(\hat{a})(\lambda + \lambda_J \hat{a}) - \xi\lambda_J \hat{a}}{\lambda + \lambda_J(\hat{a} - \xi)} = \hat{a}$, which forms a contradiction. Therefore, there is no feasible solution.

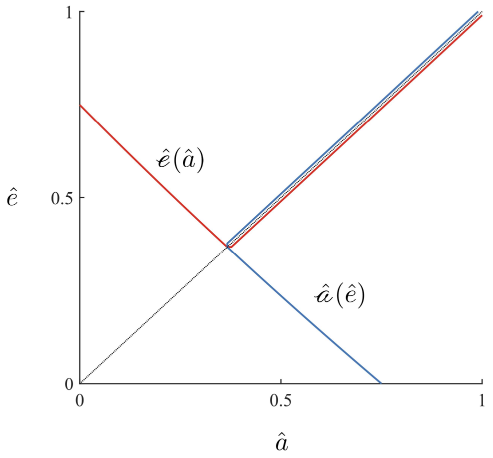


Figure EC.8 Providers' Best Response Functions

3. $\hat{e}(\hat{a}) = \frac{\hat{\delta}+\kappa-\lambda\hat{a}-\lambda_f f}{\lambda+\lambda_J\hat{a}}$, $\hat{a}(\hat{e}) = \hat{e}-\xi$. In this case, we can similarly argue that there is no equilibrium solution.
4. $\hat{e}(\hat{a}) = \frac{\hat{\delta}+\kappa-\lambda\hat{a}-\lambda_f f}{\lambda+\lambda_J\hat{a}}$, $\hat{a}(\hat{e}) = \frac{\hat{\delta}+\kappa-\lambda\hat{e}-\lambda_f f}{\lambda+\lambda_J\hat{e}}$. The reported efforts must be such that $\hat{a}-\xi \le \frac{\hat{\delta}+\kappa-\lambda\hat{a}-\lambda_f f}{\lambda+\lambda_J\hat{a}}$ and $\hat{e}-\xi \le \frac{\hat{\delta}+\kappa-\lambda\hat{e}-\lambda_f f}{\lambda+\lambda_J\hat{e}}$. Solving the above inequalities, any solution $(\hat{a},\hat{e})$ of the equation $\lambda(\hat{a}+\hat{e})+\lambda_J\hat{a}\hat{e}+\lambda_f f-\kappa = \hat{\delta}$ will be an equilibrium as long as $0 \le \hat{a} < \frac{\sqrt{\lambda^2+\lambda_J(\hat{\delta}+\kappa-\lambda_f f+\lambda_J\frac{\xi^2}{4})}-\lambda}{\lambda_J}+\frac{\xi}{2}$ and $0 \le \hat{e} < \frac{\sqrt{\lambda^2+\lambda_J(\hat{\delta}+\kappa-\lambda_f f+\lambda_J\frac{\xi^2}{4})}-\lambda}{\lambda_J}+\frac{\xi}{2}$ (See Figure EC.8). It can be verified that when $\xi \to 0$, there is a unique solution where $\hat{a} = \hat{e} = \frac{\sqrt{\lambda^2+\lambda_J(\hat{\delta}+\kappa-\lambda_f f)}-\lambda}{\lambda_J}$, that satisfies the conditions.

Therefore, there is a unique equilibrium solution

$$\left(\frac{\sqrt{\lambda^2+\lambda_J(\hat{\delta}+\kappa-\lambda_f f)}-\lambda}{\lambda_J}, \frac{\sqrt{\lambda^2+\lambda_J(\hat{\delta}+\kappa-\lambda_f f)}-\lambda}{\lambda_J}\right). \tag{EC.57}$$

□

Proof of Claim EC.5:

Now, we analyze the providers' decision on the true effort level. Let $\tilde{\alpha}(a,e)$ denote the reported effort value as defined in (EC.57). Substituting $\hat{\delta} = \lambda(a+e)+\lambda_J ae+\lambda_f f+\epsilon$ in the expression for $\tilde{\alpha}(a,e)$ above, we get

$$\tilde{\alpha}(a,e) = \frac{\sqrt{\lambda^2+\lambda_J(\lambda(a+e)+\lambda_J ae+\kappa+\epsilon)}-\lambda}{\lambda_J}.$$

It is easy to verify that

$$\frac{\partial^2[\tilde{\alpha}(a,e)|\epsilon=\hat{\epsilon}]}{\partial a^2} < 0, \quad \frac{\partial^2[\tilde{\alpha}(a,e)|\epsilon=\hat{\epsilon}]}{\partial e^2} < 0, \qquad \forall\hat{\epsilon}\in[-\kappa,\kappa].$$

Now consider the providers' decision on the true effort level. Given the investigator's effort a, the coordinator's best response is to maximize the expected benefit:

$$\max_{e} \quad \beta \mathbb{E}_{\epsilon}\left[\tilde{\alpha}(a,e)\right] + w_C - \frac{1}{2}\theta_C e^2.$$

Note that the objective function is concave, there is a unique solution to the above problem, which we denote by $\mathscr{e}(a)$. From Lemma EC.8, $\mathscr{e}(a)$ is increasing concave in a, and $\mathscr{e}(a) > 0$. Similarly, given the coordinator's effort e, the investigator's best response is to maximize the expected benefit:

$$\max_{a} \quad \nu \mathbb{E}_{\epsilon}\left[\tilde{\alpha}(a,e)\right] + w_I - \frac{1}{2}\theta_I a^2.$$

Since the objective function is concave, there is a unique solution $\mathscr{a}(e)$ to the above problem. Following a similar argument as in Lemma EC.8, $\mathscr{a}(e)$ is increasing concave in e and $\mathscr{a}(e) > 0$. Therefore, the providers' best response functions $\mathscr{a}(e)$ and $\mathscr{e}(a)$ cross at most once, thus, there is at most one equilibrium in the providers' decision game. Further, using the values of β and ν, it's easy to verify that at the centralized optimal solution (a^*, e^*), we have the following:

$$\frac{\partial}{\partial e}\left[\beta \mathbb{E}_{\epsilon}[\tilde{\alpha}(a^*,e)] + w_C - \frac{1}{2}\theta_C e^2\right]\Bigg|_{e=e^*} = 0, \quad \frac{\partial}{\partial a}\left[\nu \mathbb{E}_{\epsilon}[\tilde{\alpha}(a,e^*)] + w_I - \frac{1}{2}\theta_I a^2\right]\Bigg|_{a=a^*} = 0,$$

and hence, (a^*, e^*) is an equilibrium in the providers' decision game. Therefore, (a^*, e^*) is the unique equilibrium in the providers' decision game. Since $w_I = \frac{1}{2}\theta_I a^{*2} - \nu \mathbb{E}_{\epsilon}\left[\tilde{\alpha}(a^*,e^*)\right]$ and $w_C = \frac{1}{2}\theta_C e^{*2} - \beta \mathbb{E}_{\epsilon}\left[\tilde{\alpha}(a^*,e^*)\right]$, the providers' receive zero surplus in expectation at the equilibrium, and thus, the IR constraint is satisfied. Further, from the optimality of (a^*, e^*, f^*) under the CM model, the retention constraint is also satisfied at the equilibrium. Therefore, the above equilibrium is valid for the sponsor's problem, and the sponsor's corresponding expected retention cost is, therefore,

$$\frac{1}{2}\theta_I a^{*2} + \frac{1}{2}\theta_C e^{*2} + \tilde{\delta}(a^*, e^*, f^*) N f^*,$$

which is the same as the CM model. □

The proof of Proposition EC.7 is now complete following Claims EC.4 and EC.5. □

EC.7.4. Proof of Lemma EC.7

For any given $\hat{a}$, the compensation to the coordinator is increasing in $\hat{e}$ when $(\lambda\hat{a} + \lambda\hat{e} + \lambda_J\hat{a}\hat{e} + \lambda_f f) - \kappa \le \hat{\delta}$ and 0 afterward. Therefore, the coordinator's best response is to report $\hat{e}$ such that $(\lambda\hat{a} + \lambda\hat{e} + \lambda_J\hat{a}\hat{e} + \lambda_f f) - \kappa = \hat{\delta}$. A similar argument applies to the investigator. Therefore, any element in the set $\{(\hat{a},\hat{e}) \in [0,1]^2 : (\lambda\hat{a} + \lambda\hat{e} + \lambda_J\hat{a}\hat{e} + \lambda_f f) - \kappa = \hat{\delta}\}$ is a Nash equilibrium. □

References

Arjun, M. C., Ahamed, A., Kurpad, A. V., and Thomas, T. (2024). Blockchain based clinical trial management system: A scoping review and conceptual framework. *medRxiv*.

Block, H. W., Savits, T. H., and Singh, H. (1998). The reversed hazard rate function. *Probability in the Engineering and Informational Sciences*, 12(1):69–90.

Bolton, P. (2005). *Contract theory*. MIT Press, Cambridge, MA. Literaturverzeichnis: Seite 687-707.

Dey, D., Fan, M., and Zhang, C. (2010). Design and analysis of contracts for software outsourcing. *Information Systems Research*, 21(1):93–114.

Financial disclosure by clinical investigators (1998). 21 C.F.R. §54.

General responsibilities of investigators (2004). 21 C.F.R. §312.60.

General responsibilities of sponsors (2004). 21 C.F.R. §312.50.

Good, M. J., Hurley, P., Woo, K. M., Szczepanek, C., Stewart, T., Robert, N., Lyss, A., Gönen, M., and Lilenbaum, R. (2016). Assessing clinical trial–associated workload in community-based research programs using the asco clinical trial workload assessment tool. *Journal of Oncology Practice*, 12(5):e536–e547.

ICH (2023). Good clinical practice (gcp) e6(r3). Available at `https://database.ich.org/sites/default/files/ICH_E6%28R3%29_DraftGuideline_2023_0519.pdf`.

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

Investigator recordkeeping and record retention (2004). 21 C.F.R. §312.62.

IQVIA (2021). Clinical trial management system. Available at `https://www.iqvia.com/-/media/iqvia/pdfs/library/fact-sheets/clinical-trial-management-system.pdf`.

James, P., Bebee, P., Beekman, L., Browning, D., Innes, M., Kain, J., Royce-Westcott, T., and Waldinger, M. (2011). Creating an effort tracking tool to improve therapeutic cancer clinical trials workload management and budgeting. *Journal of the National Comprehensive Cancer Network*, 9(11):1228–1233.

Johnson & Johnson (2020). Position on the conduct of clinical trials. Available at `https://www.jnj.com/about-jnj/policies-and-positions/our-position-on-conducting-clinical-trials`.

Laffont, J.-J. and Martimort, D. (2002). *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Love, S. B., Yorke-Edwards, V., Ward, E., Haydock, R., Keen, K., Biggs, K., Gopalakrishnan, G., Marsh, L., O'Sullivan, L., Fox, L., Payerne, E., Hood, K., and Meakin, G. (2022). What is the purpose of clinical trial monitoring? *Trials*, 23(1).

Merck (2019). Clinical trial ethics. Available at `https://www.merck.com/wp-content/uploads/sites/124/2020/04/Policy_2019_Clinical-Trial-Ethics_MERCK.pdf`.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175.

Pfeiffer, J. and Wells, C. (2017). *A Practical Guide to Managing Clinical Trials*. CRC Press.

Pfizer (2021). Compensation to investigators in clinical studies. Available at `https://www.pfizer.com/science/clinical-trials/integrity-and-transparency/compensation-investigators`.

Rose, S. (2021). Investigator compensation: No one size fits all. *Clinical Researcher*, 703:7.

Siemsen, E., Balasubramanian, S., and Roth, A. V. (2007). Incentives that induce task-related effort, helping, and knowledge sharing in workgroups. *Management Science*, 53(10):1533–1550.

Simester, D. and Zhang, J. (2010). Why are bad products so hard to kill? *Management Science*, 56(7):1161–1179.

Subbiah, V. (2023). The next generation of evidence-based medicine. *Nature Medicine*, 29(1):49–58.

Veeva (2016). Veeva vault ctms. Available at `https://www.veeva.com/products/vault-ctms/`.

Zhang, H., Kong, G., and Rajagopalan, S. (2018). Contract design by service providers with private effort. *Management Science*, 64(6):2672–2689.