Online Appendix to
"Joint vs. Separate Crowdsourcing Contests"

A. Proofs.

Proof of Lemma 1. For the sub-contest l, $l=1,2$, the expected payoff to contestant i is $\mathsf{E}(u_i(e_i^l)) = A^l \int_{-\infty}^{+\infty} \Psi(e_i^l - e^{l*}(n) + \xi^l)^{n-1}\psi(\xi^l)d\xi^l - C^l(e_i^l)$. The FOC yields $A^l \int_{-\infty}^{+\infty} (n-1)\Psi(e_i^l - e^{l*}(n) + \xi^l)^{n-2}\psi(e_i^l - e^{l*}(n) + \xi^l)\psi(\xi^l)d\xi^l = C^{l\prime}(e_i^l)$. In the symmetric equilibrium, contestant i makes the same effort $e^{l*}(n)$ as other contestants. Since $C^{l\prime\prime}(\cdot) > 0$, the equilibrium effort is given by $e^{l*}(n) = {C^{l\prime}}^{-1}(A^l \int_{-\infty}^{+\infty} (n-1)\Psi(\xi^l)^{n-2}\psi(\xi^l)^2 d\xi^l) = {C^{l\prime}}^{-1}(A^l h(\xi^l;n))$. □

Proof of Lemma 2. Consider the optimization problem below:

$$\min_{e^1,e^2} \; C^1(e^1) + C^2(e^2) \quad \text{s.t. } e^1 + e^2 = e^\circ.$$

The solution to this problem can be typically found by writing the Lagrangean, $L(e^1, e^2, e^\circ; \lambda) = C^1(e^1) + C^2(e^2) + \lambda(e^\circ - e^1 - e^2)$, and the FOCs are

$$\frac{\partial L}{\partial e^1} = C^{1\prime}(\tilde{e}^1) - \tilde{\lambda} = 0 \quad (3a), \quad \frac{\partial L}{\partial e^2} = C^{2\prime}(\tilde{e}^2) - \tilde{\lambda} = 0 \quad (3b), \quad \frac{\partial L}{\partial \lambda} = e^\circ - \tilde{e}^1 - \tilde{e}^2 = 0 \quad (3c). \tag{A.1}$$

Solving the FOCs yields the Lagrange multiplier $\tilde{\lambda} = \lambda(e^\circ)$ and the optimal efforts $\tilde{e}^1(e^\circ)$, $\tilde{e}^2(e^\circ)$ along the two dimensions. Now plugging $\tilde{e}^1(e^\circ)$, $\tilde{e}^2(e^\circ)$ into the objective function and we can get a new function $C^\circ(e^\circ) = C^1(\tilde{e}^1(e^\circ)) + C^2(\tilde{e}^2(e^\circ))$ which yields the minimum value of C° for a given e°. Taking the derivative of C° with respect to e°, we obtain

$$\frac{dC^\circ(e^\circ)}{de^\circ} = C^{1\prime}(\tilde{e}^1(e^\circ))\frac{d\tilde{e}^1(e^\circ)}{de^\circ} + C^{2\prime}(\tilde{e}^2(e^\circ))\frac{d\tilde{e}^2(e^\circ)}{de^\circ}. \tag{A.2}$$

By (A.1a) and (A.1b), $C^{1\prime}(\tilde{e}^1) = \tilde{\lambda}$ and $C^{2\prime}(\tilde{e}^2) = \tilde{\lambda}$, we have (A.2)$=\tilde{\lambda}\left[\frac{d\tilde{e}^1(e^\circ)}{de^\circ} + \frac{d\tilde{e}^2(e^\circ)}{de^\circ}\right]$. By (A.1c), $e^\circ = \tilde{e}^1 + \tilde{e}^2$, and hence $\frac{d\tilde{e}^1(e^\circ)}{de^\circ} + \frac{d\tilde{e}^2(e^\circ)}{de^\circ} = 1$, thus $\frac{dC^\circ(e^\circ)}{de^\circ} = \tilde{\lambda}$. Because $C^1(\cdot)$ and $C^2(\cdot)$ are strictly increasing, again by (A.1a) and (A.1b), $\tilde{\lambda} > 0$, thus, $\frac{dC^\circ(e^\circ)}{de^\circ} > 0$, i.e., $C^\circ(e^\circ)$ is strictly increasing.

By $\frac{dC^\circ(e^\circ)}{de^\circ} = \tilde{\lambda}$, (A.1a) and (A.1b), we have

$$C^{\circ\prime}(e^\circ) = C^{1\prime}(\tilde{e}^1) = C^{2\prime}(\tilde{e}^2). \tag{A.3}$$

By the assumption $C^{2\prime\prime}(\cdot) > 0$, ${C^{2\prime}}^{-1}(\cdot)$ is well-defined, hence ${C^{2\prime}}^{-1}\left(C^{1\prime}(\tilde{e}^1)\right) = \tilde{e}^2$. By (A.1c), $\tilde{e}^1 + \tilde{e}^2 = e^\circ$, we obtain ${C^{2\prime}}^{-1}\left(C^{1\prime}(\tilde{e}^1)\right) + \tilde{e}^1 = e^\circ$. Because $C^{1\prime\prime}(\cdot) > 0$ and $C^{2\prime\prime}(\cdot) > 0$, ${C^{1\prime}}^{-1}(\cdot)$ and ${C^{2\prime}}^{-1}(\cdot)$ are strictly increasing, thus $\tilde{e}^1$ is strictly increasing in e° by (A.3). Further by (A.1a), $C^{1\prime}(\tilde{e}^1) = \tilde{\lambda}$, $\tilde{\lambda}$ is strictly increasing in e°. Then because $\frac{dC^\circ(e^\circ)}{de^\circ} = \tilde{\lambda}$ that we have just proved, $\frac{dC^\circ(e^\circ)}{de^\circ}$ is strictly increasing in e°, i.e., $C^\circ(e^\circ)$ is strictly convex. □

Proof of Lemma 3. Consider the joint contest with n contestants. Given a fixed aggregate effort level e°, all the contestants follow the optimal effort allocation, $e^\circ = \tilde{e}^1 + \tilde{e}^2$. The strategy of a contestant is his aggregate effort level, e°. By Lemma 2, $C^\circ(e^\circ)$ is strictly increasing and strictly convex. Then the expected payoff to contestant i is $\mathsf{E}(u_i(e_i^\circ)) = A\int_{-\infty}^{+\infty}\Psi^\circ(e_i^\circ - e^{\circ *}(n) + \xi^\circ)^{n-1}\psi^\circ(\xi^\circ)d\xi^\circ - C^\circ(e_i^\circ)$. Similar to the derivation in Lemma 1, the equilibrium effort in the joint contest is $e^{\circ *}(n) = {C^{\circ}}'^{-1}(A\int_{-\infty}^{+\infty}(n-1)\Psi^\circ(\xi^\circ)^{n-2}\psi^\circ(\xi^\circ)^2 d\xi^\circ) = {C^{\circ}}'^{-1}(Ah^\circ(\xi^\circ;n))$. $\square$

Proof of Lemma 4. Denote $\xi^\circ = \xi^1 + \xi^2$ and it has CDF $\Psi^\circ(\xi^\circ)$. Because the random factors along the two attributes are identical, denote the quantile function of ξ^l, $l = 1,2$, by $\Psi^{-1}(u)$ and the quantile function of ξ° by $\Psi^{\circ -1}(u)$. Write the formula of $\mathsf{E}(\xi_{(n)}^l)$, $l = 1,2$,

$$\mathsf{E}(\xi_{(n)}^l) = \int_{-\infty}^{+\infty}\xi^l n\Psi(\xi^l)^{n-1}\psi(\xi^l)d\xi^l = \int_{-\infty}^{+\infty}\xi^l n\Psi(\xi^l)^{n-1}d\Psi(\xi^l) = \int_0^1 \Psi^{-1}(u)nu^{n-1}du, \quad \text{(A.4)}$$

where the last equality is by substituting $\Psi^{-1}(u) = \xi^l$. Similarly, $\mathsf{E}(\xi_{(n)}^\circ) = \int_0^1 \Psi^{\circ -1}(u)nu^{n-1}du$. Then,

$$\begin{aligned}&\mathsf{E}(\xi_{(n)}^1) + \mathsf{E}(\xi_{(n)}^2) - \mathsf{E}((\xi^1+\xi^2)_{(n)}) = \mathsf{E}(\xi_{(n)}^1) + \mathsf{E}(\xi_{(n)}^2) - \mathsf{E}(\xi_{(n)}^\circ)\\ &= \int_0^1 2\Psi^{-1}(u)nu^{n-1}du - \int_0^1 \Psi^{\circ -1}(u)nu^{n-1}du = \int_0^1 (2\Psi^{-1}(u) - \Psi^{\circ -1}(u))nu^{n-1}du. \quad \text{(A.5)}\end{aligned}$$

Recall the assumption that ξ^l, $l = 1,2$, are identical, and they satisfy the regularity condition that there exists $u_0 \in (0,1)$ such that $2\Psi^{-1}(u) - \Psi^{\circ -1}(u) < 0$ if $u \in (0,u_0)$, and $2\Psi^{-1}(u) - \Psi^{\circ -1}(u) > 0$ if $u \in (u_0,1)$. Thus, by (A.5), we have

$$\begin{aligned}&\mathsf{E}(\xi_{(n)}^1) + \mathsf{E}(\xi_{(n)}^2) - \mathsf{E}((\xi^1+\xi^2)_{(n)})\\ &= \int_0^{u_0}(2\Psi^{-1}(u) - \Psi^{\circ -1}(u))nu^{n-1}du + \int_{u_0}^1(2\Psi^{-1}(u) - \Psi^{\circ -1}(u))nu^{n-1}du\\ &> \int_0^{u_0}(2\Psi^{-1}(u) - \Psi^{\circ -1}(u))nu_0^{n-1}du + \int_{u_0}^1(2\Psi^{-1}(u) - \Psi^{\circ -1}(u))nu_0^{n-1}du\\ &= \int_0^1(2\Psi^{-1}(u) - \Psi^{\circ -1}(u))nu_0^{n-1}du\\ &= nu_0^{n-1}\int_0^1(2\Psi^{-1}(u) - \Psi^{\circ -1}(u))du = nu_0^{n-1}[\mathsf{E}(\xi^1) + \mathsf{E}(\xi^2) - \mathsf{E}(\xi^1+\xi^2)] = 0\end{aligned}$$

The inequality is due to that u^{n-1} is strictly increasing in $(0,1)$, and $(2\Psi^{-1}(u) - \Psi^{\circ -1}(u))nu^{n-1}$ has a measure of 0 for $u = 0,1$. As a result, $\mathsf{E}(\xi_{(n)}^1) + \mathsf{E}(\xi_{(n)}^2) - \mathsf{E}((\xi^1+\xi^2)_{(n)}) > 0$. $\square$

Proof of Proposition 1. Part (i) is directly given by Lemma 4. For part (ii), by Lemmas 1 and 3, the difference of equilibrium efforts with n contestants is

$$\Delta_n^e = e^{1*} + e^{2*} - e^{\circ *} = {C^1}'^{-1}\left(A^1 h(\xi;n)\right) + {C^2}'^{-1}\left(A^2 h(\xi;n)\right) - {C^\circ}'^{-1}\left(Ah^\circ(\xi^\circ;n)\right)$$

$$= C^{1'^{-1}}\left(A^1h(\xi;n)\right) + C^{2'^{-1}}\left(A^2h(\xi;n)\right) - C^{1'^{-1}}\left(Ah^\circ(\xi^\circ;n)\right) - C^{2'^{-1}}\left(Ah^\circ(\xi^\circ;n)\right), \quad \text{(A.6)}$$

where the third equality is driven by Lemma 2 that $C^{\circ\prime}(e^\circ) = C^{1'}(\tilde{e}^1) = C^{2'}(\tilde{e}^2)$ and all the cost functions are strictly increasing and strictly convex. The sufficient condition for $\Delta_n^e < 0$ is $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n),(1-w)h(\xi;n)\}$.

Moreover, (a) is directly given by letting $w = 1/2$. For (b), denote $C(\cdot) = C^1(\cdot) = C^2(\cdot)$. If $C'(\cdot)$ is convex, then $C'^{-1}(\cdot)$ is concave. By the concavity of $C'^{-1}(\cdot)$, we have $\frac{1}{2}(C'^{-1}(y_1) + C'^{-1}(y_2)) \leq C'^{-1}\left(\frac{y_1+y_2}{2}\right)$, where y_1 and y_2 are in the domain of $C'^{-1}(\cdot)$. As a result, the inequality holds for $C^{1'^{-1}}(wAh(\xi;n)) + C^{2'^{-1}}((1-w)Ah(\xi;n)) = C'^{-1}(wAh(\xi;n)) + C'^{-1}((1-w)Ah(\xi;n)) \leq 2C'^{-1}(Ah(\xi;n)/2)$. Furthermore, by (A.6) and $C(\cdot) = C^1(\cdot) = C^2(\cdot)$, $\Delta_n^e = C'^{-1}(wAh(\xi;n)) + C'^{-1}((1-w)Ah(\xi;n)) - 2C'^{-1}(Ah^\circ(\xi^\circ;n)) \leq 2C'^{-1}(Ah(\xi;n)/2) - 2C'^{-1}(Ah^\circ(\xi^\circ;n))$. Since $C(\cdot)$ is strictly convex, $C'(\cdot)$ is strictly increasing, and then $C'^{-1}(\cdot)$ is strictly increasing. Hence, if $h^\circ(\xi^\circ;n) > h(\xi;n)/2$, $\Delta_n^e < 0$. □

Proof of Corollary 1. Denote $H(n) = \frac{h^\circ(\xi^\circ;n)}{h(\xi;n)}$. For any n, if $H(n) > \max\{w,(1-w)\}$, then $\Delta_n^e < 0$. If $\xi \sim N(0,\sigma)$, $\xi^\circ = \xi + \xi \sim N(0,\sqrt{2}\sigma)$. We have

$$h(\xi;n) = \int_{-\infty}^{+\infty}(n-1)\Psi(\xi)^{n-2}\psi(\xi)^2 d\xi = \int_{-\infty}^{+\infty}\psi(\xi)d\Psi(\xi)^{n-1} = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty}\exp(-\xi^2/(2\sigma^2))d\Psi(\xi)^{n-1}.$$

By substituting ξ/σ with y, $h(\xi;n) = \frac{1}{\sqrt{2\pi}\sigma}\int_0^1 \exp(-y^2/2)d\varphi(y)^{n-1}$, where $\varphi(y) \sim N(0,1)$. Similarly, for $h^\circ(\xi^\circ;n)$, by substituting $\xi/(\sqrt{2}\sigma)$ with $\tilde{y}$,

$$\begin{aligned} h^\circ(\xi^\circ;n) &= \int_{-\infty}^{+\infty}\psi^\circ(\xi^\circ)d\Psi^\circ(\xi^\circ)^{n-1} = \frac{1}{2\sqrt{\pi}\sigma}\int_{-\infty}^{+\infty}\exp(-\xi^{\circ 2}/(4\sigma^2))d\Psi^\circ(\xi^\circ)^{n-1} \\ &= \frac{1}{2\sqrt{\pi}\sigma}\int_{-\infty}^{+\infty}\exp(-\tilde{y}^2/2)d\varphi(\tilde{y})^{n-1}. \end{aligned}$$

Then $H(n) = \frac{h^\circ(\xi^\circ;n)}{h(\xi;n)} = 1/\sqrt{2}$. Thus if random factors follow the normal distribution, $H(n) = 1/\sqrt{2}$. By $H(n) > \max\{w,(1-w)\}$, if $w \in (1-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2})$, $\Delta_n^e < 0$ for any n. □

Corollary A.1 (Individual Performance). *For any contestant i, the performance in the joint contest first-order stochastically dominates the performance in the separate contest, i.e., $V_i^J \geq_{\text{st}} V_i^S$, if $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n),(1-w)h(\xi;n)\}$.*

Proof of Corollary A.1. For contestant i, $V_i^S = e^{1*}(n) + \xi^1 + e^{2*}(n) + \xi^2$ and $V_i^J = e^{\circ*}(n) + \xi^1 + \xi^2$. By (A.6), if $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n),(1-w)h(\xi;n)\}$, $e^{1*}(n) + e^{2*}(n) \leq e^{\circ*}(n)$. Thus, $\mathsf{P}\{V_i^J \geq z\} = \mathsf{P}\{e^{\circ*}(n) + \xi^1 + \xi^2 \geq z\} = \mathsf{P}\{\xi^1 + \xi^2 \geq z - e^{\circ*}(n)\} \geq \mathsf{P}\{\xi^1 + \xi^2 \geq z - e^{1*}(n) - e^{2*}(n)\} = \mathsf{P}\{V_i^S \geq z\}$ for any z, where the inequality is due to $e^{1*}(n) + e^{2*}(n) \leq e^{\circ*}(n)$. By the definition of usual stochastic order (see Shaked and Shanthikumar 2007, (1.A.1)), $V_i^J \geq_{\text{st}} V_i^S$. □

Many studies in the literature examine the *expected average performance*, e.g., Kalra and Shi (2001), Moldovanu and Sela (2001) and Terwiesch and Xu (2008). Since contestants are ex ante identical, the expected average performance is equivalent to the expected individual performance in our context. In some projects, e.g., a sales contest, every individual contestant's performance counts. The following corollary compares the two contest mechanisms for each contestant's performance.

Corollary A.1 shows that the performance of an individual contestant is more likely to be better in the joint contest than in the separate contest. By $V_i^J \geq_{\text{st}} V_i^S$, we immediately have $\mathsf{E}(V_i^J) \geq \mathsf{E}(V_i^S)$. In contrast to the expected best performance, the combination effect does not play a role in the expected individual performance. For each given individual, the expected performance only relies on the equilibrium effort. With the pooling effect prevailing, the expected individual performance is higher in the joint contest than in the separate contest.

Proof of Proposition 2. If the cost functions are $C^l(\beta e^l)$, $l=1,2$, and $C^\circ(\beta e^\circ)$, by (A.6), the difference of equilibrium efforts between two contests can be written as

$$\Delta_n^e = \frac{1}{\beta}C^{1'^{-1}}\left(\frac{1}{\beta}A^1 h(\xi;n)\right) + \frac{1}{\beta}C^{2'^{-1}}\left(\frac{1}{\beta}A^2 h(\xi;n)\right) - \frac{1}{\beta}C^{1'^{-1}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right) - \frac{1}{\beta}C^{2'^{-1}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right).$$

Since we focus on the case that $e^{l*} > 0$ and $e^{\circ *} > 0$, then $C^{l'^{-1}}\left(\frac{1}{\beta}A^l h(\xi;n)\right) > 0$ and $C^{l'^{-1}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right) > 0$. Denote $\Delta^l = \frac{1}{\beta}C^{l'^{-1}}\left(\frac{1}{\beta}A^l h(\xi;n)\right) - \frac{1}{\beta}C^{l'^{-1}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right)$, $l=1,2$. Note that $\Delta_n^e = \Delta^1 + \Delta^2$. Taking derivative of Δ^l with respect to β, we obtain

$$\begin{aligned}\frac{d\Delta^l}{d\beta} = &-\frac{1}{\beta^2}C^{l'^{-1}}\left(\frac{1}{\beta}A^l h(\xi;n)\right) - C^{l'^{-1'}}\left(\frac{1}{\beta}A^l h(\xi;n)\right)\frac{A^l h(\xi;n)}{\beta^3}\\ &+\frac{1}{\beta^2}C^{l'^{-1}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right) + C^{l'^{-1'}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right)\frac{Ah^\circ(\xi^\circ;n)}{\beta^3}.\end{aligned}$$

By the assumption that $C^{l''}(\cdot) > 0$, $l=1,2$, $C^{l'}(\cdot)$ is strictly increasing and thus $C^{l'^{-1}}(\cdot)$ is strictly increasing. Because $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n),(1-w)h(\xi;n)\}$, $-\frac{1}{\beta^2}C^{l'^{-1}}\left(\frac{1}{\beta}A^l h(\xi;n)\right) + \frac{1}{\beta^2}C^{l'^{-1}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right) > 0$. Now we examine two specific forms of the cost functions.

If $C^l(\beta e^l) = \exp(\rho^l \beta e^l)$, $\rho^l > 0$, then $C^{l'^{-1}}(x) = \frac{1}{\rho^l\beta}\ln(\frac{x}{\rho^l\beta})$ and $C^{l'^{-1'}}(x) = \frac{1}{\rho^l\beta x}$. Thus, $C^{l'^{-1'}}\left(\frac{1}{\beta}A^l h(\xi;n)\right)\frac{A^l h(\xi;n)}{\beta^3} = \frac{1}{\rho\beta^3}$ and $C^{l'^{-1'}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right)\frac{Ah^\circ(\xi^\circ;n)}{\beta^3} = \frac{1}{\rho\beta^3}$. As a result, $\frac{d\Delta^l}{d\beta} > 0$. If $C^l(\beta e^l) = a^l(\beta e^l)^{b^l}$, $a^l > 0$, $b^l \geq 2$, then $C^{l'^{-1'}}\left(\frac{1}{\beta}A^l h(\xi;n)\right)\frac{A^l h(\xi;n)}{\beta^3} = \frac{1}{b^l-1}(\frac{A^l h(\xi;n)}{a^l b^l})^{\frac{1}{b^l-1}}(\frac{1}{\beta})^{\frac{2b^l-1}{b^l-1}}$ and $C^{l'^{-1'}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right)\frac{Ah^\circ(\xi^\circ;n)}{\beta^3} = \frac{1}{b^l-1}(\frac{Ah^\circ(\xi^\circ;n)}{a^l b^l})^{\frac{1}{b^l-1}}(\frac{1}{\beta})^{\frac{2b^l-1}{b^l-1}}$. Since $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n),(1-w)h(\xi;n)\}$, $C^{l'^{-1'}}\left(\frac{1}{\beta}A^l h(\xi;n)\right)\frac{A^l h(\xi;n)}{\beta^3} \leq C^{l'^{-1'}}\left(\frac{1}{\beta}Ah^\circ(\xi^\circ;n)\right)\frac{Ah^\circ(\xi^\circ;n)}{\beta^3}$, and thus $\frac{d\Delta^l}{d\beta} > 0$. To summarize, for the exponential cost functions $C^l(\beta e^l) = \exp(\rho^l\beta e^l)$, $\rho^l > 0$, $l=1,2$, and the polynomial cost functions $C^l(\beta e^l) = a^l(\beta e^l)^{b^l}$, $a^l > 0$, $b^l \geq 2$, we have $\frac{d\Delta^l}{d\beta} > 0$. Therefore, Δ_n^e is strictly increasing in β.

By Proposition 1, if $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n),(1-w)h(\xi;n)\}$, $\Delta_n^e < 0$. Thus, Δ_n^e approaches 0 if β is increasing. Meanwhile, the difference of the expected best random factors Δ_n^ξ is a fixed positive value. As a result, there exists a threshold $\tilde{\beta} > 0$ such that if $\beta \geq \tilde{\beta}$, the difference $\Delta_n = \Delta_n^e + \Delta_n^\xi \geq 0$, and if $\beta < \tilde{\beta}$, the difference $\Delta_n = \Delta_n^e + \Delta_n^\xi < 0$. □

Proposition A.1. *There exists $\bar{n} \geq 2$ such that Δ_n^ξ is increasing in $n \in [2, \bar{n}]$.*

Proof of Proposition A.1. Denote the r-th order statistic of a random variable with a sample size n by subscript $_{(r:n)}$ and denote $\Psi^{-1}(\cdot)$ as the quantile function of CDF $\Psi(\cdot)$. It is sufficient to show that there exists an $\bar{n}$ such that

$$\mathsf{E}(\xi^1_{(n-1:n)}) + \mathsf{E}(\xi^2_{(n-1:n)}) - \mathsf{E}(\xi^\circ_{(n-1:n)}) \leq 0, \tag{A.7}$$

when $n \leq \bar{n}$. According to Chakraborty (1999), (A.7) holds for that the following *regularity condition* is satisfied: there exists $u_0 \in (0,1)$ such that $\Psi^{-1}(u) + \Psi^{-1}(u) - \Psi^{\circ -1}(u) < 0$ if $u \in (0, u_0)$, and $\Psi^{-1}(u) + \Psi^{-1}(u) - \Psi^{\circ -1}(u) > 0$ if $u \in (u_0, 1)$.

Recall the assumption that ξ^l, $l = 1,2$, follows a log-concave distribution. Many log-concave distributions satisfy the regularity condition, e.g., uniform, normal, logistic, Gumbel, Gamma, etc. By Bagnoli and Bergstrom (2005), Corollary 2, if the PDF is log-concave, then its hazard rate function is increasing over the support. By Watson and Gordon (1986, Theorem 1), one sufficient condition for the regularity condition to hold and $u_0 = 0.5$ is that $\Psi(\cdot)$ is a symmetric distribution with a non-decreasing hazard rate function. Therefore, the existence of $\bar{n}$ such that (A.7) holds is guaranteed if the random factors follow a symmetric log-concave distribution or an asymmetric log-concave distribution satisfying the regularity condition.

Write the formulas of $\mathsf{E}(\xi^l_{(n)})$ and $\mathsf{E}(\xi^l_{(n-1)})$, $l = 1,2$, $\mathsf{E}(\xi^l_{(n)}) = \int_{-\infty}^{+\infty} \xi^l n \Psi(\xi^l)^{n-1}\psi(\xi^l) d\xi^l$ and $\mathsf{E}(\xi^l_{(n-1)}) = \int_{-\infty}^{+\infty} \xi^l (n-1)\Psi(\xi^l)^{n-2}\psi(\xi^l) d\xi^l$. The following *recurrence relation* holds, see David and Nagaraja (2003, Chapter 3.4 Relation 1):

$$\begin{aligned} n\mathsf{E}(\xi^l_{(n-1)}) - (n-1)\mathsf{E}(\xi^l_{(n)}) &= \int_{-\infty}^{+\infty} \xi^l n(n-1)\Psi(\xi^l)^{n-2}\psi(\xi^l) d\xi^l - \int_{-\infty}^{+\infty} \xi^l n(n-1)\Psi(\xi^l)^{n-1}\psi(\xi^l) d\xi^l \\ &= \int_{-\infty}^{+\infty} \xi^l n(n-1)(\Psi(\xi^l)^{n-2} - \Psi(\xi^l)^{n-1})\psi(\xi^l) d\xi^l = \mathsf{E}(\xi^l_{(n-1:n)}), \end{aligned} \tag{A.8}$$

where the last equality is because the PDF for the $(n-1)$-th order statistics of ξ^l with the sample size n is $n(n-1)\Psi(\xi^l)^{n-2}(1-\Psi(\xi^l))\psi(\xi^l)$. A similar relation can be applied to ξ°, $n\mathsf{E}(\xi^\circ_{(n-1)}) - (n-1)\mathsf{E}(\xi^\circ_{(n)}) = \mathsf{E}(\xi^\circ_{(n-1:n)})$. By the above relations, we have $\mathsf{E}(\xi^l_{(n-1)}) - \mathsf{E}(\xi^l_{(n)}) = \frac{1}{n}[\mathsf{E}(\xi^l_{(n-1:n)}) - \mathsf{E}(\xi^l_{(n)})]$, $l = 1,2$, and $\mathsf{E}(\xi^\circ_{(n-1)}) - \mathsf{E}(\xi^\circ_{(n)}) = \frac{1}{n}[\mathsf{E}(\xi^\circ_{(n-1:n)}) - \mathsf{E}(\xi^\circ_{(n)})]$. Then,

$$\Delta^\xi_{n-1} - \Delta^\xi_n = [\mathsf{E}(\xi^1_{(n-1)}) + \mathsf{E}(\xi^2_{(n-1)}) - \mathsf{E}(\xi^\circ_{(n-1)})] - [\mathsf{E}(\xi^1_{(n)}) + \mathsf{E}(\xi^2_{(n)}) - \mathsf{E}(\xi^\circ_{(n)})]$$

$$
\begin{aligned}
&= [\mathsf{E}(\xi^1_{(n-1)}) - \mathsf{E}(\xi^1_{(n)})] + [\mathsf{E}(\xi^2_{(n-1)}) - \mathsf{E}(\xi^2_{(n)})] - [\mathsf{E}(\xi^\circ_{(n-1)}) - \mathsf{E}(\xi^\circ_{(n)})] \\
&= \frac{1}{n}\left\{[\mathsf{E}(\xi^1_{(n-1:n)}) - \mathsf{E}(\xi^1_{(n)})] + [\mathsf{E}(\xi^2_{(n-1:n)}) - \mathsf{E}(\xi^2_{(n)})] - [\mathsf{E}(\xi^\circ_{(n-1:n)}) - \mathsf{E}(\xi^\circ_{(n)})]\right\} \\
&= \frac{1}{n}\left\{[\mathsf{E}(\xi^1_{(n-1:n)}) + \mathsf{E}(\xi^2_{(n-1:n)}) - \mathsf{E}(\xi^\circ_{(n-1:n)})] - [\mathsf{E}(\xi^1_{(n)}) + \mathsf{E}(\xi^2_{(n)}) - \mathsf{E}(\xi^\circ_{(n)})]\right\} \\
&= \frac{1}{n}\left\{[\mathsf{E}(\xi^1_{(n-1:n)}) + \mathsf{E}(\xi^2_{(n-1:n)}) - \mathsf{E}(\xi^\circ_{(n-1:n)})] - \Delta^\xi_n\right\}.
\end{aligned}
$$

By Lemma 4, $\Delta^\xi_n \geq 0$. Then $\Delta^\xi_{n-1} - \Delta^\xi_n \leq 0$ if $\mathsf{E}(\xi^1_{(n-1:n)}) + \mathsf{E}(\xi^2_{(n-1:n)}) - \mathsf{E}(\xi^\circ_{(n-1:n)}) \leq 0$. By (A.7), there exists $\bar{n}$ such that when $n \leq \bar{n}$, $\mathsf{E}(\xi^1_{(n-1:n)}) + \mathsf{E}(\xi^2_{(n-1:n)}) - \mathsf{E}(\xi^\circ_{(n-1:n)}) \leq 0$. Thus, there exists $\bar{n}$ such that when $n \leq \bar{n}$, $\Delta^\xi_{n-1} - \Delta^\xi_n \leq 0$. The desired result holds. □

It is intuitive that the *best* performances in *both* contest mechanisms can be improved by having more contestants. However, it is unclear which contest mechanism benefits more from additional contestants. Proposition A.1 shows that the marginal benefit of one additional contestant in boosting the expected best random factors is more significant for the separate contest than for the joint contest if the number of contestants is not too large. The intuition is as follows. For both contests, the best performance will be enhanced only if the additional solution is better than every single one of the solutions in the existing pool. For the joint contest, contestants submit a single solution along the two dimensions, so the best performance will be improved if the additional aggregate solution is better. However, for the separate contest, the best performance will be improved if the additional solution in either attribute is better. When the pool of contestants is small, it is more likely that the additional contestant is doing better than the existing pool of contestants in one of the attributes than that he is doing better in the whole project. Thus, if the contestant pool is not too large, the firm can benefit from obtaining a higher expected best random factors, by having more contestants in the separate contest than in the joint contest. When the contestant pool is large, having more contestants may lead to diminishing returns. In Proposition A.2, we compare the two contest mechanisms in the expected best random factors when the number of contestants is large enough.

PROPOSITION A.2. *(i) If ξ^1 and ξ^2 have a bounded support $[-a, a]$,* $\lim_{n\to\infty} \Delta^\xi_n = 0$.
(ii) If ξ^1 and ξ^2 are normally distributed, $\lim_{n\to\infty} \Delta^\xi_n = \infty$.

Proof of Proposition A.2. First, we prove the following lemma.

LEMMA A.1. *If PDF $\psi(\xi^1)$ is symmetric and log-concave, then*

$$
\mathsf{E}(\xi^1_{(n)}) \geq \left(1 - \frac{1}{2^n}\right)\Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right).
$$

Proof of Lemma A.1. The expectation of the highest order statistics of ξ^1 can be written as

$$\mathsf{E}(\xi^1_{(n)}) = \int_{-\infty}^{+\infty} \xi^1 n\Psi(\xi^1)^{n-1} d\Psi(\xi^1) = \int_0^1 \Psi^{-1}(u) n u^{n-1} du,$$

by letting $\xi^1 = \Psi^{-1}(u)$. Let $B(u) = nu^{n-1}$, then we have

$$\begin{aligned}
\mathsf{E}(\xi^1_{(n)}) &= \int_0^1 \Psi^{-1}(u)B(u)du = \int_{1/2}^1 \Psi^{-1}(u)B(u)du + \int_0^{1/2} \Psi^{-1}(u)B(u)du \\
&= \int_{1/2}^1 \Psi^{-1}(u)B(u)du - \int_{1/2}^1 \Psi^{-1}(1-u)B(1-u)du \\
&= \int_{1/2}^1 \Psi^{-1}(u)B(u)du - \int_{1/2}^1 \Psi^{-1}(u)B(1-u)du \\
&= \int_{1/2}^1 \Psi^{-1}(u)[B(u) - B(1-u)]du,
\end{aligned}$$

where the fourth equality is by the symmetric property that $\psi(\xi^1) = \psi(-\xi^1)$, $\Psi(\xi^1) = 1 - \Psi(\xi^1)$ and $\Psi^{-1}(1-u) = \Psi^{-1}(u)$. By Lemma A.2, PDF $\psi(\xi^1)$ is unimodal and symmetric at 0, thus $\psi(\xi^1)$ is decreasing in $\xi^1 \geq 0$. When $\xi^1 \geq 0$, the CDF $\Psi(\xi^1)$ is concave because $\Psi''(\xi^1) = \psi'(\xi^1) \leq 0$. As a result, $\Psi^{-1}(u)$ is convex in $u \in [1/2, 1]$. Let $K = \int_{1/2}^1 B(u) - B(1-u)du = \int_{1/2}^1 [nu^{n-1} - n(1-u)^{n-1}]du = 1 - 1/2^n$, thus $\int_{1/2}^1 (B(u) - B(1-u))/K du = 1$. Since $(B(u) - B(1-u))/K$ can be considered as a PDF, then $\int_{1/2}^1 \Psi^{-1}(u)\{[B(u) - B(1-u)]/K\}du$ is the expectation of $\Psi^{-1}(u)$ with such PDF. By the convexity of $\Psi^{-1}(u)$ and Jensen's inequality, we have

$$\mathsf{E}(\xi^1_{(n)})/K = \int_{1/2}^1 \Psi^{-1}(u)\{[B(u) - B(1-u)]/K\}du \geq \Psi^{-1}\left(\int_{1/2}^1 u[B(u) - B(1-u)]/K du\right).$$

Integrating by parts, we have

$$\begin{aligned}
\int_{1/2}^1 u[B(u) - B(1-u)]du &= \int_{1/2}^1 [nu^n - n(1-u)^{n-1}u]du = \left(\frac{n}{n+1} - \frac{n}{n+1}\frac{1}{2^{n+1}}\right) + \int_{1/2}^1 ud(1-u)^n \\
&= \left(\frac{n}{n+1} - \frac{n}{n+1}\frac{1}{2^{n+1}}\right) - \frac{1}{2^{n+1}} - \int_{1/2}^1 (1-u)^n du \\
&= \left(\frac{n}{n+1} - \frac{n}{n+1}\frac{1}{2^{n+1}}\right) - \frac{1}{2^{n+1}} - \frac{1}{n+1}\frac{1}{2^{n+1}} = \frac{n}{n+1} - \frac{1}{2^n}.
\end{aligned}$$

Because $K = \int_{1/2}^1 [B(u) - B(1-u)]du = 1 - 1/2^n$, $\mathsf{E}(\xi^1_{(n)}) \geq K\Psi^{-1}\left(\int_{1/2}^1 u[B(u) - B(1-u)]/K du\right) \geq \left(1 - \frac{1}{2^n}\right)\Psi^{-1}\left(\left(\frac{n}{n+1} - \frac{1}{2^n}\right)\Big/\left(1 - \frac{1}{2^n}\right)\right) \geq \left(1 - \frac{1}{2^n}\right)\Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right)$, where the last inequality is because $\Psi^{-1}(\cdot)$ is increasing and $1 - \frac{1}{2^n} \leq 1$. Therefore, the desired inequality holds. □

Now we prove Proposition A.2.

(i) If ξ^1 and ξ^2 have a bounded support $[-a, a]$, then ξ° has the bounded support $[-2a, 2a]$. By Lemma A.1, $\left(1 - \frac{1}{2^n}\right)\Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right) \leq \mathsf{E}(\xi^1_{(n)}) \leq a$. We have $\lim\limits_{n\to\infty}\left(1 - \frac{1}{2^n}\right)\Psi^{-1}\left(\frac{n}{n+1} - \frac{1}{2^n}\right) \leq$

$\lim_{n\to\infty} \mathsf{E}(\xi^1_{(n)}) \leq \lim_{n\to\infty} a$. By the Squeeze Theorem, $\lim_{n\to\infty} \left(1-\frac{1}{2^n}\right)\Psi^{-1}\left(\frac{n}{n+1}-\frac{1}{2^n}\right) = a = \lim_{n\to\infty} \mathsf{E}(\xi^1_{(n)}) = a$. Similar results can be extended to ξ^2 and ξ°, thus $\lim_{n\to\infty} \Delta^\xi_n = \lim_{n\to\infty} \mathsf{E}(\xi^1_{(n)}) + \lim_{n\to\infty} \mathsf{E}(\xi^2_{(n)}) - \lim_{n\to\infty} \mathsf{E}(\xi^\circ_{(n)}) = a + a - 2a = 0$. The desired result holds.

(ii) If $\xi^l \sim N(0,\sigma)$, $l = 1, 2$, the quantile function can be written as

$$\varphi^{-1}(u) = \sigma\sqrt{2}\text{erf}^{-1}(2u-1), \tag{A.9}$$

where $u \in (0,1)$ and erf(u) is the error function, $\text{erf}(u) = \frac{1}{\pi}\int_{-u}^{u} e^{t^2} dt$. Then Δ^ξ_n can be written as

$$\begin{aligned}
\Delta^\xi_n &= \int_{-\infty}^{+\infty} \xi^1 n\varphi^{n-1}(\xi^1)d\varphi(\xi^1) + \int_{-\infty}^{+\infty} \xi^2 n\varphi^{n-1}(\xi^2)d\varphi(\xi^2) - \int_{-\infty}^{+\infty} \xi^\circ n\varphi^{\circ n-1}(\xi^\circ)d\varphi^\circ(\xi^\circ) \\
&= \int_0^1 n\varphi^{-1}(u)u^{n-1}du + \int_0^1 n\varphi^{-1}(u)u^{n-1}du - \int_0^1 n\varphi^{\circ -1}(u)u^{n-1}du \\
&= \int_0^1 n(2\varphi^{-1}(u) - \varphi^{\circ -1}(u))u^{n-1}du.
\end{aligned}$$

By (A.9), $2\varphi^{-1}(u) - \varphi^{\circ -1}(u) = (2\sqrt{2}\sigma - 2\sigma)\text{erf}^{-1}(2u-1)$, because $\xi^\circ \sim N(0, \sqrt{2}\sigma)$ and $\varphi^{\circ -1}(u) = 2\sigma\text{erf}^{-1}(2u-1)$. Hence, we can define a new random variable $\tilde{\xi}$ that is normally distributed with mean 0 and standard deviation $(2-\sqrt{2})\sigma$. Denote its CDF by $\tilde{\varphi}(\tilde{\xi})$, then $2\varphi^{-1}(u) - \varphi^{\circ -1}(u) = \tilde{\varphi}^{-1}(u)$. Thus

$$\Delta^\xi_n = \int_0^1 n\tilde{\varphi}^{-1}(u)u^{n-1}du = \mathsf{E}(\tilde{\xi}_{(n)}). \tag{A.10}$$

By Lemma A.1, $\lim_{n\to\infty} \left(1-\frac{1}{2^n}\right)\tilde{\varphi}^{-1}\left(\frac{n}{n+1}-\frac{1}{2^n}\right) \leq \lim_{n\to\infty} \mathsf{E}(\tilde{\xi}_{(n)})$. Because the normal distribution is defined on the $(-\infty,\infty)$, then $\lim_{n\to\infty} \left(1-\frac{1}{2^n}\right)\tilde{\varphi}^{-1}\left(\frac{n}{n+1}-\frac{1}{2^n}\right) = \infty$. Then if ξ^1 and ξ^2 follow normal distributin, $\lim_{n\to\infty} \Delta^\xi_n = \infty$. □

Proposition A.2 shows that the limiting behavior of Δ^ξ_n depends on the distributions of the random factors. It is intuitive that when the number of contestants is sufficiently large, the best performance in both contest mechanisms must be extraordinary. Since contestants are ex ante identical, the equilibrium effort is equal among all the contestants under each mechanism, so the firm ex post selects the best realized random factor in each contest. Thus there must be a realized random factor approaching its upper limit among a sufficiently large number of contestants. It is well known that when n is sufficiently large, the expectation of the highest order statistics is approximately equal to the value of $\frac{n}{n+1}$-th quantile, $\mathsf{E}(\xi_{(n)}) \approx \Upsilon^{-1}\left(\frac{n}{n+1}\right)$, where $\Upsilon^{-1}(\cdot)$ is the quantile function of $\Upsilon(\cdot)$ (see, e.g., David and Nagaraja 2003, pp. 80, (4.5.1)). When n is large enough, the term $\frac{n}{n+1}$ approaches 1 and $\mathsf{E}(\xi_{(n)})$ approaches the upper limit of the range of the random factor.

When the number of contestants is large enough and the random factors have a bounded support (e.g., two-sided truncated normal distribution), the expected best random factors in the two contest mechanisms are approximately equal, since they are both close to the upper bound of the range. However, with normally distributed random factors that have the unbounded support, the difference between the expected best random factors approaches infinity when the large pool of contestants grows even larger. Thus, the separate contest can benefit more from an increasing number of contestants than the joint contest, even when the pool of contestants is already very large. That is, the combination effect can be infinitely enhanced by more and more contestants.

LEMMA A.2. *If PDF $\psi(\xi)$ is twice continuously differentiable and log-concave, then it is unimodal.*

Proof of Lemma A.2. By the definition of twice differentiable and log-concave function, we have $\frac{\partial^2 \ln[\psi(\xi)]}{\partial \xi^2} = \frac{\partial}{\partial \xi}\left[\frac{\psi'(\xi)}{\psi(\xi)}\right] \leq 0$. Thus, for any $\xi_1 \leq \xi_2$, $\frac{\psi'(\xi_1)}{\psi(\xi_1)} \geq \frac{\psi'(\xi_2)}{\psi(\xi_2)}$, and equivalently,

$$\frac{\psi'(\xi_1)\psi(\xi_2) - \psi(\xi_1)\psi'(\xi_2)}{\psi(\xi_1)\psi(\xi_2)} \geq 0. \tag{A.11}$$

First, consider the case that there exists a ξ^* such that $\psi'(\xi^*) = 0$. In (A.11), let $\xi_2 = \xi^*$, then (A.11) implies that $\psi'(\xi_1)\psi(\xi^*) \geq 0$. Since $\psi(\xi^*) \geq 0$, $\psi'(\xi_1) \geq 0$ for $\xi_1 \leq \xi^*$. Similarly, in (A.11), let $\xi_1 = \xi^*$, then $-\psi'(\xi^*)\psi(\xi_2) \geq 0$. Thus, $\psi'(\xi_2) \leq 0$ for $\xi_2 \geq \xi^*$. Hence, if ξ^* exists, PDF $\psi(\xi)$ is increasing for $\xi \leq \xi^*$ and decreasing for $\xi \geq \xi^*$. Second, if ξ^* does not exist, because $\psi(\xi)$ is twice differentiable, $\psi(\xi)$ is either monotone increasing or decreasing. Thus, $\psi(\xi)$ is unimodal. □

LEMMA A.3. *If ξ follows a symmetric log-concave distribution with mean 0, then $h(\xi;n)$ is decreasing in n.*

Proof of Lemma A.3. By Ales et al. (2016a, p.12, Proposition 1), the equilibrium effort level e^* is non-increasing for any $n \geq 2$ if and only if the density $\psi(\xi)$ of the output shock ξ satisfies $\int_{-\infty}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi \leq 0$. Following the same approach, we now verify that any random factor that follows a symmetric log-concave distribution with mean 0 satisfies this condition. By Lemma A.2, the density function of a log-concave distribution is unimodal, which means that there exists ξ_0 such that for $\xi < \xi_0$, $\psi'(\xi) \geq 0$, for $\xi > \xi_0$, $\psi'(\xi) \leq 0$. When $n \geq 2$,

$$h(\xi;n+1) - h(\xi;n) = \int_{-\infty}^{+\infty} n\Psi(\xi)^{n-1}\psi(\xi)^2 d\xi - \int_{-\infty}^{+\infty} (n-1)\Psi(\xi)^{n-2}\psi(\xi)^2 d\xi$$

$$= \int_{-\infty}^{+\infty} \psi(\xi) d\Psi(\xi)^n - \int_{-\infty}^{+\infty} \psi(\xi) d\Psi(\xi)^{n-1} = -\int_{-\infty}^{+\infty} \psi'(\xi)\Psi(\xi)^n d\xi + \int_{-\infty}^{+\infty} \psi'(\xi)\Psi(\xi)^{n-1} d\xi$$

$$= \int_{-\infty}^{+\infty} (1-\Psi(\xi))\Psi(\xi)^{n-1}\psi'(\xi)d\xi = \int_{-\infty}^{\xi_0} (1-\Psi(\xi))\Psi(\xi)^{n-1}\psi'(\xi)d\xi + \int_{\xi_0}^{+\infty} (1-\Psi(\xi))\Psi(\xi)^{n-1}\psi'(\xi)d\xi$$

$$\leq \int_{-\infty}^{\xi_0}(1-\Psi(\xi))\Psi(\xi)\Psi(\xi_0)^{n-2}\psi'(\xi)d\xi + \int_{\xi_0}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\Psi(\xi_0)^{n-2}\psi'(\xi)d\xi$$
$$= \Psi(\xi_0)^{n-2}\int_{-\infty}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi,$$

where the third equality is due to the integration by parts. Now we verify that if the distribution is symmetric at 0, then $\int_{-\infty}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi = 0$. Because $1-\Psi(\xi)=\Psi(-\xi)$ and $\psi'(\xi)=\psi'(-\xi)$, we have

$$\begin{aligned}
&\int_{-\infty}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi\\
&= \int_{0}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi + \int_{-\infty}^{0}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi\\
&= \int_{0}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi + \int_{-\infty}^{0}(1-\Psi(-\xi))\Psi(-\xi)\psi'(-\xi)d\xi\\
&= \int_{0}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi - \int_{0}^{+\infty}(1-\Psi(\xi))\Psi(\xi)\psi'(\xi)d\xi = 0.
\end{aligned}$$

As a result, $h(\xi;n+1)-h(\xi;n)\leq 0$ and $h(\xi;n)$ is decreasing in n. □

Proof of Proposition 3. First, if the cost functions are in the polynomial form, i.e., $C^l(e^l) = a^l(e^l)^{b^l}$, $a^l>0, b^l\geq 2$, $l=1,2$, by (A.6), the difference of equilibrium efforts is

$$\begin{aligned}
\Delta_n^e &= (a^1b^1)^{-\frac{1}{b^1-1}}(wAh(\xi;n))^{\frac{1}{b^1-1}} + (a^2b^2)^{-\frac{1}{b^2-1}}((1-w)Ah(\xi;n))^{\frac{1}{b^2-1}}\\
&\quad -(a^1b^1)^{-\frac{1}{b^1-1}}(Ah^\circ(\xi^\circ;n))^{\frac{1}{b^1-1}} - (a^2b^2)^{-\frac{1}{b^2-1}}(Ah^\circ(\xi^\circ;n))^{\frac{1}{b^2-1}}\\
&= (A/a^1b^1)^{\frac{1}{b^1-1}}\left[(wh(\xi;n))^{\frac{1}{b^1-1}} - h^\circ(\xi^\circ;n)^{\frac{1}{b^1-1}}\right]\\
&\quad +(A/a^2b^2)^{\frac{1}{b^2-1}}\left[((1-w)h(\xi;n))^{\frac{1}{b^2-1}} - h^\circ(\xi^\circ;n)^{\frac{1}{b^2-1}}\right].
\end{aligned}\tag{A.12}$$

Since $\frac{h(\xi;n)}{h^\circ(\xi^\circ;n)}=\sqrt{2}$ for any n, we can rewrite (A.12) as

$$\begin{aligned}
\Delta_n^e &= (A/a^1b^1)^{\frac{1}{b^1-1}}\left[w^{\frac{1}{b^1-1}} - (\sqrt{2}/2)^{\frac{1}{b^1-1}}\right]h(\xi;n)^{\frac{1}{b^1-1}}\\
&\quad +(A/a^2b^2)^{\frac{1}{b^2-1}}\left[(1-w)^{\frac{1}{b^2-1}} - (\sqrt{2}/2)^{\frac{1}{b^2-1}}\right]h(\xi;n)^{\frac{1}{b^2-1}}.
\end{aligned}$$

By Lemma A.3, if ξ follows a symmetric log-concave distribution with mean 0, then $h(\xi;n)$ is decreasing in n. Given $w\in(1-\sqrt{2}/2,\sqrt{2}/2)$, $\Delta_n^e<0$ is increasing in n. Meanwhile, By Proposition A.2, $\lim_{n\to\infty}\Delta_n^\xi=\infty$, therefore the difference of the random factors Δ_n^ξ is a positive value that is increasing in n. As a result, there exists a threshold $\tilde{n}\in[2,\infty)$ such that when $n\leq\tilde{n}$, $\Delta_n = \Delta_n^\xi+\Delta_n^e\leq 0$, and when $n\geq\tilde{n}$, $\Delta_n=\Delta_n^\xi+\Delta_n^e\geq 0$.

Second, if the cost functions are in the exponential form, i.e., $C^l(e^l)=\exp(\rho^l e^l)$, $\rho^l>0$, $l=1,2$, by (A.6), the difference of equilibrium efforts is

$$\Delta_n^e = \ln\left(\frac{wh(\xi;n)}{h^\circ(\xi^\circ;n)}\right)/\rho^1 + \ln\left(\frac{(1-w)h(\xi;n)}{h^\circ(\xi^\circ;n)}\right)/\rho^2. \tag{A.13}$$

By Corollary 1, if $\xi^l \sim N(0,\sigma)$, $l=1,2$, then $1/H(n) = \frac{h(\xi;n)}{h^\circ(\xi^\circ;n)} = \sqrt{2}$ for any n. Since $w \in (1-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ and $\frac{h(\xi;n)}{h^\circ(\xi^\circ;n)} = \sqrt{2}$, $\Delta_n^e < 0$ by (A.13). Thus, Δ_n^e is a fixed non-positive value for any n. For Δ_n^ξ, by (A.10), $\Delta_n^\xi = \mathsf{E}(\tilde{\xi}_{(n)})$, where $\tilde{\xi} \sim N(0,(2-\sqrt{2})\sigma)$. By (A.8), $\mathsf{E}(\tilde{\xi}_{(n-1)}) - \mathsf{E}(\tilde{\xi}_{(n)}) = \frac{1}{n}[\mathsf{E}(\tilde{\xi}_{(n-1:n)}) - \mathsf{E}(\tilde{\xi}_{(n)})] \leq 0$. Thus, $\mathsf{E}(\tilde{\xi}_{(n)})$ is increasing in n. By Proposition A.2, $\lim_{n\to\infty} \Delta_n^\xi = \infty$, therefore the difference of the random factors Δ_n^ξ is a positive value that is increasing in n. There exists a threshold $\tilde{n} \in [2,\infty)$ such that when $n \leq \tilde{n}$, $\Delta_n = \Delta_n^\xi + \Delta_n^e \leq 0$, and when $n \geq \tilde{n}$, $\Delta_n = \Delta_n^\xi + \Delta_n^e \geq 0$. □

Proof of Lemma 5. By (A.6), the difference of equilibrium efforts between two contest mechanisms is $\Delta_n^e = C^{1'^{-1}}(Ah^\circ(\xi^\circ;n)) + C^{2'^{-1}}(Ah^\circ(\xi^\circ;n)) - C^{1'^{-1}}(A^1 h(\xi;n)) - C^{2'^{-1}}(A^2 h(\xi;n))$.

For part (i), if $C^l(e) = ae^b$, $l=1,2$, $a>0$, $b\geq 2$, then $C^{l'^{-1}}(x) = \left(\frac{x}{ab}\right)^{\frac{1}{b-1}}$. Thus,

$$\Delta_n^e = \left[\left(\frac{h^\circ(\xi^\circ;n)}{ab}\right)^{\frac{1}{b-1}} - \left(\frac{wh(\xi;n)}{ab}\right)^{\frac{1}{b-1}} + \left(\frac{h^\circ(\xi^\circ;n)}{ab}\right)^{\frac{1}{b-1}} - \left(\frac{(1-w)h(\xi;n)}{ab}\right)^{\frac{1}{b-1}}\right] A^{\frac{1}{b-1}},$$

where we denote $A^1 = wA$ and $A^2 = (1-w)A$. If $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n),(1-w)h(\xi;n)\}$, then

$$\left[\left(\frac{h^\circ(\xi^\circ;n)}{ab}\right)^{\frac{1}{b-1}} - \left(\frac{wh(\xi;n)}{ab}\right)^{\frac{1}{b-1}} + \left(\frac{h^\circ(\xi^\circ;n)}{ab}\right)^{\frac{1}{b-1}} - \left(\frac{(1-w)h(\xi;n)}{ab}\right)^{\frac{1}{b-1}}\right] > 0,$$

and thus Δ_n^e is strictly increasing in A.

For part (ii), if $C^l(e) = \exp(\rho e)$, $l=1,2$, $\rho>0$, then $C^{l'^{-1}}(x) = \frac{1}{\rho}\ln(\frac{x}{\rho})$. Thus,

$$\Delta_n^e = \frac{1}{\rho}\left[\ln\left(\frac{h^\circ(\xi^\circ;n)}{wh(\xi;n)}\right) + \ln\left(\frac{h^\circ(\xi^\circ;n)}{(1-w)h(\xi;n)}\right)\right],$$

where $A^1 = wA$ and $A^2 = (1-w)A$. Since Δ_n^e is independent from the total amount of prize A, Δ_n^e is fixed for any A. □

Proof of Proposition 4. We first derive a sufficient condition for the firm's utility U_o^J and U_o^S to be concave in the total prize A. Recall that the equilibrium effort level in the joint contest is $e^{\circ *}(n) = C^{\circ'^{-1}}(Ah^\circ(\xi^\circ;n))$. Since $U_o^J = C^{\circ'^{-1}}(Ah^\circ(\xi^\circ;n)) - A = C^{1'^{-1}}(Ah^\circ(\xi^\circ;n)) + C^{2'^{-1}}(Ah^\circ(\xi^\circ;n)) - A$. By denoting $C^{1'^{-1}}(\cdot) = \varphi^1(\cdot)$ and $C^{2'^{-1}}(\cdot) = \varphi^2(\cdot)$, we have that U_o^J is concave in A if $\frac{\partial^2 U_o^J}{\partial A^2} = \varphi^{1''}(Ah^\circ(\xi^\circ;n))h^{\circ 2}(\xi^\circ;n) + \varphi^{2''}(Ah^\circ(\xi^\circ;n))h^{\circ 2}(\xi^\circ;n) \leq 0$. It is equivalent to the condition that $\varphi^1(\cdot)$ and $\varphi^2(\cdot)$ are concave in the relevant range. Recall that the equilibrium effort level in the separate contest is $e^{l*}(n) = C^{l'^{-1}}(A^l h(\xi;n))$, $l=1,2$. The utility of the firm is $U_o^S = C^{1'^{-1}}(A^1 h(\xi;n)) + C^{2'^{-1}}(A^2 h(\xi;n)) - A$. The sufficient condition for the firm's utility U_o^S to be concave in the total prize A is $\frac{\partial^2 U_o^J}{\partial A^2} = \varphi^{1''}(wAh(\xi;n))(wh(\xi;n))^2 + \varphi^{2''}((1-w)Ah(\xi;n))((1-w)h(\xi;n))^2 \leq 0$. It is equivalent to the condition that $\varphi^1(\cdot)$ and $\varphi^2(\cdot)$ are concave in the relevant range.

Due to the concavity of U_o^J with respect to A, the optimal total prize A^{J*} in the joint contest is the solution to the following FOC of the firm's problem with respect to A: $\frac{\partial U_o^J}{\partial A} =$

${\varphi^1}'(A^{J*}h^\circ(\xi^\circ;n))h^\circ(\xi^\circ;n)+{\varphi^2}'(A^{J*}h^\circ(\xi^\circ;n))h^\circ(\xi^\circ;n)-1=0$. Due to the concavity of U_o^S with respect to A, the optimal total prize A^{S*} in the joint contest is the solution to the following FOC of the firm's problem with respect to A: $\frac{\partial U_o^S}{\partial A}={\varphi^1}'(wA^{S*}h(\xi;n))wh(\xi;n)+{\varphi^2}'((1-w)A^{S*}h(\xi;n))(1-w)h(\xi;n)-1=0$. Since $C^1(\cdot)=C^2(\cdot)$, we denote $C^1(\cdot)=C^2(\cdot)=C(\cdot)$ and $\varphi^1(\cdot)=\varphi^2(\cdot)=\varphi(\cdot)$ in the following proof. Now we can examine the two specific forms of the cost functions. It is simple to verify that the $\varphi(\cdot)$ of $C(e)=ae^b$, $a>0, b\geq 2$ and $C(e)=\exp(\rho e)$, $\rho>0$ is concave.

For part (i), if $C(e)=ae^b$, $a>0$, $b\geq 2$, then ${C'}^{-1}(x)=\left(\frac{x}{ab}\right)^{\frac{1}{b-1}}=\varphi(x)$. When $b>2$, $\varphi'(x)=\frac{1}{b-1}(\frac{1}{ab})^{\frac{1}{b-1}}x^{\frac{2-b}{b-1}}$. The optimal amount of prize in the joint contest A^{J*} is given by ${\varphi^1}'(A^{J*}h^\circ(\xi^\circ;n))h^\circ(\xi^\circ;n)+{\varphi^2}'(A^{J*}h^\circ(\xi^\circ;n))h^\circ(\xi^\circ;n)-1=0$, equivalently, $2\varphi'(A^{J*}h^\circ(\xi^\circ;n))h^\circ(\xi^\circ;n)=1$. Denote $X=\frac{1}{b-1}(\frac{1}{ab})^{\frac{1}{b-1}}>0$. Then, $(A^{J*})^{\frac{2-b}{b-1}}=[2X(h^\circ(\xi^\circ;n))^{\frac{1}{b-1}}]^{-1}$. Moreover, the optimal amount of prize in the separate contest A^{S*} is given by $\varphi'(wA^{S*}h(\xi;n))wh(\xi;n)+\varphi'((1-w)A^{S*}h(\xi;n))(1-w)h(\xi;n)=1$, equivalently, $X(wA^{S*}h(\xi;n))^{\frac{2-b}{b-1}}wh(\xi;n)+X((1-w)A^{S*}h(\xi;n))^{\frac{2-b}{b-1}}(1-w)h(\xi;n)=1$. Therefore, $(A^{S*})^{\frac{2-b}{b-1}}=\left[X[(wh(\xi;n))^{\frac{1}{b-1}}+((1-w)h(\xi;n))^{\frac{1}{b-1}}]\right]^{-1}$. Because $h^\circ(\xi^\circ;n)>\max\{wh(\xi;n),(1-w)h(\xi;n)\}$, $[2X(h^\circ(\xi^\circ;n))^{\frac{1}{b-1}}]^{-1}<\left[X[(wh(\xi;n))^{\frac{1}{b-1}}+((1-w)h(\xi;n))^{\frac{1}{b-1}}]\right]^{-1}$. Since $b>2$, $\frac{2-b}{b-1}<0$, we have $A^{J*}>A^{S*}$.

When $b=2$, $\varphi(x)=\frac{x}{ab}$, thus U_o^J and U_o^S are linear functions in A, i.e., $U_o^J=(\frac{2h^\circ(\xi^\circ;n)}{ab}-1)A$, and $U_o^S=(\frac{h(\xi;n)}{ab}-1)A$. There does not exist a finite optimal amount of the total prize for the two contest mechanisms.

For part (ii), if $C(e)=\exp(\rho e)$, $\rho>0$, then ${C'}^{-1}(x)=\frac{1}{\rho}\ln(\frac{x}{\rho})=\varphi(x)$. Then $\varphi'(x)=\frac{1}{\rho x}$. For the joint contest, since the optimal amount of prize A^{J*} is given by ${\varphi^1}'(A^{J*}h^\circ(\xi^\circ;n))h^\circ(\xi^\circ;n)+{\varphi^2}'(A^{J*}h^\circ(\xi^\circ;n))h^\circ(\xi^\circ;n)-1=0$, $\frac{2}{\rho A^{J*}}=1$, and thus $A^{J*}=\frac{2}{\rho}$. For the separate contest, since the optimal amount of prize A^{S*} is given by $\varphi'(wA^{S*}h(\xi;n))wh(\xi;n)+\varphi'((1-w)A^{S*}h(\xi;n))(1-w)h(\xi;n)=1$, we obtain $A^{S*}=\frac{2}{\rho}$. The optimal amounts of the total prize in the two contest mechanisms are the same. $\square$

B. Existence of Equilibrium

B.1. Separate Contest

According to Fudenberg and Tirole (1991), a pure strategy Nash equilibrium exists if the utility of contestant i, u_i^l, $l=1,2$, is quasi-concave in his effort level e_i^l in each sub-contest of the separate contest. Recall that $u_i^l(e_i^l)=A^l\mathsf{P}(e_i^l,e^{l*}(n))-C^l(e_i^l)$, where $\mathsf{P}(e_i^l,e^{l*}(n))$ denotes the winning probaility of contestant i if he or she makes effort e_i^l and other contestants make effort $e^{l*}(n)$, thus the second derivative to u_i^l is given by ${u_i^l}''(e_i^l)=A^l\frac{\partial^2 P(e_i^l,e^{l*}(n))}{\partial {e_i^l}^2}-{C^l}''(e_i^l)$. We next show the sufficient conditions for ${u_i^l}''(e_i^l)<0$, under which u_i^l will be concave which implies quasi-concavity. We follow the same approach as in Ales et al. (2016a), using a scale transformation of the random factor to $\tilde{\xi}^l = \alpha\xi^l$ with α. The α measures the dispersion of the random factor. Then the probability that contestant i wins contestant j is

$$\mathsf{P}(e_i^l + \tilde{\xi}^l \geq e_j^{l*}(n) + \tilde{\xi}^l) = \mathsf{P}(e_i^l - e_j^{l*}(n) + \tilde{\xi}^l \geq \tilde{\xi}^l) = \mathsf{P}\left(\frac{e_i^l - e_j^{l*}(n)}{\alpha} + \xi^l \geq \xi^l\right).$$

Therefore, $\mathsf{P}(e_i^l, e^{l*}(n)) = \int_{-\infty}^{\infty} \Psi\left(\frac{e_i^l - e^{l*}(n)}{\alpha} + \xi^l\right)^{n-1} \psi(\xi^l) d\xi^l$. The first derivative of $\mathsf{P}(e_i^l, e^{l*}(n))$ with respect to e_i^l is

$$\frac{\partial \mathsf{P}(e_i^l, e^{l*}(n))}{\partial e_i^l} = \frac{1}{\alpha}\int_{-\infty}^{\infty} (n-1)\Psi\left(\frac{e_i^l - e^{l*}(n)}{\alpha} + \xi^l\right)^{n-2} \psi\left(\frac{e_i^l - e^{l*}(n)}{\alpha} + \xi^l\right)\psi(\xi^l) d\xi^l.$$

Then, the second derivative of $\mathsf{P}(e_i^l, e^{l*}(n))$ with respect to e_i^l is

$$\begin{aligned}\frac{\partial^2 \mathsf{P}(e_i^l, e^{l*}(n))}{\partial {e_i^l}^2} &= \frac{1}{\alpha^2}\int_{-\infty}^{\infty} \Bigg[(n-1)(n-2)\Psi\left(\frac{e_i^l - e^{l*}(n)}{\alpha} + \xi^l\right)^{n-3} \psi^2\left(\frac{e_i^l - e^{l*}(n)}{\alpha} + \xi^l\right)\psi(\xi^l) \\ &\quad + (n-1)\Psi\left(\frac{e_i^l - e^{l*}(n)}{\alpha} + \xi^l\right)^{n-2} \psi'\left(\frac{e_i^l - e^{l*}(n)}{\alpha} + \xi^l\right)\psi(\xi^l)\Bigg] d\xi^l. \qquad \text{(B.1)}\end{aligned}$$

If $\alpha \to \infty$, the integration in (B.1) converges to a constant, thus if α is sufficiently large, the second derivative of $\mathsf{P}(e_i^l, e^{l*}(n))$ with respect to e_i^l approaches 0. Since ${C^l}''(e_i^l) > 0$, the sufficient condition for ${u_i^l}''(e_i^l)$ to be negative is that α is sufficiently large. Furthermore, another sufficient condition is that $C^l(e_i^l)$ is sufficiently convex, i.e., ${C^l}''(e_i^l)$ is sufficiently large. As a result, with those sufficient conditions, the existence of the equilibrium can be guaranteed.

B.2. Joint Contest

For the joint contest, we can follow a similar approach to derive the sufficient conditions. We show that the existence of equilibrium in the separate contest guarantees the existence of equilibrium in the joint contest. Recall that $u_i^\circ(e_i^\circ) = A\mathsf{P}^\circ(e_i^\circ, e^{\circ*}(n)) - C^\circ(e_i^\circ)$. Following a similar way in the discussion of the separate contest, we allow $\tilde{\xi}^l = \alpha\xi^l$, $l = 1, 2$, thus $\xi^\circ = \xi^1 + \xi^2 = \frac{1}{\alpha}(\tilde{\xi}^1 + \tilde{\xi}^2) = \frac{1}{\alpha}\tilde{\xi}^\circ$. Similarly, the second derivative of $\mathsf{P}^\circ(e_i^\circ, e^{\circ*}(n))$ with respect to e_i° is

$$\begin{aligned}\frac{\partial^2 \mathsf{P}^\circ(e_i^\circ, e^{\circ*}(n))}{\partial {e_i^\circ}^2} &= \frac{1}{\alpha^2}\int_{-\infty}^{\infty} \Bigg[(n-1)(n-2)\Psi^\circ\left(\frac{e_i^\circ - e^{\circ*}(n)}{\alpha} + \xi^\circ\right)^{n-3} \psi^{\circ 2}\left(\frac{e_i^\circ - e^{\circ*}(n)}{\alpha} + \xi^\circ\right)\psi^\circ(\xi^\circ) \\ &\quad + (n-1)\Psi^\circ\left(\frac{e_i^\circ - e^{\circ*}(n)}{\alpha} + \xi^\circ\right)^{n-2} \psi^{\circ\prime}\left(\frac{e_i^\circ - e^{\circ*}(n)}{\alpha} + \xi^\circ\right)\psi^\circ(\xi^\circ)\Bigg] d\xi^\circ,\end{aligned}$$

where $\psi^\circ(\cdot)$ is the PDF of ξ°. Therefore, a sufficient condition for the existence of equilibrium in the joint contest is that α is sufficiently large. Moreover, if the convexity of $C^l(\cdot)$, $l = 1, 2$ is sufficiently large, then the convexity of $C^\circ(\cdot)$ will be sufficiently large, since $C^\circ(e^\circ) = \min C^1(e^1) + C^2(e^2)$ such that $e^1 + e^2 = e^\circ$. As a result, the equilibrium exists with those sufficient conditions as claimed.

C. Individual Rationality

C.1. Separate Contest

For each sub-contest, in the symmetric equilibrium, contestants have the same effort level, and the winning probability is $1/n$ if there are n contestants participate. As a result, the individual rationality constraint is that

$$C^l(e^{l*}) = C^l[C^{l'^{-1}}(A^l h(\xi^l;n))] \leq \frac{A^l}{n}. \tag{C.1}$$

Similarly, we follow the same approach in Ales et al. (2016a) allowing $\tilde{\xi}^l = \alpha\xi^l$, then

$$\begin{aligned} h(\xi^l;n) &= \int_{-\infty}^{\infty} (n-1)\Psi(\xi^l)^{n-2}\psi(\xi^l)^2 d\xi^l = \frac{1}{\alpha}\int_{-\infty}^{\infty} (n-1)\Psi(\alpha\xi^l)^{n-2}\psi(\alpha\xi^l)^2 d\alpha\xi^l \\ &= \frac{1}{\alpha}\int_{-\infty}^{\infty} (n-1)\Psi(\tilde{\xi}^l)^{n-2}\psi(\tilde{\xi}^l)^2 d\tilde{\xi}^l = \frac{h(\tilde{\xi}^l;n)}{\alpha}. \end{aligned} \tag{C.2}$$

As a result, $C^l[C^{l'^{-1}}(A^l h(\xi^l;n))] = C^l[C^{l'^{-1}}(A^l h(\tilde{\xi}^l;n))/\alpha]$. Since $C^l(\cdot)$ and $C^{l'^{-1}}(\cdot)$ are strictly increasing, if the dispersion of the random factor measured by α is large enough, then (C.1) can be satisfied.

C.2. Joint Contest

For the joint contest, similarly, the individual rationality constraint is

$$C^\circ(e^{\circ*}) = C^\circ[C^{\circ'^{-1}}(Ah^\circ(\xi^\circ;n))] \leq \frac{A}{n}. \tag{C.3}$$

By the same approach, we allow $\tilde{\xi}^l = \alpha\xi^l$, $l = 1,2$, such that $\alpha\xi^\circ = \alpha(\xi^1 + \xi^2) = \tilde{\xi}^\circ$. As a result, by (C.2), $h^\circ(\xi^\circ;n) = h^\circ(\tilde{\xi}^\circ;n)/\alpha$. By Lemma 2, $C^\circ(\cdot)$ and $C^{\circ'^{-1}}(\cdot)$ are strictly increasing, thus if the dispersion of the random factor measured by α is large enough, the constraint must be satisfied.

C.3. Discussion

For constraints (C.1) and (C.3), if the dispersion of the random factor measured by α is large enough, then both of them can be satisfied. Now we demonstrate constraint (C.1) is neither a sufficient condition of constraint (C.3), nor vice versa.

By Lemma 2 that $C^{\circ'}(e^\circ) = C^{1'}(\tilde{e}^1) = C^{2'}(\tilde{e}^2)$ in which $\tilde{e}^1$ and $\tilde{e}^2$ are the optimal efforts in the two attributes given that $e^\circ = \tilde{e}^1 + \tilde{e}^2$, and $C^{\circ*'}(e^\circ) = Ah^\circ(\xi^\circ;n)$, there are $C^{1'^{-1}}(Ah^\circ(\xi^\circ;n)) = \tilde{e}^{1*}$ and $C^{2'^{-1}}(Ah^\circ(\xi^\circ;n)) = \tilde{e}^{2*}$. Thus, By the proof of Lemma 2, we have $C^\circ(e^\circ) = C^1(\tilde{e}^1) + C^2(\tilde{e}^2)$ $C^\circ(e^{\circ*}) = C^1(C^{1'^{-1}}(Ah^\circ(\xi^\circ;n))) + C^2(C^{2'^{-1}}(Ah^\circ(\xi^\circ;n)))$. Then, constraint (C.3) becomes

$$C^\circ(e^{\circ*}) = C^1(C^{1'^{-1}}(Ah^\circ(\xi^\circ;n))) + C^2(C^{2'^{-1}}(Ah^\circ(\xi^\circ;n))) \leq \frac{A}{n} = \frac{A^1 + A^2}{n}. \tag{C.4}$$

If the sufficient condition, $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n),(1-w)h(\xi;n)\}$, in Proposition 1 (ii) holds,

$$C^1(C^{1'^{-1}}(Ah^\circ(\xi^\circ;n))) > C^1[C^{1'^{-1}}(wAh(\xi^1;n))] = C^1[C^{1'^{-1}}(A^1h(\xi^1;n))],$$
$$C^2(C^{2'^{-1}}(Ah^\circ(\xi^\circ;n))) > C^2[C^{2'^{-1}}((1-w)Ah(\xi^1;n))] = C^2[C^{2'^{-1}}(A^2h(\xi^1;n))],$$

because $C^1(\cdot)$ and $C^2(\cdot)$ are strictly increasing and strictly convex functions. However, constraint (C.1) is generally neither a sufficient condition of (C.4), nor vice versa.

D. Budget Constraint

Assume that the firm has a budget constraint $\bar{A}$. The following discusses the optimal strategy of the firm if the cost functions are polynomial or exponential.

D.1. Polynomial Cost

By Lemma 5(i), the difference of effort levels Δ_n^e between two contest mechanisms is strictly increasing in A. And also, the difference of the expected random factors Δ_n^ξ is a fixed value given ρ. As a result, there exists a threshold $\hat{A}$ such that if $A < \hat{A}$, the separate contest is optimal and if $A > \hat{A}$, the joint contest is optimal. Note that the threshold can be 0 or $+\infty$ which means that one mechanism is optimal for any prize. If the cost functions are polynomial, then $A^{S*} < A^{J*}$, and there exist three cases: 1. $\hat{A} < A^{S*} < A^{J*}$; 2. $A^{S*} \le \hat{A} < A^{J*}$; 3. $A^{S*} < A^{J*} \le \hat{A}$. Table 4 shows the optimal mechanism and optimal prize when the budget contraint $\bar{A}$ is in different ranges if the cost functions are polynomial.

Table 4 Optimal Prize (Polynomial Cost)

Case 1: $\hat{A} < A^{S*} < A^{J*}$	Range of $\bar{A}$	$(0,\hat{A})$	$[\hat{A},A^{S*})$	$[A^{S*},A^{J*})$	$[A^{J*},+\infty)$
	Mechanism	separate	joint	joint	joint
	Optimal Prize	$\bar{A}$	$\bar{A}$	$\bar{A}$	A^{J*}
Case 2: $A^{S*} \le \hat{A} < A^{J*}$	Range of $\bar{A}$	$(0,A^{S*})$	$[A^{S*},\hat{A})$	$[\hat{A},A^{J*})$	$[A^{J*},+\infty)$
	Mechanism	separate	separate	joint	joint
	Optimal Prize	$\bar{A}$	A^{S*}	$\bar{A}$	A^{J*}
Case 3: $A^{S*} < A^{J*} \le \hat{A}$	Range of $\bar{A}$	$(0,A^{S*})$	$[A^{S*},A^{J*})$	$[A^{J*},\hat{A})$	$[\hat{A},+\infty)$
	Mechanism	separate	separate	separate	separate
	Optimal Prize	$\bar{A}$	A^{S*}	A^{S*}	A^{S*}

The detailed explanation for Table 4 is as follow:

1. ($\hat{A} < A^{S*} < A^{J*}$) If $\bar{A} \in (0,\hat{A})$, the separate contest is optimal, and the outcome is increasing in the amount of prize, then optimal amount of prize is $\bar{A}$. If $\bar{A} \in [\hat{A},A^{S*})$, the outcome of

both mechanisms is increasing in the amount of prize, and the joint contest is better than separate contest if $A > \hat{A}$, thus the firm should choose the joint contest and the optimal prize is $\bar{A}$. If $\bar{A} \in [A^{S*}, A^{J*})$, though the optimal amount of prize for the separate contest is in this range, the joint contest is better than the separate contest, so the firm should choose the joint contest and the optimal prize is $\bar{A}$. If $\bar{A} \in [A^{J*}, +\infty)$, then the firm should choose the joint contest and the optimal prize is A^{J*}.

2. $(A^{S*} \leq \hat{A} < A^{J*})$ If $\bar{A} \in (0, A^{S*})$, since the separate contest is optimal then the firm should choose the separate contest and the optimal prize is $\bar{A}$. If $\bar{A} \in [A^{S*}, \hat{A})$, the separate contest is better than the joint contest, and the optimal prize is A^{S*}. If $\bar{A} \in [\hat{A}, A^{J*})$, the outcome of the joint contest is better than that of the separate contest if $\bar{A} \geq \hat{A}$. Since the outcome of the joint contest is increasing in A, the optimal prize is $\bar{A}$. If $\bar{A} \in [A^{J*}, +\infty)$, the firm should choose the joint contest and the optimal prize is A^{J*}.
3. $(A^{S*} < A^{J*} \leq \hat{A})$ If $\bar{A} \in (0, A^{S*})$, the separate contest is optimal, and the optimal prize is $\bar{A}$. If $\bar{A} \in [A^{S*}, A^{J*})$, the separate contest is better than the joint contest, and the optimal prize is A^{S*}. If $\bar{A} \in [A^{J*}, \hat{A})$, the separate contest is still optimal, and the optimal prize is A^{S*}. If $\bar{A} \in [\hat{A}, +\infty]$, the outcome of both mechanisms is decreasing if $A \in [\hat{A}, +\infty)$, thus the separate contest is optimal and the optimal prize is A^{S*}.

D.2. Exponential Cost

By Lemma 5, the difference of effort levels Δ_n^e between two contest mechanisms is a fixed value for any A if the cost functions are exponential. Therefore, the problem which contest mechanism is optimal does not depend on the amount of prize.

Therefore, if the cost functions are exponential, then $A^{S*} = A^{J*}$, and there exist two cases: 1. $\bar{A} < A^*$; 2. $\bar{A} \geq A^*$ in which we denote $A^* = A^{S*} = A^{J*}$. If the separate (or the joint) contest is optimal and $\bar{A} < A^*$, then the optimal prize is $\bar{A}$ since the outcome is increasing in the amount of prize if $A < A^*$. If $\bar{A} \geq A^*$, then the optimal prize is A^*.

E. Elimination Contest

As an extension of the progressive contest, we examine the *elimination* contest here.

In the elimination contest, the firm selects a subset of qualified contestants from the first sub-contest to compete in the second sub-contest. Thus, there are fewer contestants competing in the second sub-contest than in the first sub-contest. On the other hand, the joint contest counterpart stays the same, because there is no way to impose a qualification. We first examine the performance in the second sub-contest with a contestant number $n_q < n$. Since the number of contestants becomes smaller than without qualification, the intensity of the competition becomes lower, and contestants have a higher incentive to make effort.[1] For the random factor part, the smaller number of contestants lowers the diversity of contestants, thus the expected best random factor is lower, i.e., a smaller sample size of the random factor leads to a lower expected highest order statistics of the random factor. The pooling effect, in favor of the joint contest, would be weakened since the equilibrium effort level in the second sub-contest of the separate contest becomes higher, and the combination effect, in favor of the separate contest, would be reduced too due to the smaller diversity of contestants in the second sub-contest. In addition, in the first sub-contest, in anticipation of elimination, contestants would work harder, in favor of the separate contest. The comparison of the two contest mechanisms depends on the relative strengths of those forces.

In the elimination contest, the firm selects a subset of qualified contestants from the first sub-contest to compete in the second sub-contest. Thus, there are fewer contestants competing in the second sub-contest than in the first sub-contest. On the other hand, the joint contest counterpart stays the same, because there is no way to impose a qualification in the joint contest. We first examine the performance in the second sub-contest with a contestant number $n_q < n$. Since the number of contestants becomes smaller than without qualification, the intensity of the competition becomes lower, and contestants have a higher incentive to make effort. For the random factor part, the smaller number of contestants lowers the diversity of contestants, thus the expected best random factor is lower, i.e., a smaller sample size of the random factor leads to a lower expected highest order statistics of the random factor. The pooling effect, in favor of the joint contest, would be weakened since the equilibrium effort level in the second sub-contest of the separate contest becomes higher, and the combination effect, in favor of the separate contest, would be reduced too due to the smaller diversity of contestants in the second sub-contest. In addition, in the first sub-contest, in anticipation of elimination, contestants would work harder, in favor of the separate contest. The comparison of the two contest mechanisms depends on the relative strengths of those forces.

LEMMA E.1. *(i) The best expectation of the random factor, $\mathsf{E}(\xi_{(n)})$, is increasing in n.*

(ii) If the random factor ξ follows a symmetric log-concave distribution with mean 0, then the equilibrium effort level is decreasing in n.

[1] How the number of contestants influences contestants' effort level has been well studied. Ales et al. (2019) show that if the random factor follows the symmetric log-concave distribution, the equilibrium effort level is decreasing in the number of contestants.

Proof. of Lemma E.1. For part (i), we verify that $\mathsf{E}(\xi_{(n)}) \geq \mathsf{E}(\xi_{(n-1)})$. By (A.8), we have $n\mathsf{E}(\xi_{(n-1)}) - (n-1)\mathsf{E}(\xi_{(n)}) = \mathsf{E}(\xi_{(n-1:n)})$. Thus, $\mathsf{E}(\xi_{(n-1)}) - \mathsf{E}(\xi_{(n)}) = [\mathsf{E}(\xi_{(n-1:n)}) - \mathsf{E}(\xi_{(n)})]/n$. Since $\mathsf{E}(\xi_{(n-1:n)}) - \mathsf{E}(\xi_{(n)}) \leq 0$, we have $\mathsf{E}(\xi_{(n-1)}) - \mathsf{E}(\xi_{(n)}) \leq 0$, and thus $\mathsf{E}(\xi_{(n)}) \geq \mathsf{E}(\xi_{(n-1)})$.

For part (ii), by Lemma A.3, $h(\xi;n)$ is decreasing in n. Since $C^l(\cdot)$, $l = 1, 2$, is strictly increasing and strictly convex, $e^{l*}(n+1) = C^{l'^{-1}}(h(\xi;n+1)) \leq C^{l'^{-1}}(h(\xi;n)) = e^{l*}(n)$. Equivalently, the equilibrium effort level is decreasing in n. ☐

The result in Lemma E.1 is consistent with the intuition. For part (i), with more contestants, it is more likely to have an outstanding realization of the random factor, thus if the number of contestants in the second sub-contest is less than that in the first sub-contest, the combination effect is reduced. For part (ii), since there is less contestants in the second sub-contest, the intensity of the competition is reduced, thus contestants have a stronger incentive to make effort in the second sub-contest. In other words, the pooling benefit of the joint contest is reduced too.

F. Synergy Effect

Synergy Effect. The solutions made by a contestant may have a synergy effect while the solutions made by different contestants do not. That is, the aggregate solution across different attributes submitted by the same contestant can have a synergy. The synergy effect intuitively favors the joint contest since every contestant submits an aggregate solution at once. To capture the synergy effect, we analyze the simplest model with an additive synergy component in Online Appendix F. We show that if the synergy effect is strong enough, the joint contest becomes optimal.

In this section, we compare the two contest mechanisms with the *synergy* effect. In some projects, the two attributes of a project can be closely related such that the aggregated solution of a contestant across different attributes is greater than the aggregation of solutions made by different contestants even if those contestants have the same effort level, i.e., $V_i^1 + V_i^2 \geq V_i^1 + V_j^2$. For example, if a project contains the two attributes as the theoretical and practical work, contestants may prefer to work on their own theoretical work when competing in doing the practical work. For the modular and progressive contests, we show that if the synergy effect is strong enough, then the joint contest is optimal.

F.1. Modular Contest

Consider a two-person model in which the contestants are indexed by i and j. If the solutions of the two attributes are made by the same contestant i, then the aggregated solution has an additive positive synergy benefit $\mu > 0$, i.e., $V_i^1 + V_i^2 + \mu$. If the solutions of the two attributes are made by different contestants, there is no such extra synergy benefit.

(i) *joint contest*

For the joint contest, contestants are required to submit an aggregate solution across the two attributes at once. Thus, the aggregate performance of every contestant has the synergy μ. Intuitively, in the joint contest, the synergy effect does not play a role in the competition since both contestants work on the two attributes at the same time. Therefore, the winning probability of each contestant is the same with what is captured in Lemma 3. By Lemma 3, the equilibrium effort of each contestant is $e^{\circ *} = {C^{\circ}}'^{-1}(Ah^{\circ}(\xi^{\circ};2))$. The expected best performance is then $V_2^J = e^{\circ *} + \mathsf{E}(\xi^{\circ}_{(2)}) + \mu$. By denoting $\gamma^{\circ} = \xi_i^1 - \xi_i^1 + \xi_i^2 - \xi_j^2$ with PDF $g^{\circ}(\gamma^{\circ})$, we have $h^{\circ}(\xi^{\circ};2) = g^{\circ}(0)$, and $e^{\circ *} = {C^{\circ}}'^{-1}(Ag^{\circ}(0))$.

(ii) *separate contest*

In the two sub-contests of the separate contest, contestants make efforts in different attributes. There are two scenarios. One is that contestant i or j wins both sub-contests. Then, the best aggregated solution is $V^S = V_t^1 + V_t^2 + \mu$, $t = i$ or j. The other is that contestant i wins the first sub-contest, and contestant j wins the second sub-contest, or the other way around. Then, for the firm. the best aggregate solution may not be the best solutions in the two sub-contests, $V^S = V_i^1 + V_j^2$ because of the synergy effect. In other words, it is possible that $V_i^1 + V_j^2 < V_i^1 + V_i^2 + \mu$. Since two contestants are assumed to be homogeneous, the two scenarios occur with equal probability, 1/2.

For the first scenario, the best aggregate performance is $V^S = e^{1*} + e^{2*} + \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(2)}) + \mu$. For the second scenario, the discussion is a little bit complicated. Suppose that contestant i wins the first sub-contest, and contestant j wins the second sub-contest. In the modular contest, the two sub-contest are parallel, contestant's effort level is not affected by the synergy effect since they are competing in each attribute (in the progressive contest, the synergy effect does influence contestants' effort level; we investigate such influence in the subsequent section). However, the aggregate solution depends on the synergy effect. Now we show that if the synergy effect is strong enough, then the joint contest is optimal.

Consider the random factor part and the synergy effect, and denote the first order statistics of ξ^l, $l = 1, 2$ with the sample size 2 by subscript $_{(1:2)}$. If the synergy effect is strong enough, then $\mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(1:2)}) + \mu > \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(2)})$. This means that if the synergy effect is strong enough, then though the winners of the two sub-contests are different contestants, the best aggregate solution is made by the same person. With probability 1/2, the difference in the random factor and synergy effect between the two contest mechanisms is $\Delta_2^{\xi} = \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(1:2)}) + \mu - \mathsf{E}(\xi^{\circ}_{(2)}) - \mu = \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(1:2)}) - \mathsf{E}(\xi^{\circ}_{(2)})$. For $\mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(1:2)})$, since ξ^1 and ξ^2 are identical and they follow a symmetric log-concave distribution $\Psi(\cdot)$ with mean 0, we have

$$\mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(1:2)}) = \int_{-\infty}^{+\infty} \xi^1 d\Psi(\xi^1)^2 + \int_{-\infty}^{+\infty} \xi^2 d(2\Psi(\xi^2)(1-\Psi(\xi^2)) + \Psi(\xi^2)^2)$$
$$= 2\int_{-\infty}^{+\infty} \xi^1 d\Psi(\xi^1) = 0.$$

With probability 1/2, the winner of both sub-contests is i or j. Therefore, the difference in the random factor and synergy effect between the two contest mechanisms is $\Delta_2^\xi = \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(2)}) + \mu - \mathsf{E}(\xi^\circ_{(2)}) - \mu = \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(2)}) - \mathsf{E}(\xi^\circ_{(2)})$. The ex ante expected difference is $\mathsf{E}(\Delta_2^\xi) = \frac{1}{2}[\mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(1:2)}) - \mathsf{E}(\xi^\circ_{(2)})] + \frac{1}{2}[\mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(2)}) - \mathsf{E}(\xi^\circ_{(2)})] = \mathsf{E}(\xi^1_{(2)}) - \mathsf{E}(\xi^\circ_{(2)})$, since ξ^1 and ξ^2 are identical. Because $\xi^\circ = \xi^1 + \xi^2$, $\xi^\circ \geq_{\text{st}} \xi^1$, where $_{\text{st}}$ means the first order stochastic dominance, thus $\mathsf{E}(\xi^1_{(2)}) - \mathsf{E}(\xi^\circ_{(2)}) \leq 0$. As a result, $\mathsf{E}(\Delta_2^\xi) \leq 0$.

By Proposition 1(ii), the pooling effect, $\Delta_2^e < 0$, we obtain $\Delta_2 = \Delta_2^e + \Delta_2^\xi < 0$. In conclusion, if the synergy effect, μ, is high enough, the joint contest is optimal.

F.2. Progressive Contest

For the progressive contest, in the second sub-contest of the separate contest, contestants are required to build their work in progress over the solutions in the first sub-contest. We discuss two scenarios in the second sub-contest: (a) contestants are required to work on the best solution generated in the first sub-contest; (b) contestants work on their own solution generated in the first sub-contest. Now, we show that if the synergy effect is strong enough, the joint contest is optimal for scenario (a).

(i) *joint contest*

For the joint contest, contestants are required to submit an aggregated solutions across the two attributes at once. Thus, the aggregated performance of every contestants has the synergy γ. Therefore, the winning probability of each contestant is the same with what is characterized in Lemma 3. By Lemma 3, the equilibrium effort of each contestant is $e^{\circ *} = {C^{\circ}}'^{-1}(Ah^\circ(\xi^\circ;2))$. The expected best performance is then $V_2^{sim} = e^{\circ *} + \mathsf{E}(\xi^\circ_{(2)}) + \mu$. By denoting $\gamma^\circ = \xi^1_i - \xi^1_i + \xi^2_i - \xi^2_j$ with CDF $G^\circ(\gamma^\circ)$, $h^\circ(\xi^\circ;2) = g^\circ(0)$, and $e^{\circ *} = {C^{\circ}}'^{-1}(Ag^\circ(0))$.

(ii) *separate contest*

Without loss of generality, consider that the winner in the first sub-contest is contestant i. In the second sub-contest, by denoting the random variable $\gamma^2 = \xi^2_i - \xi^2_j$ with CDF $G(\gamma^2)$ and PDF $g(\gamma^2)$, we have the winning probability of contestant i as $\mathsf{P}\{e^2_i + \xi^2_i + \mu > e^2_j + \xi^2_j\} = G(e^2_i - e^2_j + \mu)$, since contestant i benefits by the synergy effect but contestant j does not. Similarly, the winning probability of the contestant j is $G(-e^2_i + e^2_j - \mu)$. As a result, the expected utility functions for contestants i and j are

$$\mathsf{E}(u^2_i(e^2_i)) = A^2 G(e^2_i - e^2_j + \mu) - C^2(e^2_i),$$

$$\mathsf{E}(u_j^2(e_j^2)) = A^2G(e_j^2 - e_i^2 - \mu) - C^2(e_j^2).$$

Then, by assuming the existence of the equilibrium, FOCs are given by

$$A^2g(e_i^2 - e_j^2 + \mu) - C^{2'}(e_i^2) = 0,$$
$$A^2g(e_j^2 - e_i^2 - \mu) - C^{2'}(e_j^2) = 0.$$

Therefore, if the random factors follow a symmetric distribution, then $g(e_i^2 - e_j^2 + \mu) = g(e_j^2 - e_i^2 - \mu)$. Thus, there exists a unique solution $e_i^{2*} = e_j^{2*} = g(\mu)$. By Lemma 1, if the random factors follow a symmetric log-concave distribution with mean 0, then $h(\xi^l;2) = g(0) \geq g(\mu)$. Thus, the performance of the two contestants is $V_i^2 = C^{2'^{-1}}(A^2g(\mu)) + \xi^2 + \mu$ and $V_j^2 = C^{2'^{-1}}(A^2g(\mu)) + \xi^2$. The expected best performance in the second sub-contest contest is $\mathsf{E}(\max\{V_i^2, V_j^2\}) = C^{2'^{-1}}(A^2g(\mu)) + \mathsf{E}(\max\{\xi^2 + \mu, \xi^2\})$. Denote the random variable of $\max\{\xi^2 + \mu, \xi^2\}$ by ξ_m. The CDF of ξ_m is $\Psi(\xi^2)\Psi(\xi^2 - \mu)$ and then $\mathsf{E}(\max\{V_i^2, V_j^2\}) = \int_{-\infty}^{+\infty} \xi_m d\Psi(\xi^2)\Psi(\xi^2 - \mu)$. By the property of the log-concave distribution, $\Psi(\xi^2)\Psi(\xi^2 - \mu) \leq \Psi(\xi^2 - \mu/2)\Psi(\xi^2 - \mu/2)$. As a result, $\int_{-\infty}^{+\infty} \xi_m d\Psi(\xi^2)\Psi(\xi^2 - \mu) \leq \int_{-\infty}^{+\infty} \xi_m d\Psi(\xi^2 - \mu/2)\Psi(\xi^2 - \mu/2) = \mathsf{E}(\xi^2_{(2)} + \mu/2) = \mathsf{E}(\xi^2_{(2)}) + \mu/2$.

Now, we can compare the two contest mechanisms. The expected best performance in the joint contest is $V_2^j = e^{\circ *} + \mathsf{E}(\xi^\circ_{(2)}) + \mu$ and the expectid best performance in the separate contest is $V_2^s = e^{1*} + e^{2*} + \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\max\{\xi^2 + \mu, \xi^2\})$. The difference is $\Delta_2 = e^{1*} + e^{2*} - e^{\circ *} + \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\max\{\xi^2 + \mu, \xi^2\}) - \mathsf{E}(\xi^\circ_{(2)}) + \mu$.

For the effort part, $\Delta_2^e = e^{1*} + e^{2*} - e^{\circ *}$. By Proposition 1 (ii), if $h^\circ(\xi^\circ;2) > \max\{wh(\xi;2), (1-w)h(\xi;2)\}$, namely, $g^\circ(0) > \max\{wg(0), (1-w)g(0)\}$, $\Delta_2^e < 0$.

For the random factor part, $\Delta_2^\xi = \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\max\{\xi^2 + \mu, \xi^2\}) - \mathsf{E}(\xi^\circ_{(2)}) - \mu \leq \mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(2)}) - \mathsf{E}(\xi^\circ_{(2)}) - \mu/2$. Since $\mathsf{E}(\xi^1_{(2)}) + \mathsf{E}(\xi^2_{(2)}) - \mathsf{E}(\xi^\circ_{(2)})$ is a fixed value. If μ is large enough, then $\Delta_2^\xi \leq 0$.

Thus, if μ is large enough, then $\Delta_2 = \Delta_2^e + \Delta_2^\xi < 0$, and the joint contest is optimal.

For scenario (b), if contestants work on their own solution generated in the first sub-contest. Then the solutions of both contestants will contain the positive synergy benefit μ. Therefore, the second sub-contest is equivalent to the first sub-contest. The comparison between the two contest mechanisms is equivalent to the comparison between the two effects: the combination and pooling effects, without the synergy effect.

G. Attributes with Different Importance Levels

Now we show the result of Proposition 1 assuming that two attributes have different importance levels. In the separate contest, assume that the performance of a contestant in sub-contest l, $l = 1, 2$, is $V_i^l = e_i^l + \xi_i^l$. The total performance of a contestant is $V_i^{seq} = \theta V_i^1 + (1-\theta)V_i^2$, where

$\theta \in (0,1)$ denotes the relative importance between the two attributes. In the joint contest, the performance of a contestant is the aggregation of the performances along two attributes, denoted by $V^{sim} = \theta V_i^1 + (1-\theta)V_i^2 = \theta(e_i^1 + \xi_i^1) + (1-\theta)(e_i^2 + \xi_i^2) = \theta e_i^1 + (1-\theta)e_i^2 + \theta\xi_i^1 + (1-\theta)\xi_i^2 = e_i^\circ + \xi_i^\circ$, where $e_i^\circ = \theta e_i^1 + (1-\theta)e_i^2$ and $\xi_i^\circ = \theta\xi_i^1 + (1-\theta)\xi_i^2$.

Now we derive the optimal effort allocation by contestant with the above model setup. Similar to the proof of Lemma 2, the optimization problem for contestant is

$$\min_{e^1, e^2} C^1(e^1) + C^2(e^2) \quad \text{s.t. } \theta e^1 + (1-\theta)e^2 = e^\circ.$$

The solution to this problem can be typically found by writing the Lagrangean, $L(e^1, e^2, e^\circ; \lambda) = C^1(e^1) + C^2(e^2) + \lambda(e^\circ - \theta e^1 - (1-\theta)e^2)$ and the FOCs are

$$\begin{aligned} &\frac{\partial L}{\partial e^1} = C^{1\prime}(\tilde{e}^1) - \theta\tilde{\lambda} = 0 \quad (a), \quad \frac{\partial L}{\partial e^2} = C^{2\prime}(\tilde{e}^2) - (1-\theta)\tilde{\lambda} = 0 \quad (b), \\ &\frac{\partial L}{\partial \lambda} = e^\circ - \theta\tilde{e}^1 - (1-\theta)\tilde{e}^2 = 0 \quad (c). \end{aligned}$$

Following the same approach as in the proof of Lemma 2, one can obtain that $C^{1\prime}(\tilde{e}^1)/\theta = C^{2\prime}(\tilde{e}^2)/(1-\theta) = C^{\circ\prime}(e^\circ)$, where $C^\circ(e^\circ) = \min_{\theta e^1 + (1-\theta)e^2 = e^\circ} \{C^1(e^1) + C^2(e^2)\}$. By Lemma 1, the equilibrium effort in the separate contest is then $e^{l*} = C^{l\prime^{-1}}(A^l h(\xi; n))$, $l = 1, 2$. The expected best performance in the separate contest is then $V_n^{seq} = \theta e^{1*} + (1-\theta)e^{2*} + \theta\mathsf{E}(\xi_{(n)}) + (1-\theta)\mathsf{E}(\xi_{(n)}) = \theta C^{1\prime^{-1}}(A^1 h(\xi; n)) + (1-\theta)C^{2\prime^{-1}}(A^2 h(\xi; n)) + \theta\mathsf{E}(\xi_{(n)}) + (1-\theta)\mathsf{E}(\xi_{(n)})$. For the joint contest, with a similar proof of Lemma 3, the equilibrium effort is $e^{\circ*} = C^{\circ\prime^{-1}}(Ah^\circ(\xi^\circ; n)) = \theta C^{1\prime^{-1}}(\theta A h^\circ(\xi^\circ; n)) + (1-\theta)C^{2\prime^{-1}}((1-\theta)Ah^\circ(\xi^\circ; n))$. Therefore, the expected best performance in the joint contest is $V_n^{sim} = e^{\circ*} + \mathsf{E}(\xi^\circ_{(n)}) = \theta C^{1\prime^{-1}}(\theta A h^\circ(\xi^\circ; n)) + (1-\theta)C^{2\prime^{-1}}((1-\theta)Ah^\circ(\xi^\circ; n)) + \mathsf{E}(\xi^\circ_{(n)})$.

Now, we show the combination and pooling effects, which are the same as described in Proposition 1 parts (a) and (b). For part (a), denote the quantile function of ξ^l, $l = 1, 2$, by $\Psi^{-1}(u)$ and the quantile function of ξ° by $\Psi^{\circ -1}(u)$. Write the formula of $\mathsf{E}(\xi^l_{(n)})$, $l = 1, 2$,

$$\mathsf{E}(\xi^l_{(n)}) = \int_{-\infty}^{+\infty} \xi^1 n \Psi(\xi^l)^{n-1} \psi(\xi^l) d\xi^l = \int_{-\infty}^{+\infty} \xi^1 n \Psi(\xi^l)^{n-1} d\Psi(\xi^l) = \int_0^1 \Psi^{-1}(u) n u^{n-1} du, \quad \text{(G.1)}$$

where the last equality is obtained by substituting $\Psi^{-1}(u) = \xi^l$. Similarly, $\mathsf{E}(\xi^\circ_{(n)}) = \int_0^1 \Psi^{\circ -1}(u) n u^{n-1} du$. Then,

$$\begin{aligned} &\mathsf{E}(\theta\xi^1_{(n)}) + \mathsf{E}((1-\theta)\xi^2_{(n)}) - \mathsf{E}((\theta\xi^1 + (1-\theta)\xi^2)_{(n)}) = \theta\mathsf{E}(\xi^1_{(n)}) + (1-\theta)\mathsf{E}(\xi^2_{(n)}) - \mathsf{E}(\xi^\circ_{(n)}) \\ &= \int_0^1 \Psi^{-1}(u) n u^{n-1} du - \int_0^1 \Psi^{\circ -1}(u) n u^{n-1} du = \int_0^1 (\Psi^{-1}(u) - \Psi^{\circ -1}(u)) n u^{n-1} du. \quad \text{(G.2)} \end{aligned}$$

Assume that $\theta\xi^1$ and $(1-\theta)\xi^2$, satisfy the regularity condition that there exists $u_0 \in (0,1)$ such that $\Psi^{-1}(u) - \Psi^{\circ -1}(u) < 0$ if $u \in (0, u_0)$, and $\Psi^{-1}(u) - \Psi^{\circ -1}(u) > 0$ if $u \in (u_0, 1)$. Thus, by (G.2), we have

$$\begin{aligned}
&\int_0^{u_0} (\Psi^{-1}(u) - \Psi^{\circ -1}(u)) n u^{n-1} du + \int_{u_0}^1 (\Psi^{-1}(u) - \Psi^{\circ -1}(u)) n u^{n-1} du \\
> &\int_0^{u_0} (\Psi^{-1}(u) - \Psi^{\circ -1}(u)) n u_0^{n-1} du + \int_{u_0}^1 (\Psi^{-1}(u) - \Psi^{\circ -1}(u)) n u_0^{n-1} du \\
= &\int_0^1 (\Psi^{-1}(u) - \Psi^{\circ -1}(u)) n u_0^{n-1} du = n u_0^{n-1} \int_0^1 (\Psi^{-1}(u) - \Psi^{\circ -1}(u)) du \\
= &\, n u_0^{n-1} [\theta \mathsf{E}(\xi^1) + (1-\theta)\mathsf{E}(\xi^2) - \mathsf{E}(\theta\xi^1 + (1-\theta)\xi^2)] \\
= &\, 0.
\end{aligned}$$

Thus, the strict inequality holds for non-identical random factors by a similar argument to that for identical random factors. For the pooling effect, the sufficient condition is then $h^\circ(\xi^\circ; n) > \max\{wh(\xi;n)/\theta, (1-w)h(\xi;n)/(1-\theta)\}$, where $A^1 = wA$ and $A^2 = (1-w)A$, $w \in (0,1)$. As a result, the only difference is that the sufficient condition for the pooling effect now depends on w and θ. Since we allow an arbitrary w throughout the whole paper, the pooling effect must exist for a range of w.

Now we show the range of w for the pooling effect to hold with normally distributed random factors. Since $\xi^\circ = \theta\xi + (1-\theta)\xi$, if $\xi \sim N(0,\sigma)$ then $\xi^\circ \sim N(0, \sqrt{\theta^2 + (1-\theta)^2}\sigma)$. We have

$$h(\xi;n) = \int_{-\infty}^{+\infty} (n-1)\Psi(\xi)^{n-2}\psi(\xi)^2 d\xi = \int_{-\infty}^{+\infty} \psi(\xi) d\Psi(\xi)^{n-1} = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \exp(-\xi^2/(2\sigma^2)) d\Psi(\xi)^{n-1}.$$

Substitute ξ/σ with y, then $h(\xi;n) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^1 \exp(-y^2/2) d\varphi(y)^{n-1}$, where $\varphi(y) \sim N(0,1)$. Similarly, for $h^\circ(\xi^\circ;n)$, by substituting $\xi^\circ/(\sqrt{\theta^2+(1-\theta)^2}\sigma)$ with $\tilde{y}$,

$$\begin{aligned}
h^\circ(\xi^\circ;n) &= \int_{-\infty}^{+\infty} \psi^\circ(\xi^\circ) d\Psi^\circ(\xi^\circ)^{n-1} = \frac{1}{\sqrt{2(\theta^2+(1-\theta)^2)\pi}\sigma} \int_{-\infty}^{+\infty} \exp(-\xi^{\circ 2}/(4\sigma^2)) d\Psi^\circ(\xi^\circ)^{n-1} \\
&= \frac{1}{\sqrt{2(\theta^2+(1-\theta)^2)\pi}\sigma} \int_{-\infty}^{+\infty} \exp(-\tilde{y}^2/2) d\varphi(\tilde{y})^{n-1}.
\end{aligned}$$

Then, $\frac{h^\circ(\xi^\circ;n)}{h(\xi;n)} = \frac{1}{\sqrt{\theta^2+(1-\theta)^2}}$. The condition $h^\circ(\xi^\circ;n) > \max\{wh(\xi;n)/\theta, (1-w)h(\xi;n)/(1-\theta)\}$ is equivalent to $1 > \max\{w\sqrt{\theta^2+(1-\theta)^2}/\theta, (1-w)\sqrt{\theta^2+(1-\theta)^2}/(1-\theta)\}$. Then, the range of w is $(1 - \frac{1-\theta}{\sqrt{\theta^2+(1-\theta)^2}}, \frac{\theta}{\sqrt{\theta^2+(1-\theta)^2}})$. Specifically, if $\theta = 1/2$, the range of w is $(1-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, which is the same as what we derived in the base model.

H. Heterogeneous Contestants

H.1. Model Setup

We consider a two-person model with two expertise types (high and low) in each attribute. In contrast to the base model in which all the contestants are assumed to be identical for each attribute, we assume here that contestants are endowed with expertise either x_H or x_L, $x_H \geq x_L > 0$ in each attribute. The expertise in each attribute follows a two-point distribution. The probability that a contestant is endowed with x_L in each attribute is $\eta \in (0,1)$, and the probability that a contestant is endowed with x_H is $1-\eta$. Suppose we index the two contestants by i and j. The random factor in each attribute follows the normal distribution $N(0,\sigma)$, and thus the difference of random factors between two contestants i and j is denoted by $\gamma^l = \xi_i^l - \xi_j^l$, $l = 1,2$ following the normal distribution $N(0,\sqrt{2}\sigma)$ with PDF $g(\cdot)$ and CDF $G(\cdot)$.

We assume that the cost functions along the two dimensions are identical in the exponential form, $C(\cdot) = C^1(\cdot) = C^2(\cdot) = \exp(\rho e)$. In each sub-contest, every contestant knows only his own expertise type and that his opponents' expertise is drawn independently from the two-point distribution. The game is a Bayesian game in the Harsanyi sense (see Harsanyi 1968), where "types" are defined by contestants' expertise. In the symmetric Bayesian equilibrium, contestants' behavior is determined by their types, regardless of their identities. Hence, we use type H or L to refer to a contestant's behavior in equilibrium. Also, we call a contestant with high or low expertise in each attribute as H-type or L-type contestant respectively.

There are several ways to model the behavior of heterogeneous contestants. A commonly used setting is that different levels of expertise results in different efficiencies in making efforts.[2] For exerting the same amount of effort, an H-type contestant incurs a lower cost than an L-type contestant. Such a model characterizes the heterogeneity of contestants in their innovation ability. Contestants with higher talents tend to spend less time in developing novel ideas. It is reasonable to use the heterogeneous cost model to characterize contestant behavior for projects that require innovative thinking, such as ideation and art designing contests. For each attribute, if a contestant is type $t = H, L$, his cost function is $C(\cdot)/x_t$, $l = 1,2$. Similar characterizations have been adopted in Lazear and Rosen (1981), Moldovanu and Sela (2001) and Fey (2008) with slightly different model setups. The following lemma characterizes the equilibrium effort level in each sub-contest of the separate contest if contestants' expertise is independent along the two attributes, or contestants'

[2] Another setting is that different levels of expertise provides different starting points. That model characterizes the heterogeneity of contestants in their skill levels or experience. A skilled programmer may have mastered several well-developed programming frameworks. An experienced salesperson may have kept in contact with several clients so that in the sales contest he can guarantee a certain sales volume at the beginning of the competition. Such a characterization was adopted in Terwiesch and Xu (2008) and Körpeoğlu and Cho (2018).

expertise is correlated but contestants are not strategic. Based on the equilibrium, we discuss the strategic behavior of contestants in the subsequent section.

Lemma H.1 (Bayesian Equilibrium). *In the sub-contest l, $l=1,2$, there exists an equilibrium such that $e_H^{l*} \geq e_L^{l*}$. If $\eta = 1/2$, such equilibrium is unique and the effort levels have the following closed forms: $e_L^{l*} = \ln(A^l K_0 x_L/\rho)\rho$ and $e_H^{l*} = \ln(A^l K_0 x_H/\rho)\rho$, where $K_0 = g(0)/2 + g(\ln(x_H/x_L)/\rho)/2$.*

Proof of Lemma H.1. In the sub-contest l, $l=1,2$, of the separate contest, if contestant j with type H or L makes effort e_H^{l*} or e_L^{l*} in equilibrium, respectively, the winning probability of contestant i is $\eta G(e_i^l - e_L^{l*}) + (1-\eta) G(e_i^l - e_H^{l*})$. The expected payoff to contestant i is $\mathsf{E}(u_i^l(e_i^l|x_i)) = A^l[\eta G(e_i^l - e_L^{l*}) + (1-\eta) G(e_i^l - e_H^{l*})] - C(e_i^l)/x_i$. The FOC yields $A^l[\eta g(e_i^{l*} - e_L^{l*}) + (1-\eta) g(e_i^{l*} - e_H^{l*})] = C'(e_i^{l*})/x_i$. In a symmetric equilibrium, contestant i makes effort e_L^{l*} if he is L-type and e_H^{l*} if he is H-type, which lead to

$$A^l[\eta g(0) + (1-\eta) g(e_L^{l*} - e_H^{l*})] = C'(e_L^{l*})/x_L, \tag{H.1}$$

$$A^l[\eta g(e_H^{l*} - e_L^{l*}) + (1-\eta) g(0)] = C'(e_H^{l*})/x_H. \tag{H.2}$$

Now we prove that there exists an equilibrium such that $e_L^{l*} \leq e_H^{l*}$. For notation simplicity, we suppress the superscript l in the proof of $e_L^{l*} \leq e_H^{l*}$. Denote $e_H^* - e_L^* = \delta_{H-L}$. We want to show that there exists a $\delta_{H-L} \geq 0$. Dividing (H.1) by (H.2), we have

$$\frac{\eta g(0) + (1-\eta) g(-\delta_{H-L})}{\eta g(\delta_{H-L}) + (1-\eta) g(0)} = \frac{x_H C'(e_L^*)}{x_L C'(e_H^*)}. \tag{H.3}$$

By the symmetric assumption of $g(\cdot)$, we have $g(\delta_{H-L}) = g(-\delta_{H-L})$. Then (H.3) becomes

$$\frac{\eta g(0) + (1-\eta) g(\delta_{H-L})}{\eta g(\delta_{H-L}) + (1-\eta) g(0)} - \frac{x_H C'(e_L^*)}{x_L C'(e_H^*)} = 0. \tag{H.4}$$

By $e_H^* = e_L^* + \delta_{H-L}$, the left hand side (LHS) of (H.4) is

$$\text{LHS of (H.4)} = \frac{\eta g(0) + (1-\eta) g(\delta_{H-L})}{\eta g(\delta_{H-L}) + (1-\eta) g(0)} - \frac{x_H C'(e_L^*)}{x_L C'(e_L^* + \delta_{H-L})}. \tag{H.5}$$

If $\delta_{H-L} = 0$, (H.5)$= 1 - \frac{x_H}{x_L} \leq 0$. If $\delta_{H-L} \to \infty$, then (H.5)$\to \frac{\eta}{1-\eta} > 0$ because $\lim_{\delta_{H-L} \to \infty} g(\delta_{H-L}) = 0$. Since (H.5) is continuous in δ_{H-L}, there exists an intermediate point $\delta_{H-L} \geq 0$ such that (H.4) is satisfied. As a result, there exists an equilibrium such that $e_H^* \geq e_L^*$.

If $\eta = 1/2$, then the LHSs of (H.1) and (H.2) are the same. With the exponential cost function, we have $\frac{x_H}{x_L} = \frac{\exp(\rho e_H^*)}{\exp(\rho e_L^*)}$. Then, $e_H^* - e_L^* = \ln(x_H/x_L)/\rho$. By (H.1) and (H.2),

$$A^l[g(0)/2 + g(\ln(x_H/x_L)/\rho)/2] = C'(e_L^{l*})/x_L = \rho \exp(\rho e_L^{l*})/x_L,$$

$$A^l[g(\ln(x_H/x_L)/\rho)/2+g(0)/2] = C'(e_H^{l*})/x_H = \rho\exp(\rho e_H^{l*})/x_H.$$

Therefore, the closed form of the equilibrium effort levels are $e_L^{l*} = \ln(A^l K_0 x_L/\rho)\rho$ and $e_H^{l*} = \ln(A^l K_0 x_H/\rho)\rho$, where $K_0 = g(0)/2 + g(\ln(x_H/x_L)/\rho)/2$. □

Lemma H.1 shows that there exists an equilibrium such that the H-type contestants exert more efforts than the L-type contestants. Since their equilibrium performances are $V_H^{l*} = e_H^{l*} + \xi^l$ and $V_L^{l*} = e_L^{l*} + \xi^l$, $l = 1, 2$, then $V_H^{l*} \geq_{\text{st}} V_L^{l*}$ because $e_H^{l*} \geq e_L^{l*}$. That is, the H-type contestants are more likely to have a higher performance than the L-type contestants. Lemma H.1 also shows that there exists a unique equilibrium if $\eta = 1/2$ and the closed form of the equilibrium effort level can be obtained. For the mathematical tractability, we assume $\eta = 1/2$ in the following analysis.

H.2. Correlated Expertise

We examine two cases that the expertise of a contestants is completely *positively* or *negatively* correlated.

Positive correlation. The contestant who is endowed with high expertise x_H (resp., low expertise x_L) in the first attribute will have high expertise x_H (resp., low expertise x_L) in the second attribute.

Negative correlation. The contestant who is endowed with high expertise x_H (resp., low expertise x_L) in the first attribute will have low expertise x_L (resp., high expertise x_H) in the second attribute.

Since there exists a correlation between the expertise of the two attributes, if the firm discloses the performance of contestants after the first sub-contest, then contestants can infer the type of their opponents and the belief of the prior expertise distribution is updated at the beginning of the second sub-contest. Because the performance of a contestant in the first sub-contest consists of the effort and the random factor, the way how contestants infer their opponent's type depends on what kind of signal they can obtain from the first sub-contest, e.g., the effort, or the performance (i.e., the effort plus the random factor). If contestants can learn the performances in the first sub-contest, they may not be able to accurately infer the type of their opponents due to the random factor, which makes the characterization of contestants' behavior extremely complicated. Thus, we focus on the case that contestants can learn his opponent's effort level by the performance in the first sub-contest. This is indeed the case when the randomness comes from the preferences or the private tastes of judges, while the quality of the solution depends on the effort level. In the symmetric equilibrium, contestant of different types have different effort levels, therefore with the disclosed information, contestants can accurately learn their opponents' type.

If contestants are strategic, they may try to hide their type in the first sub-contest, because their truthful revelation may put them into a disadvantageous position in the second sub-contest. The only way for contestants to hide their types is that in the first sub-contest, both contestants have the same effort level regardless of their true types. This is analogous to Hausch (1986), in which the only way for a type of players to hide is that they make the same bid as other types of players and no one has an incentive to deviate. With their identities successfully hidden in the first sub-contest, in the second sub-contest, neither contestants' types are revealed, and thus their performances in the second sub-contest can be characterized by Lemma H.1. For the separate contest, we have the following results.

LEMMA H.2. *Assume* $\eta = 1/2$.

(i) For positively or negatively correlated expertise, contestants truthfully behave in each sub-contest of the separate contest.

(ii) In the second sub-contest of the separate contest, if both contestants are of the same type, then the equilibrium effort is $e_t^{2*} = \ln(x_t A^2 g(0)/\rho)/\rho$, $t = H$ *or* L. *If one contestant is H-type and the other is L-type, then the equilibrium efforts are* $e_H^{2*} = \ln(x_H A^2 g(\ln(x_H/x_L)/\rho)/\rho)/\rho$ *and* $e_L^{2*} = \ln(x_L A^2 g(\ln(x_L/x_H)/\rho)/\rho)/\rho$.

Proof of Lemma H.2. (i) First, we derive the equilibrium effort levels in different sub-contests of the separate contest at first when contestants truthfully behave in the first sub-contest. In the first sub-contest, since contestants have no idea about their opponents' type, thus the equilibrium effort levels are the same as what is characterized by Lemma H.1. Therefore, if $\eta = 1/2$, the equilibrium effort levels for the H-type and L-type contestants are

$$e_L^{1*} = \ln\left(x_L A^1[g(0)/2 + g(\ln(x_L/x_H)/\rho)/2]/\rho\right)/\rho,$$
$$e_H^{1*} = \ln\left(x_H A^1[g(0)/2 + g(\ln(x_L/x_H)/\rho)/2]/\rho\right)/\rho.$$

In the second sub-contest, both contestants will know his opponent's type. With probability 1/2, both contestants are either H-type or L-type, and thus the model turns to be the contest with homogeneous contestants. By Lemma 1, the equilibrium effort levels are then $e_L^{2*} = \ln(x_L A^2 g(0)/\rho)/\rho$ and $e_H^{2*} = \ln(x_H A^2 g(0)/\rho)/\rho$.

With probability 1/2, one contestant is high-type and the other is low-type, and the expected payoff to contestant i is $\mathsf{E}(u_i^2(e_i^2|x_i)) = A^2 G(e_i^2 - e_H^{2*}) - C(e_i^2)/x_i$ if $i = L$ or $\mathsf{E}(u_i^2(e_i^2|x_i)) = A^2 G(e_i^2 - e_L^{2*}) - C(e_i^2)/x_i$ if $i = H$, which leads to the FOC for the H-type and L-type contestants: $A^2 g(e_L^{2*} - e_H^{2*}) - C'(e_L^{2*})/x_L = 0$ and $A^2 g(e_H^{2*} - e_L^{2*}) - C'(e_H^{2*})/x_H = 0$. Since $g(\cdot)$ is symmetric at 0, $C'(e_L^{2*})/x_L = C'(e_H^{2*})/x_H$, equivalently, $e_H^{2*} - e_L^{2*} = \ln(x_H/x_L)/\rho$. Solving the FOC yields the equilibrium effort levels in the second sub-contest, $e_L^{2*} = \ln(x_L A^2 g(\ln(x_L/x_H)/\rho)/\rho)/\rho$ and $e_H^{2*} = \ln(x_H A^2 g(\ln(x_H/x_L)/\rho)/\rho)/\rho$.

Second, we compute the expected utilities of contestants in the first and second sub-contests if they truthfully behave in the separate contest. By the equilibrium effort levels derived above, in the first sub-contest, the chance for an H-type or an L-type contestant to compete with an H-type or L-type contestant is 1/2. Therefore, the ex ante expected utilities for H-type and L-type contestants are

$$\mathsf{E}(u_L^1(e_L^{1*}|x_L)) = \frac{A^1}{2}[G(0)+G(\ln(x_L/x_H)/\rho)] - \frac{A^1}{2\rho}[g(0)+g(\ln(x_L/x_H)/\rho)], \tag{H.6}$$
$$\mathsf{E}(u_H^1(e_H^{1*}|x_H)) = \frac{A^1}{2}[G(0)+G(\ln(x_H/x_L)/\rho)] - \frac{A^1}{2\rho}[g(0)+g(\ln(x_H/x_L)/\rho)]. \tag{H.7}$$

In the second sub-contest, if there is an H-type contestant and another is an L-type contestant, then the expected utilities for H-type and L-type contestants are

$$\mathsf{E}(u_L^2(e_L^{2*}|x_L)) = A^2G(\ln(x_L/x_H)/\rho) - \frac{A^2}{\rho}g(\ln(x_H/x_L)/\rho),$$
$$\mathsf{E}(u_H^2(e_H^{2*}|x_H)) = A^2G(\ln(x_H/x_L)/\rho) - \frac{A^2}{\rho}g(\ln(x_H/x_L)/\rho).$$

Note that the expected utility of an L-type contestant can be negative in the second sub-contest since $x_L < x_H$. That is, if we allow the entry decision to be reconsidered, it is possible that an L-type contestant self-interestedly drops out of the second sub-contest when she finds that her opponent is an H-type.

If both contestants are L-type or H-type contestants, the expected utility is $\mathsf{E}(u_H^2(e_H^{2*}|x_H)) = \mathsf{E}(u_L^2(e_L^{2*}|x_L)) = A^2G(0) - \frac{A^2}{\rho}g(0)$. Therefore, the ex ante expected utilities of contestants in the second sub-contest are

$$\mathsf{E}(u_L^2(e_L^{2*}|x_L)) = \frac{1}{2}\left[A^2G(\ln(x_L/x_H)/\rho) - \frac{A^2}{\rho}g(\ln(x_H/x_L)/\rho)\right] + \frac{1}{2}\left[A^2G(0) - \frac{A^2}{\rho}g(0)\right], \tag{H.8}$$
$$\mathsf{E}(u_H^2(e_H^{2*}|x_H)) = \frac{1}{2}\left[A^2G(\ln(x_H/x_L)/\rho) - \frac{A^2}{\rho}g(\ln(x_H/x_L)/\rho)\right] + \frac{1}{2}\left[A^2G(0) - \frac{A^2}{\rho}g(0)\right]. \tag{H.9}$$

Third, we show that the utilities of contestants are the same in the second sub-contest, no matter whether they hide or do not hide their types in the first sub-contest. If contestants hide their types in the first sub-contest, then the second sub-contest is equivalent to the first sub-contest in which contestants truthfully behave. As a result, by (H.6) and (H.7), we have the ex ante expected utilities in the second sub-contest if contestants hide their types as

$$\mathsf{E}(u_L^2(e_L^{2*}|x_L)) = \frac{A^2}{2}[G(0)+G(\ln(x_L/x_H)/\rho)] - \frac{A^2}{2\rho}[g(0)+g(\ln(x_L/x_H)/\rho)],$$
$$\mathsf{E}(u_H^2(e_H^{2*}|x_H)) = \frac{A^2}{2}[G(0)+G(\ln(x_H/x_L)/\rho)] - \frac{A^2}{2\rho}[g(0)+g(\ln(x_H/x_L)/\rho)],$$

which are the same as (H.8) and (H.9). In conclusion, no matter whether contestants hide their types or not, the ex ante expected utilities stay the same in the second sub-contest. By Lemma H.1, if $\eta = 1/2$, the equilibrium that contestants with different types make different efforts is unique. Since hiding one's type in the first sub-contest is not beneficial for contestants, they will deviate to the unique equilibrium in which both contestants truthfully behave in the first sub-contest.

The same result holds for the case that the expertise of the contestant along the two attributes is negatively correlated.

(ii) The equilibrium effort in the second sub-contest is derived in the first part of proof in (i). □

For Lemma H.2(i), we find that if $\eta = 1/2$, contestants have no incentive to hide their types in the first sub-contest. Since contestants decide on whether to hide their types at the beginning of the first sub-contest, they need to take into account their ex ante utility in the second sub-contest. We find that no matter whether contestants hide or do not hide their types in the first sub-contest, the ex ante utilities in the second sub-contest remain the same if $\eta = 1/2$. As a result, hiding one's type is not beneficial for contestants. Moreover, by Lemma H.1, if $\eta = 1/2$, in the first sub-contest, the equilibrium is unique such that contestants with different types make different efforts. Therefore, if contestants hide their types in the first sub-contest, they will deviate because the utilities in the second sub-contest is the same no matter they hide or not. Thus, contestants' behavior in the first sub-contest is characterized by Lemma H.1, i.e., they truthfully behave without any strategic behavior. In the second sub-contest, contestants know their opponent's type, the equilibrium efforts are characterized by Lemma H.2(b).

The expected utility of an L-type contestant can be negative in the second sub-contest of the separate contest since $x_L < x_H$. That is, if we allow the entry decision to be reconsidered, it is possible that an L-type contestant self-interestedly drops out of the second sub-contest when she finds that her opponent is an H-type. When this happens, the slack-off effect in the second sub-contest will be stronger than when the L-type contestant is forced to participate, since the H-type contestant is the sole remaining participant. As a result, if we allow a contestant to reconsider her participation in the second sub-contest of the separate contest, the joint contest, which is not affected by this relaxation, is more likely to be favored than when all contestants are forced to participate.

H.3. Joint Contest

With the model setup of the positively correlated and negative correlated expertise across the two attributes, we can derive the contestant behavior in the joint contest. Denote the difference of random factors between contestants i and j along the two dimensions by $\gamma^\circ = \xi_i^1 + \xi_i^2 - \xi_j^1 - \xi_j^2 =$

$(\xi_i^1-\xi_j^1)+(\xi_i^2-\xi_j^2)=\gamma^1+\gamma^2$. By the symmetric property of γ^1 and γ^2, the random variable γ° has a symmetric PDF $g^\circ(\gamma^\circ)$ and a CDF $G^\circ(\gamma^\circ)$. We assume $g^\circ(0)>g(0)/2$, equivalently, $h^\circ(\xi^\circ;2)>h(\xi;2)/2$, which is naturally satisfied by normal distributions.

LEMMA H.3. *(i) If contestants' expertise along the two attributes is positively correlated, the equilibrium efforts for the L-type and H-type contestants are* $e_L^{\circ *}=2\ln(AK_1x_L/\rho)/\rho$ *and* $e_H^{\circ *}=2\ln(AK_1x_H/\rho)/\rho$ *respectively, where* $K_1=g^\circ(2\ln(x_H/x_L)/\rho)/2+g^\circ(0)/2$.

(ii) If contestants' expertise along the two attributes is negatively correlated, then equilibrium effort is $e^{\circ *}=2\ln(Ag^\circ(0)\sqrt{x_Hx_L}/\rho)\rho$ *for both contestants.*

Proof of Lemma H.3. (i) First, we derive the expressions of $C_i^{\circ\prime}(e_i^\circ)$. If the expertise along the two attributes is positively correlated, then a contestant is either H-type or L-type in both attributes. Therefore, the cost function is $C^\circ(e_i^\circ)=\min\{C(\tilde{e}_i^1)/x_i+C(\tilde{e}_i^2)/x_i\}$, $i=L,H$. By Lemma 2, the optimal allocation of efforts satisfies $C'(\tilde{e}_i^1)/x_i=C'(\tilde{e}_i^2)/x_i$, $i=L,H$. Then, $C'(\tilde{e}_i^1)/C'(\tilde{e}_i^2)=1$, equivalently, $\tilde{e}_i^1=\tilde{e}_i^2$. When $C'(e_i)=\rho\exp(\rho e_i)$, since the total cost is $C_i^\circ(e_i^\circ)=C(\tilde{e}_i^1)/x_i+C(\tilde{e}_i^2)/x_i$, the derivative of the total cost function is

$$C_i^{\circ\prime}(e_i^\circ)=\frac{\rho}{2}\exp\left(\frac{\rho e_i^\circ}{2}\right)/x_i+\frac{\rho}{2}\exp\left(\frac{\rho e_i^\circ}{2}\right)/x_i=\rho\exp\left(\frac{\rho e_i^\circ}{2}\right)/x_i, \tag{H.10}$$

$i=L,H$. Similar to the proof of Lemma H.1, if contestant j with type H or L makes effort $e_H^{\circ *}$ or $e_L^{\circ *}$ in equilibrium, respectively, the winning probability of contestant i is $G^\circ(e_i^\circ-e_H^{\circ *})/2+G^\circ(e_i^\circ-e_L^{\circ *})/2$. The expected payoff to contestant i is $\mathsf{E}(u^\circ(e_i^\circ|x_i))=A[G^\circ(e_i^\circ-e_H^{\circ *})+G^\circ(e_i^\circ-e_L^{\circ *})]/2-C^\circ(e_i^\circ)$. In the symmetric equilibrium, contestant i makes effort $e_L^{\circ *}$ if he is L-type and $e_H^{\circ *}$ if he is H-type, which lead to

$$A[g^\circ(0)+g^\circ(e_L^{\circ *}-e_H^{\circ *})]/2=C_L^{\circ\prime}(e_L^{\circ *}), \tag{H.11}$$

$$A[g^\circ(e_H^{\circ *}-e_L^{\circ *})+g^\circ(0)]/2=C_H^{\circ\prime}(e_H^{\circ *}). \tag{H.12}$$

Since $g^\circ(\cdot)$ is symmetric at 0, the LHSs of (H.11) and (H.12) are the same. With the exponential cost function, we have $\frac{x_H}{x_L}=\frac{\exp(\rho e_H^{\circ *}/2)}{\exp(\rho e_L^{\circ *}/2)}$, Then, $e_H^{\circ *}-e_L^{\circ *}=2\ln(x_H/x_L)/\rho$. As a result, by (H.11) and (H.12), we have

$$A[g^\circ(0)+g^\circ(2\ln(x_H/x_L)/\rho)]/2=\rho\exp\left(\rho e_L^{\circ *}/2\right)/x_L, \tag{H.13}$$

$$A[g^\circ(2\ln(x_H/x_L)/\rho)+g^\circ(0)]/2=\rho\exp\left(\rho e_H^{\circ *}/2\right)/x_H. \tag{H.14}$$

By (H.13) and (H.14), the equilibrium efforts for the H-type and L-type contestants are $e_L^{\circ *}=2\ln(AK_1x_L/\rho)/\rho$ and $e_H^{\circ *}=2\ln(AK_1x_H/\rho)/\rho$, where $K_1=[g^\circ(2\ln(x_H/x_L)/\rho)+g^\circ(0)]/2$.

(ii) Now we derive the expression of $C_i^{\circ\prime}(e_i^\circ)$ if the expertise along the two attributes is negatively correlated. Denote the type of contestant by (i^1, i^2), $i^1, i^2 = H, L$, where i^1 is the expertise in the first attribute, and i^2 is the expertise in the second attribute. If the expertise along the two attributes is negatively correlated, a contestant can either be type (H, L) or (L, H). Since the cost functions for those two types are the same, we derive the expression of the $C_i^{\circ\prime}(e_i^\circ)$ for type (L, H), which is the same for type (H, L). If contestant i has expertise x_L in the first attribute and x_H in the second attribute, given the aggregate effort e_i°, there exists an optimal allocation of efforts $e_i^\circ = \tilde{e}_i^1 + \tilde{e}_i^2$ such that $C^\circ(e_i^\circ) = \min\{C(\tilde{e}_i^1)/x_L + C(\tilde{e}_i^2)/x_H\}$. By Lemma 2, the optimal allocation of efforts satisfies $C'(\tilde{e}_i^1)/x_L = C'(\tilde{e}_i^2)/x_H$. Then,

$$C'(\tilde{e}_i^1)/C'(\tilde{e}_i^2) = x_L/x_H. \tag{H.15}$$

When $C'(e_i) = \rho\exp(\rho e_i)$, (I.1) becomes $\exp(\rho(\tilde{e}_i^1 - \tilde{e}_i^2)) = x_L/x_H$, equivalently $\tilde{e}_i^1 - \tilde{e}_i^2 = \ln(x_L/x_H)/\rho$. Since $e_i^\circ = \tilde{e}_i^1 + \tilde{e}_i^2$, we have $\tilde{e}_i^1 = [e_i^\circ + \ln(x_L/x_H)/\rho]/2$ and $\tilde{e}_i^2 = [e_i^\circ - \ln(x_L/x_H)/\rho]/2$. Since the total cost is $C^\circ(e_i^\circ) = C(\tilde{e}_i^1)/x_L + C(\tilde{e}_i^2)/x_H$, the derivative of the total cost function is

$$\begin{aligned} C^{\circ\prime}(e_i^\circ) &= \frac{\rho}{2}\exp\left(\frac{\rho e_i^\circ + \ln(x_L/x_H)}{2}\right)/x_L + \frac{\rho}{2}\exp\left(\frac{\rho e_i^\circ - \ln(x_L/x_H)}{2}\right)/x_H \\ &= \rho\exp(\rho e_i^\circ/2)/\sqrt{x_H x_L}. \end{aligned} \tag{H.16}$$

Following the proof in Lemma 3, we obtain the symmetric equilibrium effort $e^{\circ*} = 2\ln(Ag^\circ(0)\sqrt{x_H x_L}/\rho)\rho$. □

Lemma H.3(i) is similar to Lemma H.1. It shows that the equilibrium effort in the joint contest is similar to that in the first sub-contest of the separate contest, if the contestants' expertise along the two attributes is positively correlated. The only difference is that in the joint contest, contestants make effort in both attributes at once, and the H-type contestant is better than the L-type contestant in both attributes, due to the positively correlated expertise. Since the two cost functions across the two attributes of each type of a contestant are the same, the effort levels of a contestant in the two attributes are the same. Moreover, Lemma H.3(ii) is similar to Lemma 3. Since the expertise across the two attributes is negative correlated and the expertise levels x_H and x_L are the same for both attributes, then in the joint contest contestants make the same effort since their aggregate cost functions are the same. With Lemmas H.2 and H.3, we can compare the equilibrium efforts between the two contest mechanisms.

Proof of Proposition 5. (i) First, we compare the equilibrium effort levels of the H-type contestant between two contest mechanisms.

By Lemma H.3(i), in the joint contest, the equilibrium effort of the H-type contestant is $e_H^{\circ *} = 2\ln(AK_1x_H/\rho)/\rho$, where $K_1 = [g^\circ(2\ln(x_H/x_L)/\rho) + g^\circ(0)]/2$. By Lemma H.1, the effort of the H-type contestant in the first sub-contest is $e_H^{1*} = \ln(A^1K_0x_H/\rho)\rho$, where $K_0 = [g(\ln(x_H/x_L)/\rho) + g(0)]/2$. By Lemma H.2, the equilibrium effort level of the H-type contestant in the second sub-contest is $e_H^{2*} = \ln(A^2g(\ln(x_H/x_L)/\rho)x_H/\rho)/\rho$ if his opponent is L-type. By Lemma 1, the equilibrium effort level of the H-type contestant is $e_H^{2*} = \ln(A^2g(0)x_H/\rho)/\rho$ if his opponent is H-type. The probability that an H-type contestant competes with an H-type or L-type contestant is 1/2. Thus, the difference of the expected equilibrium effort levels between two contest mechanisms is

$$\begin{aligned}\mathsf{E}(e_H^{\circ *}) - \mathsf{E}(e_H^{1*} + e_H^{2*}) &= 2\ln(AK_1x_H/\rho)/\rho - [\ln(A^1K_0x_H/\rho)/\rho + \ln(A^2g(0)x_H/\rho)/\rho]/2 \\ &\quad -[\ln(A^1K_0x_H/\rho)/\rho + \ln(A^2g(\ln(x_H/x_L)/\rho)x_H/\rho)/\rho]/2 \\ &= \frac{1}{2\rho}\left[\ln\left(\frac{4K_1^2}{K_0g(0)}\frac{4K_1^2}{K_0g(\ln(x_H/x_L)/\rho)}\right)\right] = \frac{1}{2\rho}\left[\ln\left(\frac{16K_1^4}{K_0^2g(0)g(\ln(x_H/x_L)/\rho)}\right)\right],\end{aligned}$$

where the second equality is due to $A/2 = A^1 = A^2$. Now we examine the value of $\frac{16K_1^4}{K_0^2g(0)g(\ln(x_H/x_L)/\rho)}$. Since $g^\circ(x) = \frac{1}{2\sqrt{2\pi}\sigma}e^{-\frac{x^2}{8\sigma^2}}$ and $g(x) = \frac{1}{2\sqrt{\pi}\sigma}e^{-\frac{x^2}{4\sigma^2}}$, by denoting $\delta_x = \ln(x_H/x_L)/\rho$, $\delta_x \in (0, +\infty)$, we have the following result,

$$\frac{16K_1^4}{K_0^2g(0)g(\ln(x_H/x_L)/\rho)} = \frac{4(g^\circ(2\delta_x) + g^\circ(0))^4}{(g(\delta_x) + g(0))^2g(0)g(\delta_x)} = \frac{\left[\exp\left(-\frac{\delta_x^2}{2\sigma^2}\right) + 1\right]^4}{\left[\exp\left(-\frac{\delta_x^2}{4\sigma^2}\right) + 1\right]\left[\exp\left(-\frac{\delta_x^2}{2\sigma^2}\right) + \exp\left(-\frac{\delta_x^2}{4\sigma^2}\right)\right]}.$$

Since $\exp\left(-\frac{\delta_x^2}{2\sigma^2}\right) = \left[\exp\left(-\frac{\delta_x^2}{4\sigma^2}\right)\right]^2$, we can denote $z = \exp\left(-\frac{\delta_x^2}{4\sigma^2}\right) \in [0, 1]$. Note that if $z = 0$, $x_H/x_L \to +\infty$, and if $z = 1$, $x_H/x_L = 1$. We have that $\frac{16K_1^4}{K_0^2g(0)g(\ln(x_H/x_L)/\rho)} = \frac{(z^2+1)^4}{(z+1)(z^2+z)} \geq \frac{(z^2+1)^4}{(z+1)(z^2+1)} \geq \frac{(z^2+1)^3(z+1)}{(z+1)(z^2+1)} = (z^2+1)^2 \geq 1$. As a result, if the expertise along the two attributes is completely positively correlated, the H-type contestant makes a higher expected equilibirium effort in the joint contest than in the separate contest.

Second, we compare the equilibrium effort levels of the L-type contestant between two contest mechanisms. By Lemma H.3(i), in the joint contest, the equilibrium effort of the L-type contestant is $e_L^{\circ *} = 2\ln(AK_1x_L/\rho)/\rho$, where $K_1 = [g^\circ(2\ln(x_H/x_L)/\rho) + g^\circ(0)]/2$. By Lemma H.1, the effort of the L-type contestant in the first sub-contest is $e_L^{1*} = \ln(A^1K_0x_L/\rho)\rho$, where $K_0 = [g(\ln(x_H/x_L)/\rho) + g(0)]/2$. By Lemma H.2, the equilibrium effort level of the L-type contestant in the second sub-contest is $e_L^{2*} = \ln(A^2g(\ln(x_H/x_L)/\rho)x_L/\rho)/\rho$ if his opponent is H-type. By Lemma 1, the equilibrium effort level of the H-type contestant is $e_L^{2*} = \ln(A^2g(0)x_L/\rho)/\rho$ if his opponent is L-type. The probability that an H-type contestant competes with an H-type or L-type contestant is

1/2. Thus, the difference of the expected equilibrium effort levels between two contest mechanisms is

$$\begin{aligned}\mathsf{E}(e_L^{\circ*}) - \mathsf{E}(e_L^{1*} + e_L^{2*}) &= 2\ln(AK_1x_L/\rho)/\rho - [\ln(A^1K_0x_L/\rho)/\rho + \ln(A^2g(0)x_H/\rho)/\rho]/2 \\ &\quad -[\ln(A^1K_0x_L/\rho)/\rho + \ln(A^2g(\ln(x_H/x_L)/\rho)x_L/\rho)/\rho]/2 \\ &= \frac{1}{2\rho}\left[\ln\left(\frac{4K_1^2}{K_0g(0)}\frac{4K_1^2}{K_0g(\ln(x_H/x_L)/\rho)}\right)\right] = \frac{1}{2\rho}\left[\ln\left(\frac{16K_1^4}{K_0^2g(0)g(\ln(x_H/x_L)/\rho)}\right)\right],\end{aligned}$$

where the second equality is due to $A/2 = A^1 = A^2$. Since the value of $\mathsf{E}(e_L^{\circ*}) - \mathsf{E}(e_L^{1*} + e_L^{2*})$ is the same with that of $\mathsf{E}(e_H^{\circ*}) - \mathsf{E}(e_H^{1*} + e_H^{2*})$. The same result holds for the L-type contestant.

(ii) If the expertise along the two attributes is negatively correlated, then by Lemma H.3(ii), the effort level will be the same for either type of contestants, i.e., $e^{\circ*} = 2\ln(Ag(0)\sqrt{x_Hx_L}/\rho)\rho$. Consider the contestant with type (H, L). By Lemma H.1, the equilibrium effort level of the contestant in first sub-contest is $e_H^{1*} = \ln(A^1K_0x_H/\rho)\rho$. By Lemma H.2, the equilibrium effort level of the contestant in the second sub-contest is $e_L^{2*} = \ln(A^2g(\ln(x_H/x_L)/\rho)x_L/\rho)/\rho$ if his opponent is type (L, H). By Lemma 1, the equilibrium effort level of the H-type contestant is $e_L^{2*} = \ln(A^2g(0)x_L/\rho)/\rho$ if his opponent is type (H, L). The probability that the contestant with type (H, L) competes with an contestant with type (H, L) or (L, H) is 1/2. Thus, the difference of the expected equilibrium effort levels between two contest mechanisms is

$$\begin{aligned}\mathsf{E}(e^{\circ*}) - \mathsf{E}(e_H^{1*} + e_L^{2*}) &= 2\ln(Ag^\circ(0)\sqrt{x_Hx_L}/\rho)/\rho - [\ln(A^1K_0x_H/\rho)/\rho + \ln(A^2g(0)x_L/\rho)/\rho]/2 \\ &\quad -[\ln(A^1K_0x_H/\rho)/\rho + \ln(A^2g(\ln(x_H/x_L)/\rho)x_L/\rho)/\rho]/2 \\ &= \frac{1}{2\rho}\ln\left(\frac{16g^\circ(0)^4}{K_0^2g(0)g(\ln(x_H/x_L)/\rho)}\right) \geq \frac{1}{2\rho}\ln\left(\frac{16g^\circ(0)^4}{g^4(0)}\right) > 0,\end{aligned}$$

where the first inequality is due to $g(\ln(x_H/x_L)/\rho) \leq g(0)$ and the second inequality is due to $g^\circ(0) > g(0)/2$. Therefore, we have $\mathsf{E}(e^{\circ*}) - \mathsf{E}(e_H^{1*} + e_L^{2*}) > 0$. For the negatively correlated expertise, for the contestant with type (H, L), the effort level is higher in the joint contest than in the separate contest. The same result holds for a contestant with type (L, H) due to the symmetry. □

Proposition 5 shows that if the expertise along the two attributes is perfectly positively correlated, then each contestant has a higher expected effort level in the joint contest than in the separate contest. Since $\eta = 1/2$, with probability 1/2, one contestant is H-type and the other contestant is L-type. In the second sub-contest of the separate contest, contestants know their opponents' types. We find that if one is H-type and the other is L-type, then they make relatively little effort because the marginal winning probability for additional effort is low. For an L-type contestant, if he knows that his opponent is H-type, then the winning probability is slim, and thus he will slack off because making great effort incurs a high cost but gains little improvement in the chance of winning. For an H-type contestant, if he knows that his opponent is L-type who will make little effort, there is no need for him to make great effort as well. In summary, the *information revelation* leads to the low effort level for both types of contestants. With probability $1/2$, both contestants will make little effort in equilibrium in the second sub-contest.

In the first sub-contest, there is no information revelation, and thus either contestant does not know exactly what type his opponent is, but each contestant takes into account the probability (which is $1/2$) that his opponent is of a different type from him. Therefore, they make little effort, an effort level lower than if the opponent is definitely of the same type, but higher than the effort level if the opponent is definitely of a different type. With the analysis and explanation above, we find that the *heterogeneity* in expertise leads to a *lower* effort level than a homogeneous pool of contestants who have the expertise equal to the average expertise of the heterogeneous pool.

For the joint contest, since contestants are required to submit the aggregate solution, they take into account the possibility that their opponents are of different types. Then, the joint contest is analogous to the first sub-contest of the separate contest. However, the heterogeneity of the expertise is higher in the joint contest than the first sub-contest of the separate contest due to the positive correlation of expertise. For the L-type (H-type) contestant, the winning probability is lower (higher) when competing in both attributes at once than competing in one attribute. Thus, both contestants have a lower effort level in the joint contest than in the first sub-contest of the separate contest. In summary, though there is no information revelation in the joint contest, the relatively *high heterogeneity* in expertise also leads to relatively low effort levels for both contestants.

With the discussion above, we find that the contestants makes little effort in the separate contest mainly because of the information revelation in the second sub-contest, and the heterogeneity in the first sub-contest. Moreover, contestants make little effort in the joint contest because of the high heterogeneity. For the general case, the comparison between the separate and joint contests can be ambiguous. But, if the random factors follow the normal distribution and $\eta = 1/2$, then the effect of information revelation and the heterogeneity in the separate contest, and the effect of the high heterogeneity in the joint contest will be in a similar level. Since those effects mentioned above can be neutralized, the pooling effect becomes the dominating force, and therefore, each contestant has a higher equilibrium effort level in the joint contest than in the separate contest.

If contestants' expertise along the two attributes is negatively correlated, each contestant always has a higher equilibrium effort level in the joint contest than in the separate contest. Since contestants are strong at one attribute but weak at the other, the effect of information revelation, which leads to slack off in the second sub-contest of the separate contest, always exists. However, for the joint contest, there is no effect of the high heterogeneity since contestants are ex ante identical in the expertise. As a result, under the joint influence of the pooling effect, each contestant has a higher equilibrium effort level in the joint contest than in the separate contest.

Discussion. If the project is highly effort-based, then the contestant with a higher effort level will win with high probability. If the expertise along the two attributes is negatively correlated, the best efforts in the two sub-contests of the separate contest can be made by different contestants. To compare the expected best effort levels between the two contest mechanisms, we do following analysis.

If the expertise along the two attributes is negatively correlated, then by Lemma H.3(ii), the effort level will be the same for either type of contestants, i.e., $e^{\circ *} = 2\ln(Ag(0)\sqrt{x_H x_L}/\rho)\rho$. Consider the contestant with type (H, L). By Lemma H.1, the equilibrium effort level of the contestant in first sub-contest is $e_H^{1*} = \ln(A^1 K_0 x_H/\rho)\rho$. By Lemma H.2, the equilibrium effort level of the contestant in the second sub-contest is $e_L^{2*} = \ln(A^2 g(\ln(x_H/x_L)/\rho) x_L/\rho)/\rho$ if his opponent is type (L, H). By Lemma 1, the equilibrium effort level of the H-type contestant is $e_L^{2*} = \ln(A^2 g(0) x_L/\rho)/\rho$ if his opponent is type (H, L). The probability that the contestant with type (H, L) competes with an contestant with type (H, L) or (L, H) is 1/2. Note that if two contestants are the same type, then the expected best effort level will be either contestants' effort level in the separate contest. If two contestants are of different types, then the expected best effort level will the combination of effort levels owned by the contestant with H-type in each attribute.

Thus, if the the difference of the expected best equilibrium effort levels between two contest mechanisms is

$$
\begin{aligned}
& \mathsf{E}(e^{\circ *}) - \mathsf{E}(e_H^{1*} + e_L^{2*})/4 - \mathsf{E}(e_L^{1*} + e_H^{2*})/4 - \mathsf{E}(e_H^{1*} + e_H^{2*})/2 \\
= & \; 2\ln(Ag^{\circ}(0)\sqrt{x_H x_L}/\rho)/\rho - [\ln(A^1 K_0 x_H/\rho)/\rho + \ln(A^2 g(0) x_L/\rho)/\rho]/4 \\
& - [\ln(A^1 K_0 x_L/\rho)/\rho + \ln(A^2 g(0) x_H/\rho)/\rho]/4 \\
& - [\ln(A^1 K_0 x_H/\rho)/\rho + \ln(A^2 g(\ln(x_H/x_L)/\rho) x_H/\rho)/\rho]/2 \\
= & \; \frac{1}{2\rho} \ln\left(\frac{16 g^{\circ}(0)^4}{K_0^2 g(0) g(\ln(x_H/x_L)/\rho)} \sqrt{\frac{x_L}{x_H}} \right).
\end{aligned}
$$

Since $K_0 < g(0)$ and $g^{\circ}(0) > g(0)/2$, we have $K_0^2 g(0) g(\ln(x_H/x_L)/\rho) < g(0)^4 < 16 g^{\circ}(0)^4$. If x_L and x_H are close, then $\mathsf{E}(e^{\circ *}) - \mathsf{E}(e_H^{1*} + e_L^{2*}) > 0$. If the difference between x_L and x_H is sufficiently large, then $\mathsf{E}(e^{\circ *}) - \mathsf{E}(e_H^{1*} + e_L^{2*}) \leq 0$. $\square$

H.4. A Given Sample Path

Though contestants' expected equilibrium effort levels are higher in the joint contest than in the separate contest. The two contest mechanisms perform differently in incentivizing contestants' efforts given different realizations of the pairs of contestants, still under the assumption that ex ante the contestants only know their own types.

Proposition H.1. *(i) For the perfectly positively correlated expertise, (a) if an H-type contestant's opponent is H-type, there exists a threshold on x_H/x_L, under which $e_H^{\circ *} - e_H^{1*} - e_H^{2*} \geq 0$ and above which $e_H^{\circ *} - e_H^{1*} - e_H^{2*} < 0$; (b) if an H-type contestant's opponent is L-type, for any x_H/x_L, $e_H^{\circ *} - e_H^{1*} - e_H^{2*} \geq 0$; (c) if an L-type contestant's opponent is L-type, there exists a threshold on x_H/x_L, under which $e_L^{\circ *} - e_L^{1*} - e_L^{2*} \geq 0$ and above which $e_L^{\circ *} - e_L^{1*} - e_L^{2*} < 0$; (d) if an L-type contestant's opponent is H-type, for any x_H/x_L, $e_L^{\circ *} - e_L^{1*} - e_L^{2*} \geq 0$.*

(ii) For the perfectly negatively correlated expertise, contestants with type (H, L) or (L, H) have the higher equilibrium effort level in the joint than in the separate contest.

Proof of Proposition H.1. (i)(a) We compare the equilibrium effort levels of the H-type contestant between the two contest mechanisms if his opponent is H-type.

By Lemma H.3(i), in the joint contest, the equilibrium effort of the H-type contestant is $e_H^{\circ *} = 2\ln(AK_1 x_H/\rho)/\rho$, where $K_1 = [g^\circ(2\ln(x_H/x_L)/\rho) + g^\circ(0)]/2$. By Lemma H.1, the effort of the H-type contestant in the first sub-contest is $e_H^{1*} = \ln(A^1 K_0 x_H/\rho)\rho$, where $K_0 = [g(\ln(x_H/x_L)/\rho) + g(0)]/2$. By Lemma 1, the equilibrium effort level of the H-type contestant is $e_H^{2*} = \ln(A^2 g(0) x_H/\rho)/\rho$ if his opponent is H-type. The difference of the equilibrium effort levels between two contest mechanisms is

$$e_H^{\circ *} - e_H^{1*} - e_H^{2*} = 2\ln(AK_1 x_H/\rho)/\rho - \ln(A^1 K_0 x_H/\rho)\rho - \ln(A^2 g(0) x_H/\rho)/\rho = \ln\left(\frac{4K_1^2}{K_0 g(0)}\right)/\rho,$$

where the last equality is due to that $A/2 = A^1 = A^2$. Now we examine the value of $\frac{4K_1^2}{K_0 g(0)}$. Since $g^\circ(x) = \frac{1}{2\sqrt{2\pi}\sigma} e^{-\frac{x^2}{8\sigma^2}}$ and $g(x) = \frac{1}{2\sqrt{\pi}\sigma} e^{-\frac{x^2}{4\sigma^2}}$, by denoting $\delta_x = \ln(x_H/x_L)/\rho$, $\delta_x \in (0, +\infty)$, we have the following result,

$$\frac{4K_1^2}{K_0 g(0)} = \frac{4(g^\circ(2\delta_x) + g^\circ(0))^2}{2(g(\delta_x) + g(0))g(0)} = \frac{\left[\exp\left(-\frac{\delta_x^2}{2\sigma^2}\right) + 1\right]^2}{\left[\exp\left(-\frac{\delta_x^2}{4\sigma^2}\right) + 1\right]}.$$

Since $\exp\left(-\frac{\delta_x^2}{2\sigma^2}\right) = \left[\exp\left(-\frac{\delta_x^2}{4\sigma^2}\right)\right]^2$, by denoting $z = \exp\left(-\frac{\delta_x^2}{4\sigma^2}\right) \in [0, 1]$, we have $\frac{4K_1^2}{K_0 g(0)} = \frac{(z^2+1)^2}{z+1}$. Note that if $z = 0$, $x_H/x_L \to +\infty$, and if $z = 1$, $x_H/x_L = 1$. Taking derivative with respect to z, we find that $\left[\frac{(z^2+1)^2}{z+1}\right]' = \frac{4(z^2+1)z(z+1) - (z^2+1)^2}{(z+1)^2} = \frac{(z^2+1)(3z^2+4z-1)}{(z+1)^2}$. Therefore, if $z \in [0, \frac{\sqrt{7}-2}{3})$, then $3z^2 +$

$4z-1<0$, i.e., $\left[\frac{(z^2+1)^2}{z+1}\right]'<0$ and $\frac{4K_1^2}{K_0g(0)}$ is striclty decreasing. If $z\in[\frac{\sqrt{7}-2}{3},1]$, then $3z^2+4z-1\geq 0$, i.e., $\left[\frac{(z^2+1)^2}{z+1}\right]'\geq 0$ and $\frac{4K_1^2}{K_0g(0)}$ is increasing.

If $z=0$, i.e., $x_H/x_L\to+\infty$, then $\lim\limits_{x_H/x_L\to+\infty}\frac{4K_1^2}{K_0g(0)}=1$ and thus for $z\in(0,\frac{\sqrt{7}-2}{3})$, $\frac{4K_1^2}{K_0g(0)}<1$. Moreover, if $z=1$, i.e., $x_H=x_L$, $\frac{4K_1^2}{K_0g(0)}=2$ and the value of $\frac{4K_1^2}{K_0g(0)}$ is strictly increasing in z from a value less than 1 to 2. As a summary, when x_H/x_L ranges from 1 to $+\infty$, there exists a threshold below which $\frac{4K_1^2}{K_0g(0)}\geq 1$, and above which $\frac{4K_1^2}{K_0g(0)}<1$. As a result, there exists a threshold on x_H/x_L above which $e_H^{\circ *}-e_H^{1*}-e_H^{2*}<0$ and under which $e_H^{\circ *}-e_H^{1*}-e_H^{2*}\geq 0$.

(i)(b) By Lemma H.2, the equilibrium effort level of the H-type contestant in the second sub-contest is $e_H^{2*}=\ln(A^2g(\ln(x_H/x_L)/\rho)x_H/\rho)/\rho$ if his opponent is L-type. The difference of the equilibrium efforts between two contest mechanisms is

$$\begin{aligned}e_H^{\circ *}-e_H^{1*}-e_H^{2*} &= 2\ln(AK_1x_H/\rho)/\rho-\ln(A^1K_0x_H/\rho)\rho-\ln(A^2g(\ln(x_H/x_L)/\rho)x_H/\rho)/\rho\\ &=\ln\left(\frac{4K_1^2}{K_0g(\ln(x_H/x_L)/\rho)}\right)/\rho,\end{aligned}$$

where the last equality is due to that $A/2=A^1=A^2$. Now we examine the value of $\frac{4K_1^2}{K_0g(\ln(x_H/x_L)/\rho)}$. With the notation $\delta_x=\ln(x_H/x_L)/\rho$, $\delta_x\in(0,+\infty)$, we can obtain

$$\frac{4K_1^2}{K_0g(\ln(x_H/x_L)/\rho)}=\frac{4(g^\circ(2\delta_x)+g^\circ(0))^2}{2(g(\delta_x)+g(0))g(\delta_x)}=\frac{\left[\exp\left(-\frac{\delta_x^2}{2\sigma^2}\right)+1\right]^2}{\left[\exp\left(-\frac{\delta_x^2}{2\sigma^2}\right)+\exp\left(-\frac{\delta_x^2}{4\sigma^2}\right)\right]}.$$

By denoting $z=\exp\left(-\frac{\delta_x^2}{4\sigma^2}\right)\in(0,1)$, we have $\frac{4K_1^2}{K_0g(\ln(x_H/x_L)/\rho)}=\frac{(z^2+1)^2}{(z^2+z)}>\frac{(z^2+1)^2}{(z^2+1)}=z^2+1\geq 1$. Therefore, $e_H^{\circ *}-e_H^{1*}-e_H^{2*}>0$ for any x_H/x_L.

(i)(c) By Lemma H.3(i), in the joint contest, the equilibrium effort of the L-type contestant is $e_L^{\circ *}=2\ln(AK_1x_L/\rho)/\rho$, where $K_1=[g^\circ(2\ln(x_H/x_L)/\rho)+g^\circ(0)]/2$. By Lemma H.1, the effort of the L-type contestant in the first sub-contest is $e_L^{1*}=\ln(A^1K_0x_L/\rho)\rho$, where $K_0=[g(\ln(x_H/x_L)/\rho)+g(0)]/2$. By Lemma H.1, the equilibrium effort level in the second sub-contest is $e_L^{2*}=\ln(A^2g(0)x_L/\rho)/\rho$ if his opponent is L-type, and the difference of the equilibrium effort levels between two contest mechanisms is

$$e_L^{\circ *}-e_L^{1*}-e_L^{2*}=2\ln(AK_1x_L/\rho)/\rho-\ln(A^1K_0x_L/\rho)/\rho-\ln(A^2g(0)x_L/\rho)/\rho=\ln\left(\frac{4K_1^2}{K_0g(0)}\right)/\rho,$$

where the last equality is due to that $A/2=A^1=A^2$. Since the value of $e_L^{\circ *}-e_L^{1*}-e_L^{2*}$ depends on the $\frac{4K_1^2}{K_0g(0)}$, the discussion of the value is similar to (a). Therefore, there exists a threshold on x_H/x_L above which $e_L^{\circ *}-e_L^{1*}-e_L^{2*}<0$ and under which $e_L^{\circ *}-e_L^{1*}-e_L^{2*}\geq 0$.

(i)(d) By Lemma H.2, the equilibrium effort level in the second sub-contest is $e_L^{2*} = \ln(A^2 g(\ln(x_L/x_H)/\rho)x_L/\rho)/\rho$ if his opponent is H-type. Thus, the difference of the equilibrium effort levels between two contest mechanisms is

$$\begin{aligned} e_L^{\circ *} - e_L^{1*} - e_L^{2*} &= 2\ln(AK_1 x_L/\rho)/\rho - \ln(A^1 K_0 x_L/\rho)/\rho - \ln(A^2 g(\ln(x_L/x_H)/\rho)x_L/\rho)/\rho \\ &= \ln\left(\frac{4K_1^2}{K_0 g(\ln(x_L/x_H)/\rho)}\right)/\rho = \ln\left(\frac{4K_1^2}{K_0 g(\ln(x_H/x_L)/\rho)}\right)/\rho, \end{aligned}$$

where the last equality is due to the symmetry of $g(\cdot)$. Since the value of $e_L^{\circ *} - e_L^{1*} - e_L^{2*}$ depends on the $\frac{4K_1^2}{K_0 g(\ln(x_H/x_L)/\rho)}$, the discussion of the value is similar to (b). Therefore, $e_L^{\circ *} - e_L^{1*} - e_L^{2*} > 0$ for any x_H/x_L.

(ii) If the expertise along the two attributes is negatively correlated, then by Lemma H.3, the effort level will be the same for either type of contestants, i.e., $e^{\circ *} = 2\ln(Ag(0)\sqrt{x_H x_L}/\rho)\rho$. If both contestants are type (H, L), the equilibrium effort level of the contestant in the second sub-contest is $e_H^{2*} = \ln(A^2 g(0)x_H/\rho)/\rho$. The difference of the equilibrium efforts between two contest mechanisms is

$$\begin{aligned} e^{\circ *} - e_H^{1*} - e_L^{2*} &= 2\ln(Ag^\circ(0)\sqrt{x_H x_L}/\rho)\rho - \ln(A^1 K_0 x_H/\rho)\rho - \ln(A^2 g(0)x_L/\rho)/\rho \\ &= \ln\left(\frac{4g^\circ(0)^2}{K_0 g(0)}\right)/\rho \geq \ln\left(\frac{4g^\circ(0)^2}{g(0)^2}\right)/\rho > 0, \end{aligned}$$

where the first inequality is due to $g(\ln(x_H/x_L)/\rho) \leq g(0)$ and the second inequality is due to $g^\circ(0) > g(0)/2$. Therefore, we have $e^{\circ *} - e_H^{1*} - e_L^{2*} > 0$. If a contestant with type (H, L) has an opponent with type (L, H), then

$$\begin{aligned} e^{\circ *} - e_H^{1*} - e_L^{2*} &= 2\ln(Ag^\circ(0)\sqrt{x_H x_L}/\rho)\rho - \ln(A^1 K_0 x_H/\rho)\rho - \ln(A^2 g(\ln(x_L/x_H))x_L/\rho)/\rho \\ &= \ln\left(\frac{4g^\circ(0)^2}{K_0 g(\ln(x_H/x_L)/\rho)}\right)/\rho \geq \ln\left(\frac{4g^\circ(0)^2}{g(0)^2}\right)/\rho > 0, \end{aligned}$$

where the first inequality is due to $g(\ln(x_H/x_L)/\rho) \leq g(0)$ and the second inequality is due to $g^\circ(0) > g(0)/2$.

Therefore, for the negatively correlated expertise, for the contestant with type (H, L), the effort level is higher in the joint contest than in the separate contest. The same result holds for a contestant with type (L, H) due to the symmetry. □

The explanation is as follows.

1. Perfectly positive correlation:

   If the two contestants are of the same type, by information revelation, in the second sub-contest, both contestants know their opponents' type. Therefore, both contestants have a chance to win so they are motivated to make high effort. If the two contestants are of different types, in the second sub-contest, both contestants will make relatively low effort since the low type contestant feels little chance to win and the high type contestant feels no need to make high effort. Moreover, in this case, the effort level is decreasing in the difference between the high and low expertise because if the difference is large, the low type contestant barely has any chance to win. However, in the first sub-contest or in the joint contest, contestants do not know their opponents' type but they will take into account all the possibilities that their opponents can be a high type or low type.

For Proposition H.1(i)(a) and (i)(c), if both contestants are of the same type and the difference between the high and low expertise is large, then both contestants will make relatively low effort in the first sub-contest and the joint contest since they take into account the possibility that they may compete with a contestant with a different type. However, the information revelation leads to that both contestants will make high effort in the second sub-contest. Overall, contestants will make high effort in the separate contest than in the joint contest. In other words, the effect of information revelation dominates the pooling effect if both contestants are of the same type and the difference between the high and low expertise is large. If both contestants are of the same type and the difference between the high and low expertise is small, then both contestants will make relatively low effort in both contest mechanisms. Thus, the effect of information revelation is relatively weak and dominated by the pooling effect.

For Proposition H.1(i)(b) and (i)(d), if the two contestants are of different types, then the information revelation leads to low effort in the second sub-contest of the separate contest. Therefore, the information revelation has a negative effect on the effort level of contestants, so each contestant's effort level is higher in the joint contest than in the separate contest.

2. Perfectly negative correlation:

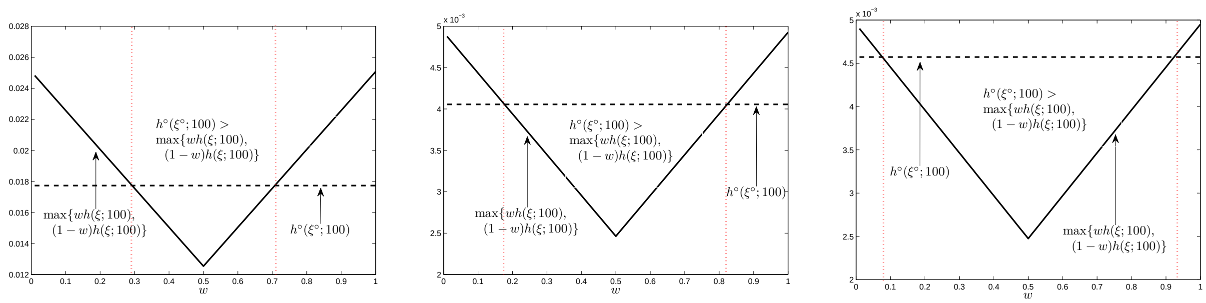
   If contestants' expertise along the two attributes is negatively correlated, each contestant always has a higher equilibrium effort level in the joint contest than in the separate contest, since the effect of information revelation, which leads to slack off in the second sub-contest of the separate contest, always exists. However, for the joint contest, contestants are ex ante identical in the expertise, there is no heterogeneity effect. As a result, under the additional influence of the pooling effect, each contestant has a higher equilibrium effort level in the joint contest than in the separate contest.

I. Numerical Study

I.1. Distributions

We perform a numerical study of several commonly used distributions for random factors and display in Figure I.1 the comparison of the functions $h^\circ(\xi^\circ;n)$ and $\max\{wh(\xi;n),(1-w)h(\xi;n)\}$. In Figure I.1(a), the normal distribution has mean 0 and $\sigma=1$. In Figure I.1(b), the logistic distribution has mean 0 and scale 1. In Figure I.1(c), the Gumbel distribution has mean 0 and scale 1. The contestant number is set to 100, i.e., $n=100$. The horizontal dash line is the value of $h^\circ(\xi^\circ;100)$, and the solid line is the value of $\max\{wh(\xi;100),(1-w)h(\xi;100)\}$ with respect to w. The segment between the two vertical dotted lines on the w-axis contains the values of w that satisfy the condition $h^\circ(\xi^\circ;100)>\max\{wh(\xi;100),(1-w)h(\xi;100)\}$.

Figure I.1 Comparison between $h^\circ(\xi^\circ;n)$ and $\max\{wh(\xi;n),(1-w)h(\xi;n)\}$



(a) Normal distribution (b) Logistic distribution (c) Gumbel distribution

I.2. Independent Expertise

In the joint contest, there are four types of contestants: (L,L), (L,H), (H,L) and (H,H), and the prior probability for those types is η^2, $\eta(1-\eta)$, $\eta(1-\eta)$ and $(1-\eta)^2$ respectively. Since the expertise levels are the same for both attributes, we recognize the contestants with types (L,H) and (H,L) in the joint contest as M-type, and he makes effort $e_M^{\circ*}$ in the symmetric equilibrium. For the contestant with type (L,L) or (H,H), he makes effort $e_L^{\circ*}$ or $e_H^{\circ*}$ respectively in the symmetric equilibrium. The expected payoff to contestant i is $\mathsf{E}(u_i^\circ(e_i^\circ))=A[\eta^2G^\circ(e_i^\circ-e_L^{\circ*})+(1-\eta)^2G^\circ(e_i^\circ-e_H^{\circ*})+2\eta(1-\eta)G^\circ(e_i^\circ-e_M^{\circ*})]-C_i^\circ(e_i^\circ)$, where $i=L,M,H$. Therefore, the FOC is given by $A[\eta^2g^\circ(e_i^{\circ*}-e_L^{\circ*})+(1-\eta)^2g^\circ(e_i^{\circ*}-e_H^{\circ*})+2\eta(1-\eta)g^\circ(e_i^{\circ*}-e_M^{\circ*})]=C_i^{\circ\prime}(e_i^{\circ*})$.

Now we derive the expressions of $C_i^{\circ\prime}(e_i^\circ)$, $i=L,M,H$. If contestant i is M-type, he can be type (H,L) or (L,H). Since the cost functions for those two types are the same, we derive the expression of the $C_i^{\circ\prime}(e_i^\circ)$ for type (L,H), which will be the same for type (H,L). If contestant i has expertise x_L in the first attribute and x_H in the second attribute, given the aggregate effort e_i°, there exists an optimal allocation of efforts $e_i^\circ = \tilde{e}_i^1 + \tilde{e}_i^2$ such that $C_M^\circ(e_i^\circ) = \min\{C(\tilde{e}_i^1)/x_L + C(\tilde{e}_i^2)/x_H\}$. By Lemma 2, the optimal allocation of efforts satisfies $C'(\tilde{e}_i^1)/x_L = C'(\tilde{e}_i^2)/x_H$. Then,

$$C'(\tilde{e}_i^1)/C'(\tilde{e}_i^2) = x_L/x_H. \tag{I.1}$$

When $C'(e_i) = \rho\exp(\rho e_i)$, (I.1) becomes $\exp(\rho(\tilde{e}_i^1 - \tilde{e}_i^2)) = x_L/x_H$, equivalently $\tilde{e}_i^1 - \tilde{e}_i^2 = \ln(x_L/x_H)/\rho$. Since $e_i^\circ = \tilde{e}_i^1 + \tilde{e}_i^2$, we have $\tilde{e}_i^1 = [e_i^\circ + \ln(x_L/x_H)/\rho]/2$ and $\tilde{e}_i^2 = [e_i^\circ - \ln(x_L/x_H)/\rho]/2$. Since the total cost is $C_M^\circ(e_i^\circ) = C(\tilde{e}_i^1)/x_L + C(\tilde{e}_i^2)/x_H$, the derivative of the total cost function is

$$\begin{aligned} C_M^{\circ\prime}(e_i^\circ) &= \frac{\rho}{2}\exp\left(\frac{\rho e_i^\circ + \ln(x_L/x_H)}{2}\right)/x_L + \frac{\rho}{2}\exp\left(\frac{\rho e_i^\circ - \ln(x_L/x_H)}{2}\right)/x_H \\ &= \rho\exp(\rho e_i^\circ/2)/\sqrt{x_H x_L}. \end{aligned} \tag{I.2}$$

By (I.2), the expressions of the derivative of the cost functions for H-type and L-type contestants are $C_H^{\circ\prime}(e_i^\circ) = \rho\exp(\rho e_i^\circ/2)/x_H$, and $C_L^{\circ\prime}(e_i^\circ) = \rho\exp(\rho e_i^\circ/2)/x_L$.

Now we can derive the equilibrium effort levels in the joint contest. In the symmetric equilibrium, contestants with the same type make the same effort, thus we obtain

$$\begin{aligned} A[\eta^2 g^\circ(0) + (1-\eta)^2 g^\circ(e_L^{\circ*} - e_H^{\circ*}) + 2\eta(1-\eta)g^\circ(e_L^{\circ*} - e_M^{\circ*})] &= C_L^{\circ\prime}(e_L^{\circ*}) = \rho\exp(\rho e_L^{\circ*}/2)/x_L, \\ A[\eta^2 g^\circ(e_H^{\circ*} - e_L^{\circ*}) + (1-\eta)^2 g^\circ(0) + 2\eta(1-\eta)g^\circ(e_H^{\circ*} - e_M^{\circ*})] &= C_H^{\circ\prime}(e_H^{\circ*}) = \rho\exp(\rho e_H^{\circ*}/2)/x_H, \\ A[\eta^2 g^\circ(e_M^{\circ*} - e_L^{\circ*}) + (1-\eta)^2 g^\circ(e_M^{\circ*} - e_H^{\circ*}) + 2\eta(1-\eta)g^\circ(0)] &= C_M^{\circ\prime}(e_M^{\circ*}) = \rho\exp(\rho e_M^{\circ*}/2)/\sqrt{x_H x_L}. \end{aligned}$$

The solution of the above equations yields the symmetric equilibrium of the contestants' effort levels in the joint contest. The closed form solution can be obtained by assuming $\eta = 1/2$. By $\eta = 1/2$, those equations become

$$A[g^\circ(0)/4 + g^\circ(e_L^{\circ*} - e_H^{\circ*})/4 + g^\circ(e_L^{\circ*} - e_M^{\circ*})/2] = C_L^{\circ\prime}(e_L^{\circ*}) = \rho\exp(\rho e_L^{\circ*}/2)/x_L, \tag{I.3}$$

$$A[g^\circ(e_H^{\circ*} - e_L^{\circ*})/4 + g^\circ(0)/4 + g^\circ(e_H^{\circ*} - e_M^{\circ*})/2] = C_H^{\circ\prime}(e_H^{\circ*}) = \rho\exp(\rho e_H^{\circ*}/2)/x_H, \tag{I.4}$$

$$A[g^\circ(e_M^{\circ*} - e_L^{\circ*})/4 + g^\circ(e_M^{\circ*} - e_H^{\circ*})/4 + g^\circ(0)/2] = C_M^{\circ\prime}(e_M^{\circ*}) = \rho\exp(\rho e_M^{\circ*}/2)/\sqrt{x_H x_L}. \tag{I.5}$$

Divide (I.5) by (I.3) and divide (I.4) by (I.5),

$$\frac{g^\circ(e_M^{\circ*} - e_L^{\circ*})/4 + g^\circ(e_M^{\circ*} - e_H^{\circ*})/4 + g^\circ(0)/2}{g^\circ(0)/4 + g^\circ(e_L^{\circ*} - e_H^{\circ*})/4 + g^\circ(e_L^{\circ*} - e_M^{\circ*})/2} = \exp(\rho(e_M^{\circ*} - e_L^{\circ*})/2)\sqrt{\frac{x_L}{x_H}}, \tag{I.6}$$

$$\frac{g^\circ(e_H^{\circ*} - e_L^{\circ*})/4 + g^\circ(0)/4 + g(e_H^{\circ*} - e_M^{\circ*})/2}{g^\circ(e_M^{\circ*} - e_L^{\circ*})/4 + g^\circ(e_M^{\circ*} - e_H^{\circ*})/4 + g^\circ(0)/2} = \exp(\rho(e_H^{\circ*} - e_M^{\circ*})/2)\sqrt{\frac{x_L}{x_H}}. \tag{I.7}$$

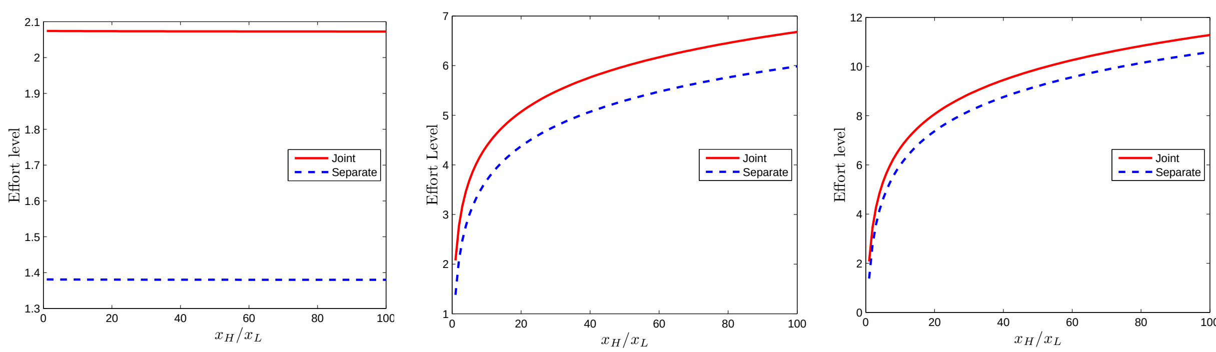
We perform the numerical study to solve (I.6) and (I.7) so that we can obtain the effort level of the H-type, L-type and M-type contestants.

For the separate contest, since the expertise of the two attributes is independent, there is no learning behavior and the equilibrium effort level is given by Lemma H.1. In sub-contest l, the equilibrium effort levels are $e_L^{l*} = \ln(A^l K_0 x_L/\rho)\rho$ and $e_H^{l*} = \ln(A^l K_0 x_H/\rho)\rho$, where $K_0 = g(0)/2 + g(\ln(x_H/x_L)/\rho)/2$.

For the H-type contestant, the effort level in the joint contest is $e_H^{\circ*}$ and the effort level in the separate contest is $e_H^{1*} + e_H^{2*}$. For the L-type contestant, the effort level in the joint contest is $e_L^{\circ*}$ and the effort level in the separate contest is $e_L^{1*} + e_L^{2*}$. For the M-type contestant, the effort level in the joint contest is $e_M^{\circ*}$ and the effort level in the separate contest is $e_L^{1*} + e_H^{2*}$ or $e_H^{1*} + e_L^{2*}$.

To guarantee the equilibrium existence, we allow σ to be high, so $\sigma = 100$. Since we focus on the positive effort level, the total prize should be high enough such that the inverse cost function, which is a logarithm function, can induce a positive value. As a result, $A = 1000$. Moreover, we let $\rho = 1$ and $x_H/x_L \in [1, 100]$.

Figure I.2 Comparison of effort levels for independent expertise



(a) Low type contestant (b) Median type contestant (c) High type contestant

Figure I.2 shows the effort levels of low type, median type and high type contestant in the two contest mechanisms respectively. The dashed line shows the effort level in the separate contest, while the solid line shows the effort level in the joint contest. By Figure I.2, we observe that the effort level of each contestant is higher in the joint contest than in the separate contest.

For the independent expertise case, the heterogeneity exists in both the joint and separate contests. Meanwhile, the two sub-contests and the joint contest are games with incomplete information. Since both contest mechanisms have the effect of heterogeneity which leads to the slack-off behavior of contestants, the pooling effect is expected to be dominating. Thus, contestants are expected to have a higher effort level in the joint contest than in the separate contest. □

J. Two-Person Model

Consider a two-person model. If $\xi^l \sim N(0,\sigma)$, $l=1,2$, and $n=2$, by the formula of $h(\xi^l;n)$,

$$h(\xi^l;2)=\int_{-\infty}^{+\infty}\psi(\xi^l)^2 d\xi^l=\int_{-\infty}^{+\infty}\psi(\xi^l)\psi(-\xi^l)d\xi^l=\frac{1}{2\sqrt{\pi}\sigma}.$$

By Lemma 1, the equilibrium effort for sub-contest l in the separate contest is $e^{l*}=C^{l'^{-1}}(A^l/(2\sqrt{\pi}\sigma))$. In the joint contest, $\xi^\circ=\xi^1+\xi^2$ follows $N(0,2\sigma)$. Thus, $h^\circ(\xi^\circ;2)=1/(2\sqrt{2\pi}\sigma)$. By Lemma 3, the equilibrium effort in the joint contest is $e^{\circ *}=C^{\circ'^{-1}}(A/(2\sqrt{2\pi}\sigma))$. For Proposition 1(ii), the difference between the equilibrium efforts in the two contest mechanisms is $\Delta^e=C^{1'^{-1}}(wA/(2\sqrt{\pi}\sigma))+C^{2'^{-1}}((1-w)A/(2\sqrt{\pi}\sigma))-C^{1'^{-1}}(A/(2\sqrt{2\pi}\sigma))-C^{2'^{-1}}(A/(2\sqrt{2\pi}\sigma))$. One sufficient condition for $\Delta^e<0$ is $h^\circ(\xi^\circ;2)>\max\{wh(\xi;2),(1-w)h(\xi;2)\}$, which is satisfied if $w\in(1-\sqrt{2}/2,\sqrt{2}/2)\approx(0.29,0.71)$. The inequality is due to $C^{l'}(\cdot)>0$, $l=1,2$. For Proposition 1 (ii), the condition $h^\circ(\xi^\circ;2)>h(\xi;2)/2$ can be naturally satisfied by the normal distribution because $2\sqrt{2}\sigma<4\sigma$.

References

Ales, L, SH Cho, E Körpeoğlu. 2016. Innovation tournaments with multiple contributors. https://ssrn.com/abstract=2277215.

Bagnoli, M, T Bergstrom. 2005. Log-concave probability and its applications. *Econ. Theory.* **26**(2) 445–469.

Chakraborty, I 1999. Bundling decisions for selling multiple objects. *Econ. Theory.* **13**(3) 723–733.

David, HA, HN Nagaraja. 2003. *Stochastic Orders.* Wiley, New York.

Fudenberg, D, J Tirole. 1991. *Game Theorey.* MIT Press, Cambridge, MA.

Shaked, M, JG Shanthikumar. 2007. *Stochastic Orders.* Springer Science & Business Media.

Watson, R, I Gordon. 1986. On quantiles of sums. *Austral. J. Statist.* **28**(2) 192–199.