Appendix to

"The First Deal: The Division of Founder Equity in New Ventures"

Thomas Hellmann and Noam Wasserman

January 2016

1 Model assumptions

We briefly restate the main assumptions of the model. There are two founders, A and B. A can have one of two possible skill levels: with probability q, A has high skills (ϕ_H), otherwise low skills (ϕ_L with $\phi_H > \phi_L$). B always has low skills ϕ_L. After discovery founders are either symmetric (i.e., $\phi_L^A = \phi_L^B$) or asymmetric (i.e., $\phi_H^A > \phi_L^B$). Founders contribute some of their own resources (financial capital and/or intellectual property). The total cost of providing resources is R, the relative shares are given by $(0.5 + r)R$ for A and $(0.5 - r)R$ for B, so that r ($\in [-0.5, 0.5]$) describes founder resource asymmetries.

At the first stage founders decide whether or not to engage in costly discovery or not. Discovery costs are borne jointly by the two founders and incurred whenever it is jointly efficient to do so. We denote discovery costs by k and assume that they have some cumulative distribution function $K(k)$ over $[0, \infty]$. With or without discovery, the founders then agree on a division of equity. We denote A's share premium σ, so that A gets an equity stake of $0.5 + \sigma$ and B of $0.5 - \sigma$.

At the second stage founders can provide effort e_i ($i = A, B$) at a private cost of $c(e_i) = 0.5\psi e_i^2$. At the third stage the venture either reaches a performance milestone, creating a return normalized to 1, or else fails and generates no returns. The probability of success is given by $p = \phi_A e_A + \phi_B e_B$, so that success is a function of efforts and skills.

As discussed in the main text, founders have outcome-inequality aversion (OIA). In the base model there is a single parameter α which has a distribution $\Omega(\alpha)$ with $\alpha \in [0, \infty])$. In case of success, founders' utilities are denoted by $w_A = (0.5+\sigma) - \alpha|\sigma|$ and $w_B = (0.5-\sigma) - \alpha|\sigma|$. At the private effort stage the utilities are given by have $v_i^j = p^j w_i - c_i$ where $p^j = \phi_A^j e_A + \phi_B e_B$, which all depends on the skill realization $j = \{H, L\}$. At the initial contracting stage we have $u_i = qv_i^H + \overline{q}v_i^L - 0.5kd$, where $d = \{0, 1\}$ ($d = 1$ means discovery costs are incurred). We denote the joint utilities by $v^j = v_A^j + v_B^j$ and $u = u_A + u_B$. We also distinguish between v_d^j for $d = \{0, 1\}$, and similarly for v_d and u_d. The optimal equity split with discovery is found by maximizing v_1^j for each realization of j. Without discovery the optimal equity split is found by maximizing u without knowledge of the realization of j. For now we assume that $u_A \geq (0.5+r)R$ and $u_B \geq (0.5-r)R$, so that both founders' participation constraint is always satisfied.

2 Derivation of optimal contracts

2.1 Optimal effort

We solve the model backwards, starting with the optimal effort choice. Founders choose e_i to maximize $v_i^j = p^j w_i - c_i$. Solving the first order condition we obtain $e_i = \frac{\phi_i w_i}{\psi}$. We define $\Phi_A = \Phi_j = \frac{\phi_j^2}{\psi}$, $\Phi_B = \Phi_L$, $\Phi = \frac{\Phi_H}{\Phi_L}$ and $\theta = \frac{\Phi - 1}{\Phi + 1}$ (where $\frac{d\theta}{d\Phi} = \frac{2}{(\Phi+1)^2} > 0$). The optimized value of v_i^j is given by $v_A^j = \frac{1}{2}\Phi_j w_A^2 + \Phi_L w_A w_B$ and $v_B^j = \frac{1}{2}\Phi_L w_B^2 + \Phi_j w_A w_B$.

2.2 Optimal contracts with discovery

Consider now optimal contracts with discovery. This consists of choosing an optimal equity premium σ^j, anticipating its subsequent effect on private efforts, as reflected in the optimized values of v_i^j . For the asymmetric case where $j = H$ it is easy to see that $\sigma \geq 0$, so that $w_A = \frac{1}{2} + \sigma - \alpha\sigma$ and $w_B = \frac{1}{2} - \sigma - \alpha\sigma$. We use this in $v_1^H = v_A^H + v_B^H$ and obtain 'after transformations' ('a.t.' henceforth) $v_1^H = \frac{\Phi_H + \Phi_L}{2}[\frac{3}{4} + \theta\sigma_1^H - 3\alpha\sigma_1^H - (1 + 2\alpha\theta - 3\alpha^2)(\sigma_1^H)^2]$. From the first order condition of maximizing v_1^H we find that the optimal equity premium is given by $\sigma_1^H = \frac{1}{2}\frac{\theta - 3\alpha}{1 + 2\alpha\theta - 3\alpha^2}$. Totally differentiating σ_1^H we find $\frac{d\sigma_1^H}{d\alpha} = -\frac{1}{2}\frac{3 + \theta^2 + (\theta - 3\alpha)^2}{(1 + 2\alpha\theta - 3\alpha^2)^2} < 0$ and $\frac{d\sigma_1^H}{d\Phi} = \frac{1}{(\Phi+1)^2}\frac{1 + 3\alpha^2}{(1 + 2\alpha\theta - 3\alpha^2)^2} > 0$. This says that the optimal premium is decreasing in α (intuitively, higher OIA reduces the optimal premium), and increasing in Φ (intuitively, greater skill differences increase the optimal premium).

The above expression for the optimal equity premium remains valid as along as $\sigma_1^H \geq 0 \Leftrightarrow \alpha \leq \frac{\theta}{3} \equiv \alpha_1$. For $\alpha > \alpha_1$ the optimal split is simply $\sigma_1^H = 0$. We note that $\frac{d\alpha_1}{d\Phi} = \frac{2}{3(\Phi+1)^2} > 0$ (intuitively, greater skill differences increase the range where unequal splitting remains optimal).

For the remainder of the proof it is also useful to collect expressions for the optimized utilities. For $\alpha \leq \alpha_1$, the optimized utility is given by $v_1^H = \frac{\Phi_H + \Phi_L}{2}[\frac{3}{4} + \frac{1}{4}\frac{(\theta - 3\alpha)^2}{1 + 2\alpha\theta - 3\alpha^2}]$. For $\alpha > \alpha_1$, the optimized utility is given by $v_1^H = \frac{3}{4}\frac{\Phi_H + \Phi_L}{2}$ (which is constant in α).

For the case of discovery with symmetric skills we use $v_1^L = \Phi_L[\frac{3}{4} - 3\alpha\sigma - (1-3\alpha^2)\sigma^2]$ and note that $\frac{dv_1^L}{d\sigma} = \Phi_L[-3\alpha - 2(1-3\alpha^2)\sigma] < 0$ so that the optimal split always satisfies $\sigma_1^L = 0$. The optimized utility is given by $v_1^L = \frac{3}{4}\Phi_L$.

2.3 Optimal contracts without discovery

Consider next the case of no discovery. The optimal contract maximizes $u_0 = qv^H + \overline{q}v$. We define $\chi_H = q\frac{\Phi_H + \Phi_L}{2}$, $\chi_L = \overline{q}\Phi_L$, $\chi = \frac{\chi_L}{\chi_H} = \frac{\overline{q}}{q}\frac{2}{1+\Phi}$ and $\gamma = 1 + \chi$. Note that $\frac{d\gamma}{dq} = -\frac{1}{q^2}\frac{2}{1+\Phi} < 0$. We rewrite $u_0 = \chi_H[\frac{3}{4} + \theta\sigma - 3\alpha\sigma - (1 + 2\alpha\theta - 3\alpha^2)\sigma^2] + \chi_L[\frac{3}{4} - 3\alpha\sigma - (1-3\alpha^2)\sigma^2]$. Solving the first-order condition we obtain $\sigma_0^* = \frac{1}{2}\frac{\theta - 3\alpha\gamma}{\gamma - 3\alpha^2\gamma + 2\alpha\theta}$. It is easy to verify that $\frac{d\sigma_0^*}{d\alpha} < 0$ and $\frac{d\sigma_0^*}{dq} < 0$ so that the optimal premium is again decreasing in α and increasing in q. The condition $\sigma_0^* \geq 0$ is now satisfied for $\alpha \leq \frac{\theta}{3\gamma} \equiv \alpha_0$. Since $\gamma > 1$ we have $\alpha_1 > \alpha_0 > 0$. Again we find that $\frac{d\alpha_0}{d\Phi} > 0$.

The optimized utility for $\alpha < \alpha_0$ is given by $u_0 = \chi_H[\frac{3}{4}\gamma + \frac{1}{4}\frac{[\theta - 3\alpha\gamma]^2}{\gamma - 3\alpha^2\gamma + 2\alpha\theta}]$, with $\frac{du_0}{d\alpha} = -\frac{\chi_H}{2}\frac{(\theta - 3\alpha\gamma)(3\gamma^2 + \theta^2)}{(\gamma - 3\alpha^2\gamma + 2\alpha\theta)^2} < 0$. For $\alpha > \alpha_0$ we simply have $\sigma_0^* = 0$ so that $u_0 = \frac{3}{4}(\chi_H + \chi_L)$.

2.4 Renegotiation

We briefly show why renegotiation never occurs in this model. With discovery there is no point in time where additional information emerges, so renegotiation is irrelevant. Without discovery, however, we may ask if the founders want to renegotiate at the point where they discover their actual skills. Consider first the case where the founders discover that $j = L$. The founders originally agreed to σ_0 but ex-post an equal split would have been more efficient. However, since there is no transferable utility in this model, renegotiation can only occur if A is willing to voluntarily give up some equity. We can find the highest level of equity that A wants to have by maximizing $v_A = \frac{1}{2}\Phi_A(\frac{1}{2} + \sigma - \alpha\sigma)^2 + \Phi_L(\frac{1}{2} + \sigma - \alpha\sigma)(\frac{1}{2} - \sigma - \alpha\sigma)$ with respect to σ. Standard calculations show that this occurs at $\sigma_L^{Max} = \frac{1-3\alpha}{2(1+2\alpha-3\alpha^2)}$. We note that $\sigma_L^{Max} > \sigma_0^* \Leftrightarrow \frac{1-3\alpha}{2(1+2\alpha-3\alpha^2)} > \frac{1}{2}\frac{\theta - 3\alpha\gamma}{\gamma - 3\alpha^2\gamma + 2\alpha\theta}$ which yields after transformations $\gamma > \theta$. Thus for $j = L$, A never wants to renegotiate the original equity split. Next we consider the case of $j = H$. Here the question is whether B wants to give up any equity to make A more productive. Maximizing $v_B = \frac{1}{2}\Phi_L(\frac{1}{2} - \sigma - \alpha\sigma)^2 + \Phi_H(\frac{1}{2} + \sigma - \alpha\sigma)(\frac{1}{2} - \sigma - \alpha\sigma)$ with respect to σ we find a.t. $\sigma^{\min} = -\frac{1}{2}\frac{1+\alpha+2\Phi\alpha}{2\Phi - 1 - 2\alpha - \alpha^2 - 2\Phi\alpha^2}$. We note that $\sigma^{\min} \leq 0$ as long as $2\Phi - 1 - 2\alpha - \alpha^2 - 2\Phi\alpha^2 \geq 0$. It is easy to verify that this is true for all $\alpha \in [0, \alpha_0]$. Thus, even if A has high skills, B never voluntary gives A a positive share premium. Thus no renegotiation ever occurs in the model.

3 Results about optimal contracts

3.1 Proposition 1

We are now in a position to examine our main Propositions. We first formally restate all Propositions and then provide their proof.

Founders incur discovery costs whenever $k < u_1(\alpha) - u_0(\alpha) \equiv u_{10}(\alpha)$. We define $\mu(\alpha) = 1 - K(u_{10}(\alpha))$.

Proposition 1: The probability of no discovery $\mu(\alpha)$ is weakly increasing in α. It is strictly increasing for $\alpha < \alpha_1$ and constant for $\alpha > \alpha_1$.

Proof: We note that $\mu(\alpha) = Prob[k > u_{10}(\alpha)] = 1 - K(u_{10}(\alpha))$. For the case of $\alpha > \alpha_1$, we find that $u_{10}(\alpha) = 0$ so that $\mu(\alpha) = 1 - K(0)$ which does not depend on α. For $\alpha \in (\alpha_0, \alpha_1)$ we have $u_0 = q\frac{3}{4}\frac{\Phi_H + \Phi_L}{2} + \overline{q}\frac{3}{4}\Phi_L$ and $u_1 = qv_1^H + \overline{q}\frac{3}{4}\Phi_L$ so that $\frac{du_{10}(\alpha)}{d\alpha} = q\frac{dv_1^{HL}}{d\alpha} < 0$. Thus $\frac{d\mu(\alpha)}{d\alpha} = -K'\frac{du_{10}(\alpha)}{d\alpha} > 0$. For $\alpha < \alpha_0$ we note that since $\frac{dv_1^L}{d\alpha} = 0$ we have $\frac{du_1}{d\alpha} = q\frac{dv_1^H}{d\alpha}$, so that $\frac{du_{10}}{d\alpha} = \frac{du_1}{d\alpha} - \frac{du_0}{d\alpha}$ which after transformation yields $\frac{du_{10}}{d\alpha} = q\frac{\Phi_L + \Phi_H}{4}[\frac{(\theta - 3\alpha\gamma)(3\gamma^2 + \theta^2)}{(\gamma - 3\alpha^2\gamma + 2\alpha\theta)^2} - \frac{(\theta - 3\alpha)(3 + \theta^2)}{(1 + 2\alpha\theta - 3\alpha^2)^2}]$. We claim that $\frac{du_{10}}{d\alpha} < 0$, which is equivalent to claiming that $\xi_1 = \frac{d}{d\gamma}\frac{(\theta - 3\alpha\gamma)(3\gamma^2 + \theta^2)}{(\gamma - 3\alpha^2\gamma + 2\alpha\theta)^2} < 0$ for all γ. After transformation we get $\xi_1 = \frac{(2\theta^2 - 6\alpha\theta\gamma)(6\alpha\gamma + 3\alpha^2\theta - \theta) - (9\alpha\gamma^2 + 9\alpha\theta^2)(\gamma - 3\alpha^2\gamma + 2\alpha\theta)}{[(\gamma - 3\alpha^2\gamma + 2\alpha\theta)]^3}$ so that $\xi_1 < 0 \Leftrightarrow \xi_2 = 9\alpha\gamma^3 - 27\alpha^3\gamma^3 + 54\alpha^2\theta\gamma^2 + 12\alpha^2\theta^3 + 2\theta^3 - 9\alpha\theta^2\gamma - 9\alpha^3\theta^2\gamma > 0$. We note that at $\alpha = 0$ we have $\xi_2 = 2\theta^3 > 0$ and at $\alpha = \alpha_0$ we have $\xi_2 = 3\theta\gamma^2 + 4\theta^3 + \frac{\theta^5}{\gamma^2} > 0$. To show that $\xi_2 > 0$ for all interior $\alpha \in (0, \alpha_0)$ we use $\frac{d\xi_2}{d\alpha} = 9\gamma^3 - 81\alpha^2\gamma^3 + 108\alpha\theta\gamma^2 + 24\alpha\theta^3 - 9\theta^2\gamma - 27\alpha^2\theta^2\gamma$ and $\frac{d^2\xi_2}{d\alpha^2} = -162\alpha\gamma^3 + 108\theta\gamma^2 + 24\theta^3 - 54\alpha\theta^2\gamma$. We note that at $\alpha = 0$ we have $\frac{d\xi_2}{d\alpha} =$

$9\gamma^3 - 9\theta^2\gamma = 9\gamma(\gamma^2 - \theta^2) > 0$ and $\dfrac{d^2\xi_2}{d\alpha^2} = 108\theta\gamma^2 + 24\theta^3 > 0$. Moreover, at $\alpha = \alpha_0$ we have $\dfrac{d\xi_2}{d\alpha} = 9\gamma^3 + 18\theta^2\gamma + 5\dfrac{\theta^4}{\gamma} > 0$ and $\dfrac{d^2\xi_2}{d\alpha^2} = 54\theta\gamma^2 + 8\theta^3 > 0$. This says that ξ_2 is increasing and convex over the entire range $\alpha \in [0, \alpha_0]$, which implies $\xi_2 > 0$ for all interior $\alpha \in (0, \alpha_0)$.$\square$

3.2 Proposition 2

Next we denote the probability of equal splitting by $\lambda(\alpha)$. As a building block we note that conditional on $d = \{0, 1\}$, the probability of equal splitting is given by $\lambda_0(\alpha) = 0$ if $\alpha < \alpha_0$ and $\lambda_0(\alpha) = 1$ if $\alpha \geq \alpha_0$, and similarly $\lambda_1(\alpha) = 0$ if $\alpha < \alpha_1$ and $\lambda_1(\alpha) = 1$ if $\alpha \geq \alpha_1$.

Proposition 2: $\lambda(\alpha)$ is decreasing for $\alpha < \alpha_0$, has a positive discontinuity at α_0, is increasing in α for $\alpha \in [\alpha_0, \alpha_1]$, and has another positive discontinuity at $\alpha = \alpha_1$.

Proof: For $\alpha < \alpha_0$ we have $\lambda(\alpha) = \overline{\mu}(\alpha)\overline{q}$ (note that throughout the appendix a bar over a variable denotes one minus that variable, i.e., $\overline{\mu}(\alpha)$ denotes $1 - \mu(\alpha)$). This is decreasing in α since we know from Proposition 1 that $\mu(\alpha)$ is increasing in α. For $\alpha \in [\alpha_0, \alpha_1]$ we have $\lambda(\alpha) = \mu(\alpha) + \overline{\mu}(\alpha)\overline{q}$ which is increasing in α. Finally for $\alpha \geq \alpha_1$ we have $\lambda(\alpha) = 1$. It is easy to see that λ also increases discontinuously at α_0 and α_1. $\square$

3.3 Proposition 3

The $\lambda(\alpha)$ function is defined conditional on a given α. Now we define the unconditional probability of equal splitting across the entire distribution of α, denoted by Λ and given by $\Lambda = \int_0^{\alpha_0} \overline{\mu}(\alpha)\overline{q} d\Omega(\alpha) + \int_{\alpha_0}^{\alpha_1} [\mu(\alpha) + \overline{\mu}(\alpha)\overline{q}] d\Omega(\alpha) + \int_{\alpha_1}^{1} d\Omega(\alpha)$. We further decompose this into the probability of equal splitting without and with discovery, given by $\Lambda_0 = \int_{\alpha_0}^{1} \dfrac{\mu(\alpha)}{M} d\Omega(\alpha)$ and $\Lambda_1 = \overline{q} + q \int_{\alpha_1}^{1} \dfrac{\overline{\mu}(\alpha)}{1 - M} d\Omega(\alpha)$ where $M = \int_0^1 \mu(\alpha) d\Omega(\alpha)$.

Proposition 3: The probability of equal splitting is higher with discovery ($d = 0$) than without discovery ($d = 0$), provided that $\Omega(\alpha_0)$ is not too large.

Proof: We first define the probabilities of equal splitting using a normalized distribution function for the cases $d = \{0, 1\}$. Specifically, we define $\omega_0(\alpha) = \dfrac{\mu(\alpha)}{M}\omega(\alpha)$ as the density, and $\Omega_0(\alpha) = \int_0^{\alpha} \omega_0(\alpha) d\alpha$ as the cumulative distribution function under no discovery; similarly for $\omega_1(\alpha) = \dfrac{\overline{\mu}(\alpha)}{1 - M}\omega(\alpha)$ and $\Omega_1(\alpha) = \int_0^{\alpha} \omega_1(\alpha) d\alpha$. By construction $\Omega_0(0) = \Omega_1(0) = 0$ and $\Omega_0(1) = \Omega_1(1) = 1$. From Proposition 1 we know that $\mu(\alpha)$ is an increasing function of α, and thus $\overline{\mu}(\alpha)$ is an increasing function of α. Thus for any $\alpha \in (0, 1)$ we must have $\Omega_0(\alpha) < \Omega_1(\alpha)$, which says that Ω_0 is first-order stochastic dominant to Ω_1. We rewrite $\Lambda_0 = \int_{\alpha_0}^{1} d\Omega_0(\alpha)$ and $\Lambda_1 = \overline{q} + q \int_{\alpha_1}^{1} d\Omega_1(\alpha)$, and obtain a.t. $\Lambda_0 - \Lambda_1 = [\Omega_1(\alpha_0) - \Omega_0(\alpha_0)] + q\Omega_1(\alpha_1) - \Omega_1(\alpha_0)$.

The first term in the brackets is always positive because Ω_0 is first-order stochastic dominant to Ω_1. The second terms is also clearly positive, and the third term clearly negative. The condition $\Lambda_0 - \Lambda_1 \geq 0$ holds whenever $\Omega_1(\alpha_0)$ is not too large, which is equivalent to $\Omega(\alpha_0)$ being not too large. It is easy to verify that at $q \to 1$ we have $\alpha_0 \to \alpha_1$ so that $\Lambda_0 - \Lambda_1 = \Omega_1(\alpha_0) - \Omega_0(\alpha_0)] > 0$. This implies that for sufficiently large values of q the condition that $\Omega(\alpha_0)$ not too large is always satisfied. For $q \to 0$ we have $\Lambda_0 - \Lambda_1 \to 0$. For sufficiently small values of q the condition that $\Omega(\alpha_0)$ is not too large is binding. To see this consider the extreme case of $\Omega(\alpha_0) = 1$, i.e., when there are no medium or high OIA types. In this case we get $\Lambda_0 = 0$ and $\Lambda_1 = \overline{q}$ so that $\Lambda_0 - \Lambda_1 = -\overline{q} < 0$.□

3.4 Proposition 4

In our model founders use joint efficiency bargaining, not Nash bargaining. Resources therefore only matter when they affect the participation constraint. We now relax the assumption that $u_A \geq (0.5 + r)R$ and $u_B \geq (0.5 - r)R$. We ask what happens when a founder's participation constraint ('PC' henceforth) becomes binding. This allows us to examine the role of resources asymmetries, as captured by r.

Consider first the case with discovery and symmetric skills. Without binding PCs we know that $\sigma_1^L = 0$ is optimal. To satisfy the PCs we need $v_A^L = \frac{3}{8}\Phi_L \geq (0.5 + r)R \Leftrightarrow r \leq \overline{r}_1^L \equiv \frac{3\Phi_L}{8R} - \frac{1}{2}$ and $v_B^L = \frac{3}{8}\Phi_L \geq (0.5 - r)R \Leftrightarrow r \geq \underline{r}_1^L \equiv \frac{1}{2} - \frac{3\Phi_L}{8R}$. For $r > \overline{r}_1^L$, A's PC becomes binding, so that $v_A^L = \frac{1}{2}\Phi_L(\frac{1}{2} + \sigma - \alpha\sigma)^2 + \Phi_L(\frac{1}{2} + \sigma - \alpha\sigma)(\frac{1}{2} - \sigma - \alpha\sigma) = (0.5 + r)R$. It is easy to see that this requires $\sigma_1^L > 0$. Similarly, for $r < \underline{r}_1^L$, B's PC becomes binding, so that $v_B^l = \frac{1}{2}\Phi_L(\frac{1}{2} - \sigma - \alpha\sigma)^2 + \Phi_L(\frac{1}{2} + \sigma - \alpha\sigma)(\frac{1}{2} - \sigma - \alpha\sigma) = (0.5 - r)R$, which requires $\sigma_1^L < 0$. We note that lower values of $|r|$ are associated with $\sigma = 0$, whereas higher values of $|r|$ (i.e., sufficiently large positive or large negative values of r) with $\sigma \neq 0$.

Consider next the model with discovery and asymmetric skills. In the high OIA region where $\alpha > \alpha_1$, the symmetric equilibrium $\sigma_1^H = 0$ remains optimal as long as $v_A^H = \frac{\Phi_H + 2\Phi_L}{8} \geq (0.5 + r)R \Leftrightarrow r \leq \overline{r}_1^H \equiv \frac{\Phi_H + 2\Phi_L}{8R} - \frac{1}{2}$ and $v_B^H = \frac{2\Phi_H + \Phi_L}{8} \geq (0.5 - r)R \Leftrightarrow r \geq \underline{r}_1^H \equiv \frac{1}{2} - \frac{2\Phi_H + \Phi_L}{8R}$. For $r > \overline{r}_1^H$, A's PC becomes binding again, requiring $\sigma_1^H > 0$. Similarly, for $r < \underline{r}_1^H$, B's PC becomes binding, requiring $\sigma_1^H < 0$. We note again that more extreme values of r (i.e., higher values of $|r|$) are associated with $\sigma \neq 0$.

For $\alpha < \alpha_1$, the optimal premium without binding PCs is already positive. It is easy to see that for sufficiently high values of r, A's PC becomes binding, requiring the optimal premium to increase further. For sufficiently low values of r, B's PC becomes binding, requiring lower optimal premium. For a single parameter value (which we think of as a measure zero event) we obtain $\sigma = 0$, but for all others we obtain $\sigma \neq 0$. Overall we note that binding PCs do not change the basic result that $\sigma \neq 0$ for $\alpha < \alpha_1$.

The model without discovery is very similar to the model with asymmetric discovery. For $\alpha > \alpha_0$ the symmetric equilibrium $\sigma_0 = 0$ remains optimal as long as $v_A^H = \frac{q\Phi_H + 2q\Phi_L + 3\overline{q}\Phi_L}{8} \geq (0.5 + r)R \Leftrightarrow r \leq \overline{r}_0 \equiv \frac{q\Phi_H + 2q\Phi_L + 3\overline{q}\Phi_L}{8R} - \frac{1}{2}$ and $v_B^H = \frac{2q\Phi_H + q\Phi_L + 3\overline{q}\Phi_L}{8} \geq (0.5 - r)R \Leftrightarrow r \geq \underline{r}_0 \equiv \frac{1}{2} - \frac{2q\Phi_H + q\Phi_L + 3\overline{q}\Phi_L}{8R}$. For $r > \overline{r}_0$, A's PC becomes binding again, requiring $\sigma_0 > 0$. Similarly, for $r < \underline{r}_0$, B's PC becomes binding, requiring $\sigma_0 < 0$. We note again that more extreme values of r (i.e., higher values of $|r|$) are associated with $\sigma \neq 0$. Finally, for $\alpha < \alpha_0$ the analysis is analogous to the analysis of $\alpha < \alpha_1$ in the model with asymmetric discovery.

We have thus shown the following result:

Proposition 4: The probability of equal splitting μ is independent of r as long as both PCs are not binding. It is lower for sufficiently high values of r (where A's PC becomes binding) or for sufficiently low values of r (where B's PC becomes binding).

It is worth briefly commenting on the connection between Proposition 4 and our empirical identification strategy. Proposition 4 says that equal splitting becomes less likely for more extreme values of r, i.e., for more unequal founder resource commitments. This establishes the rank condition for identification. It is easy to see that the exclusion restriction also holds in this model. This is because the probability of success p is not directly a function of r. It can only be affected by r indirectly through the endogenous efforts e_i which are a function of the optimal premium σ.

For simplicity our theory holds R constant, but one can also think of a model extension where R is related to performance. While we might expect total resources R to matter for performance, this is not true for r, which only concerns asymmetries about which founders contributed the resources R. In the empirical analysis we control for total resources (related to R) through our average founder variables (such as AVG IDEAS and AVG CAPITAL). The heterogeneity measures (COV IDEAS and COV CAPITAL) then capture the asymmetries in founder contributions (related to r).

4 Results about performance

4.1 Proposition 5

We now turn to the performance analysis. For this we first derive explicit analytical functions for the value of p at the optimal contract. Consider the case with discovery first. Using the optimal effort choices in $p = \phi_A e_A + \phi_B e_B$ yields $p = \Phi_A(\frac{1}{2} + \sigma - \alpha\sigma) + \Phi_L(\frac{1}{2} - \sigma - \alpha\sigma)$. For $j = L$ equal splitting is optimal so that performance is given by $p_1^L = \Phi^L$. For $j = H$ and $\alpha \geq \alpha_1$ equal spitting is also optimal, so that $p_1^H = \frac{\Phi_A + \Phi_L}{2}$. For $j = H$ and $\alpha < \alpha_1$ we use $\sigma_1^* = \frac{1}{2}\frac{\theta - 3\alpha}{1 - 3\alpha^2 + 2\alpha\theta}$ and obtain a.t. $p_1^H = \frac{\Phi_H + \Phi_L}{2}\frac{1 - 2\alpha\theta + \theta^2}{1 - 3\alpha^2 + 2\alpha\theta}$. We note that $\frac{dp_1^H}{d\alpha} = -\frac{\Phi_H + \Phi_L}{[1 - 3\alpha^2 + 2\alpha\theta]^2}\xi_3$ where $\xi_3 = 2\theta + \theta^3 - 3\alpha - 3\alpha\theta^2 + 3\alpha^2\theta$. We note that $\frac{dp_1^H}{d\alpha} < 0 \Leftrightarrow \xi_3 \geq 0$. We use $\frac{d\xi_3}{d\alpha} = -3 - 3\theta^2 + 6\alpha\theta$, which at $\alpha = 0$ satisfies $\frac{d\xi_3}{d\alpha} = -3 - 3\theta^2 < 0$ and at $\alpha = \alpha_1$ satisfies $\frac{d\xi_3}{d\alpha} = -3 - \theta^2 < 0$. This says that ξ_3 is decreasing in α over the entire range $\alpha \in [0, \alpha_1]$. At $\alpha = 0$ we have $\xi_3 = 2\theta + \theta^3 > 0$, and at $\alpha = \alpha_1$ we have a.t. $\xi_3 = \theta + \frac{1}{3}\theta^3 > 0$. Thus ξ_3 remains strictly positive over the entire range $\alpha \in [0, \alpha_1]$.

Next we consider the optimal performance without discovery. For $\alpha \geq \alpha_0$ an equal split is optimal, so from the above we know that $p_0^L = \Phi^L$ and $p_0^H = \frac{\Phi_A + \Phi_L}{2}$. Consider thus the case of $\alpha < \alpha_0$ where $\sigma_0^* = \frac{1}{2}\frac{\theta - 3\alpha\gamma}{\gamma - 3\alpha^2\gamma + 2\alpha\theta}$. Suppose $j = H$, then we find a.t. that $p_0^H = \frac{\Phi_H + \Phi_L}{2}\frac{\gamma + \theta^2 + \alpha\theta(1 - 3\gamma)}{(1 - 3\alpha^2)\gamma + 2\alpha\theta}$. We obtain $\frac{dp_0^H}{d\alpha} = -\frac{\Phi_H + \Phi_L}{2}\frac{\xi_4}{[\gamma - 3\alpha^2\gamma + 2\alpha\theta]^2}$ where $\xi_4 = 2\theta^3 + \theta\gamma + 3\theta\gamma^2 - 6\alpha\theta^2\gamma - 6\alpha\gamma^2 + 9\alpha^2\theta\gamma^2 - 3\alpha^2\theta\gamma$. We note that $\frac{dp_0^H}{d\alpha} < 0 \Leftrightarrow \xi_4 > 0$. ξ_4 is a convex quadratic function of α with $\frac{d\xi_4}{d\alpha} = -6\theta^2\gamma - 6\gamma^2 + 18\alpha\theta\gamma^2 - 6\alpha\theta\gamma$ and $\frac{d\xi_4}{d\alpha} = 18\theta\gamma^2 - 6\theta\gamma > 0$. At $\alpha = 0$ we have $\xi_4 = 2\theta^3 + \theta\gamma + 3\theta\gamma^2 > 0$ and at $\alpha = \alpha_1$ we have a.t. $\xi_4 = \theta(3\gamma^2 - \gamma) + \theta^3\frac{3\gamma - 1}{3\gamma} > 0$. Moreover, at $\alpha = 0$ we also have $\frac{d\xi_4}{d\alpha} = 6\gamma(\gamma - \theta^2) > 0$. Thus ξ_4 is positive and increasing throughout the range $\alpha \in [0, \alpha_1]$. In the case of $j = L$ we get a.t. $p_0^L = \Phi_L\frac{\gamma + \alpha\theta}{(1 - 3\alpha^2)\gamma + 2\alpha\theta}$. We have $\frac{dp_0^L}{d\alpha} = -\frac{\Phi_L}{[\gamma - 3\alpha^2\gamma + 2\alpha\theta]^2}\xi_5$ where $\xi_5 = \theta\gamma - 3\alpha^2\theta\gamma - 6\alpha\gamma^2$. At $\alpha = 0$ we have $\xi_5 = \theta\gamma$ so that $\frac{dp_0^L}{d\alpha} < 0$ but at $\alpha = \alpha_0$ we actually have $\xi_5 = -\frac{\theta^3}{3\gamma} - \theta\gamma < 0$. The sign of $\frac{dp_0^L}{d\alpha}$ changes from negative to positive over the range, so that p_0^L is first decreasing and then increasing in α. It is easy to show that the minimum of p_0^L occurs at $\alpha = -\frac{\gamma}{\theta} + \sqrt{(\frac{\gamma}{\theta})^2 + \frac{1}{3}}$.

Under no discovery the behavior of p_0^H and p_0^L is informative, but the most important is the behavior of the expected performance, which we define as $P_0 = qp_0^H + \overline{q}p_0^L$. Given the non-monotonicity of p_0^L we are particularly interested in how P_0 changes with α. After transformations we obtain $P_0 = \chi_H \frac{\gamma(\gamma + \theta^2 - 2\alpha\theta)}{\gamma - 3\alpha^2\gamma + 2\alpha\theta}$. This yields $\frac{dP_0}{d\alpha} = \frac{-2\chi_H\gamma\xi_6}{[\gamma - 3\alpha^2\gamma + 2\alpha\theta]^2}$ where $\xi_6 = 2\theta\gamma + \theta^3 + 3\alpha^2\theta\gamma - 3\alpha\gamma^2 - 3\alpha\theta^2\gamma$. Thus $\frac{dP_0}{d\alpha} < 0 \Leftrightarrow \xi_6 > 0$. ξ_6 is a quadratic and convex function of α with $\frac{d\xi_6}{d\alpha} = 6\alpha\theta\gamma - 3\gamma^2 - 3\theta^2\gamma$. Furthermore at $\alpha = 0$ we have $\frac{d\xi_6}{d\alpha} = -3\gamma^2 - 3\theta^2\gamma < 0$ and at $\alpha = \alpha_1$ we also have a.t. $\frac{d\xi_6}{d\alpha} = -2\theta^2\chi - 3\gamma^2 - \theta^2\gamma < 0$. It follows that ξ_6 is decreasing in α over $\alpha \in [0, \alpha_1]$. At $\alpha = 0$ we have $\xi_6 = 2\theta\gamma + \theta^3 > 0$ and at $\alpha = \alpha_0$ we have a.t. $\xi_6 = \theta\gamma + \frac{\theta^3}{3\gamma} > 0$. Thus $\xi_6 > 0$ for all $\alpha \in [0, \alpha_1]$.

Finally we note that the expected performance under discovery is given by $P_1 = qp_1^H + \overline{q}p_1^L$. For $\alpha < \alpha_1$ we have $\frac{dP_1}{d\alpha} = q\frac{dp_1^H}{d\alpha} < 0$. We have thus proven the following Proposition.

Proposition 5: Expected performance is a decreasing function of α, both with or without discovery, i.e., $\frac{dP_1(\alpha)}{d\alpha} \leq 0$ and $\frac{dP_0(\alpha)}{d\alpha} \leq 0$.

4.2 Proposition 6

Next we consider the expected performance under equal (unequal) splitting for a given α, denoted by $P^{\sigma=0}$ ($P^{\sigma\neq 0}$). We also define $\delta(\alpha) \equiv P^{\sigma=0}(\alpha) - P^{\sigma\neq 0}(\alpha)$ as the performance premium to equal splitting, which could in principle be positive or negative. Note that in the main text we sometimes find it easier to talk about the premium to unequal splitting, which is $-\delta$.

Proposition 6, Part 1: For a given $\alpha < \alpha_1$, equal splitting is associated with lower performance, i.e., $\delta(\alpha) < 0$.

Proof: For $\alpha > \alpha_1$ there is no unequal splitting, so that $P^{\sigma\neq 0}$ is not defined, we can therefore ignore this case. For $\alpha \in [\alpha_0, \alpha_1]$ we get equal splitting under three scenarios: $\{d = 0, j = L\}$, $\{d = 0, j = H\}$ and $\{d = 1, j = L\}$, so that the expected performance is given by $P^{\sigma=0} = \frac{1}{\mu q + \overline{q}}[\mu q \frac{\Phi_H + \Phi_L}{2} + \mu\overline{q}\Phi_L + \overline{\mu q}\Phi_L]$ which can be rewritten as $P^{\sigma=0} = \frac{\Phi_H + \Phi_L}{2} - \frac{\overline{q}}{\mu q + \overline{q}}\frac{\Phi_H - \Phi_L}{2}$. Unequal splitting is only associated with $\{d = 1, j = H\}$ so that $P^{\sigma\neq 0} =$

$p_1^H = \dfrac{\Phi_H + \Phi_L}{2}\dfrac{1-2\alpha\theta+\theta^2}{1-3\alpha^2+2\alpha\theta}$. From $P^{\sigma\neq 0} = \dfrac{\Phi_H + \Phi_L}{2}\dfrac{1-2\alpha\theta+\theta^2}{1-3\alpha^2+2\alpha\theta} > \dfrac{\Phi_H + \Phi_L}{2} > \dfrac{\Phi_H + \Phi_L}{2} - \dfrac{\overline{q}}{\mu q + \overline{q}}\dfrac{\Phi_H - \Phi_L}{2} = P^{\sigma=0}$ we deduce that $\delta < 0$.

For $\alpha < \alpha_0$, equal splitting can only come from $\{d = 1, j = L\}$ so that $P^{\sigma=0} = \Phi_L$. Unequal splitting can come from $\{d = 0, j = L\}$, $\{d = 0, j = H\}$ and $\{d = 1, j = H\}$. We obtain $P^{\sigma\neq 0} = \dfrac{1}{\mu + \overline{\mu}q}[\mu P_0 + \overline{\mu}q\dfrac{\Phi_H + \Phi_L}{2}\dfrac{1-2\alpha\theta+\theta^2}{1-3\alpha^2+2\alpha\theta}]$ where it is useful to rewrite $P_0 = (q\dfrac{\Phi_H + \Phi_L}{2} + \overline{q}\Phi_L)\dfrac{\gamma+\theta^2-2\alpha\theta}{\gamma-3\alpha^2\gamma+2\alpha\theta}$. To show that $\delta < 0$ we use $\dfrac{\Phi_H + \Phi_L}{2} > \Phi_L$, so that a sufficient condition is $\dfrac{\gamma+\theta^2-2\alpha\theta}{\gamma-3\alpha^2\gamma+2\alpha\theta} > 1 \Leftrightarrow \xi_7 = \theta^2 + 3\alpha^2\gamma - 4\alpha\theta > 0$. At $\alpha = 0$ we have $\xi_7 = \theta^2 > 0$ and at $\alpha = \alpha_0$ we have $\xi_7 = \theta^2(1 - \dfrac{1}{\gamma}) > 0$. Moreover, $\dfrac{d\xi_7}{d\alpha} = 6\alpha\gamma - 4\theta$ so that $\dfrac{d\xi_7}{d\alpha} = -4\theta$ at $\alpha = 0$ and $\dfrac{d\xi_7}{d\alpha} = -2\theta < 0$ at $\alpha = \alpha_0$. Thus ξ_7 is decreasing in α over the range $\alpha \in [0, \alpha_0]$ but remains positive throughout.□

For Part 2 we define again the expected performance allowing for self-selection across α. We define

$$\Pi^{\sigma=0} = \frac{\int_0^{\alpha_0}\overline{\mu}(\alpha)\overline{q}\Phi_L d\Omega(\alpha) + \int_{\alpha_0}^{\alpha_1}[\mu(\alpha)P_0^{(\sigma=0)} + \overline{\mu}(\alpha)\overline{q}\Phi_L]d\Omega(\alpha) + \int_{\alpha_1}^{1}P_0^{(\sigma=0)}d\Omega(\alpha)}{\int_0^{\alpha_0}\overline{\mu}(\alpha)\overline{q}d\Omega(\alpha) + \int_{\alpha_0}^{\alpha_1}[\mu(\alpha)q + \overline{q}]d\Omega(\alpha) + \int_{\alpha_1}^{1}1d\Omega(\alpha)}$$ and

$$\Pi^{\sigma\neq 0} = \frac{\int_0^{\alpha_0}[\mu(\alpha)P_0 + \overline{\mu}(\alpha)qp_1^H]d\Omega(\alpha) + \int_{\alpha_0}^{\alpha_1}\overline{\mu}(\alpha)qp_1^H d\Omega(\alpha)}{\int_0^{\alpha_0}[\mu(\alpha) + \overline{\mu}(\alpha)q]d\Omega(\alpha) + \int_{\alpha_0}^{\alpha_1}\overline{\mu}(\alpha)qd\Omega(\alpha)},$$

where $P_0^{(\sigma=0)} = q\dfrac{\Phi_H + \Phi_L}{2} + \overline{q}\Phi_L$ and $P_0^{(\sigma\neq 0)} = (q\dfrac{\Phi_H + \Phi_L}{2} + \overline{q}\Phi_L)\dfrac{\gamma+\theta^2-2\alpha\theta}{\gamma-3\alpha^2\gamma+2\alpha\theta}$.

Proposition 6, Part 2: Allowing for self-selection across α, equal splitting is associated with lower performance, i.e., $\Pi^{\sigma=0} < \Pi^{\sigma\neq 0}$.

Proof: We note that $\Pi^{\sigma=0} < P_0^{(\sigma=0)} < \Pi^{\sigma\neq 0}$ since $\Phi_L < P_0^{(\sigma=0)}$, $p_1^H > P_0^{(\sigma=0)}$ and $P_0^{(\sigma\neq 0)} > P_0^{(\sigma=0)}$.□

4.3 Proposition 7

To examine the causal effect of equal splitting on performance, we consider random variation in two variables, r and k. Proposition 4 shows that sufficiently high or low values of r lead to unequal splitting. Proposition 3 implies that higher realizations of discovery costs k lead to a higher probability of equal splitting.

Proposition 7: Consider a given α and a given realization of skills $j = \{H, L\}$.

Part 1: There are two cases where the realization of founder resource asymmetries r induces teams to switch between equal and unequal splitting. For symmetrically skilled teams ($j = L$), an equal split leads to higher performance. For asymmetrically skilled teams ($j = H$), with sufficiently high values of α, an equal split can lead to higher or lower performance.

Part 2: There are two cases where the realization of discovery costs k induces teams to switch between equal and unequal splitting. The first case is for $\alpha < \alpha_0$ and $j = L$, in which case performance is higher under equal splitting (with discovery) than under unequal splitting (without discovery). The second case is for $\alpha \in (\alpha_0, \alpha_1)$ and $j = H$, in which case performance is higher under unequal splitting (with discovery) than under equal splitting (without discovery).

Proof:

Proof of part 1: We know from the proof of Proposition 4 that variation in r causes changes from $\sigma = 0$ to $\sigma \neq 0$ under two types of circumstances. First, for symmetric teams ($j = L$) with discovery ($d = 1$), we get $\sigma \neq 0$ for either $r > \overline{r}_1^L$ or $r < \underline{r}_1^L$. Yet for symmetric skills, p achieves its global maximum at $\sigma = 0$. Thus a random allocation to unequal splitting reduces p. Second, for asymmetric teams ($j = H$) with sufficiently high values of α (i.e., $\alpha > \alpha_0$ without discovery, $\alpha > \alpha_1$with discovery), $r > \overline{r}_1^H$ generates $\sigma > 0$, $r \in (\underline{r}_1^H, \overline{r}_1^H)$ generates $\sigma = 0$, and $r < \underline{r}_1^H$ generates $\sigma < 0$. Relative to the base case of $\sigma = 0$, sufficiently large values of r improve performance, whereas sufficiently negative values of r lower performance. This is because for $j = H$, p is increasing in σ in a range around $\sigma = 0$.

Proof of part 2: Consider now cases where discovery costs randomly allocate teams into equal versus unequal splitting. From Figure 1 in the main text we immediately note that this can never happen in the high OIA region, nor can it happen in the medium OIA region with symmetric teams, or the low OIA region with asymmetric teams. We are thus left to examine the effects in the medium OIA region with asymmetric teams, and the low OIA region with symmetric teams. For $\alpha < \alpha_0$ and $j = L$, sufficiently low k leads to discovery, $\sigma = 0$ and $p_1^L = \Phi^L$ whereas high k leads to no discovery, $\sigma > 0$ and $p_0^L = \Phi_L \dfrac{\gamma + \alpha\theta}{(1 - 3\alpha^2)\gamma + 2\alpha\theta}$. Performance under equal splitting is higher, because $p_1^L > p_0^L$. For $\alpha \in (\alpha_0, \alpha_1)$ and $j = H$, sufficiently glow k leads to discovery, $\sigma > 0$ and $p_1^H = \dfrac{\Phi_H + \Phi_L}{2} \dfrac{1 - 2\alpha\theta + \theta^2}{1 - 3\alpha^2 + 2\alpha\theta}$ whereas high k leads to no discovery, $\sigma = 0$ and $p_0^H = \dfrac{\Phi_A + \Phi_L}{2}$. Performance under equal splitting is lower, because $p_0^H < p_1^H$.$\square$

5 Model Extensions

5.1 Two-sided skill uncertainty

Consider a model with two-sided skill uncertainty. Each founder has a distinct probability of success, given by q_A and q_B. W.l.o.g. we assume $1 \geq q_A \geq q_B \geq 0$. Using obvious notation we have four skill states, denoted by $j = HH, HL, LH, LL$, and their associated state probability by q^j. At the private effort stage we have $v_i^j = p^j w_i - c_i$ where $p^j = \phi_A^j e_A + \phi_B^j e_B$. At the initial contracting stage we have $u_i = \sum_j q^j v_i^j - dk/2$. Using the standard first order conditions for optimal efforts, we obtain a.t. the following utilities: $v^{HH} = \Phi_L(\frac{1}{2}\Phi w_A^2 + \frac{1}{2}\Phi w_B^2 + 2\Phi w_A w_B)$, $v^{HL} = \Phi_L(\frac{1}{2}\Phi w_A^2 + \frac{1}{2}w_B^2 + (\Phi + 1)w_A w_B)$, $v^{LH} = \Phi_L(\frac{1}{2}w_A^2 + \frac{1}{2}\Phi w_B^2 + (\Phi + 1)w_A w_B)$ and $v^{LL} = \Phi_L(\frac{1}{2}w_A^2 + \frac{1}{2}w_B^2 + 2w_A w_B)$.

With discovery the two-sided model behaves just like the one-sided model, the only addition being for $j = LH$, where we obtain the same results as for $j = HL$, except that the roles of A and B are switched. Let us now derive the optimal contracts without discovery by maximizing u_0 w.r.t. σ. A.t. we obtain $u_0 = \Phi_L(w_A^2 Q_1 + w_B^2 Q_2 + w_A w_B Q_3)$ where $Q_1 = \frac{1}{2}(q_A \Phi + q_B)$, $Q_2 = \frac{1}{2}(q_B \Phi + q_A)$ and $Q_3 = (q_A + q_B)(\Phi - 1) + 2$. The first order condition for σ is given by $\frac{du_0}{d\sigma} = \Phi_L(\frac{dw_A^2}{d\sigma}Q_1 + \frac{dw_B^2}{d\sigma}Q_2 + \frac{dw_A w_B}{d\sigma}Q_3) = 0$. A.t. we obtain $\sigma^* = \frac{(1-\alpha)Q_1 - (1+\alpha)Q_2 - \alpha Q_3}{-\Delta}$ where $\Delta = 2(1-\alpha)^2 Q_1 + 2(1+\alpha)^2 Q_2 - 2(1-\alpha^2)Q_3 < 0$. We find that $\sigma^* = 0$ at $\alpha_0 = \frac{Q_1 - Q_2}{Q_1 + Q_2 + Q_3}$, which a.t. yields $\alpha_0 = \frac{(q_A - q_B)(\Phi - 1)}{(q_A + q_B)(3\Phi - 1) + 4}$. This says that α_0 is a function of the difference in high skills probabilities $(q_A - q_B)$. In the symmetric case of $q_A = q_B$ we find $\alpha_0 = 0$.

5.2 Continuous skill distribution

Instead of assuming that there are only two states, consider a model extension with a continuum of states. For tractability we revert back to one-sided uncertainty, and assume that upon discovery A's has skills $\widetilde{\phi}$ that are continuously distributed over the internal $[\phi_L, \phi_H]$. Symmetry of skills between A and B thus becomes a measure zero event. Our main question is whether equal splitting still occurs when partners are always asymmetric.

Our focus here is to show that equal splitting is optimal for a non-degenerate range of parameter values of $\widetilde{\phi}$. To see this, consider again the critical value $\alpha_1 = \frac{\theta}{3} = \frac{1}{3}\frac{\phi_H - \phi_L}{\phi_H + \phi_L}$ from the base model. For the continuous skills case, we define for each $\widetilde{\phi}$ a critical value

$\alpha_1(\widetilde{\phi}) = \frac{1}{3}\frac{\widetilde{\phi} - \phi_L}{\widetilde{\phi} + \phi_L}$, so that equal splitting is optimal whenever $\alpha > \alpha_1(\widetilde{\phi})$. We note that $\alpha_1(\widetilde{\phi}) = 0$ at $\widetilde{\phi} = \phi_L$, and that $\alpha_1(\widetilde{\phi})$ is strictly increasing in $\widetilde{\phi}$. Thus, for every $\alpha > 0$, there exists some $\widehat{\phi} > \phi_L$ so that $\alpha_1(\widehat{\phi}) = \alpha$. And for all $\widetilde{\phi} < \widehat{\phi}$ we have $\alpha_1(\widetilde{\phi}) < \alpha$. But this is precisely the condition for equal splitting to be optimal. We have thus shown that for every $\alpha > 0$, there exists a lower parameter range $\widetilde{\phi} \in [\phi_L, \widehat{\phi}]$ for which equal splitting is optimal. This helps to justify our simplifying symmetry assumption in the base model.

5.3 Utility inequality aversion

We briefly consider the model with utility inequality aversion (UIA). Without discovery we assume that preferences are given by $U_A = u_A - \beta|u_A - u_B|$ and $U_B = u_B - \beta|u_A - u_B|$, where u_A and u_B are the same as before, except that we assume no OIA and set $\alpha = 0$. With discovery we assume that preferences are given by $V_A^j = v_A^j - \beta|v_A^j - v_B^j|$ and $V_B^j = v_B^j - \beta|v_A^j - v_B^j|$, $j = \{H, L\}$.

A complete derivation of the model is beyond the scope of this paper, we only focus on the question when equal splitting occurs. In the model with discovery it is easy to see that equal splitting is optimal whenever the partners are symmetric, i.e., when $j = L$. However, for the asymmetric case of $j = H$, we examine the maximization of $V^H = V_A^H + V_B^H = (v_A^H + v_B^H) - 2\beta|v_A^H - v_B^H|$. From the OIA model we know that with optimal efforts we get $v_A^H = \frac{1}{2}\Phi_H w_A^2 + \Phi_L w_A w_B$ and $v_B^H = \frac{1}{2}\Phi_L w_B^2 + \Phi_H w_A w_B$. For the UIA model we simply use $\alpha = 0$, so that $w_A = \frac{1}{2} + \sigma$ and $w_B = \frac{1}{2} - \sigma$. We obtain a.t. $v_A^H + v_B^H = \frac{\Phi_H + \Phi_L}{2}[\frac{3}{4} + \sigma\theta - \sigma^2]$ and $v_A^H - v_B^H = \frac{\Phi_H + \Phi_L}{2}[-\frac{\theta}{4} + \sigma + 3\sigma^2\theta]$. We note that near $\sigma = 0$ we have $v_A^H - v_B^H = \frac{\Phi_H + \Phi_L}{2}[-\frac{\theta}{4}] < 0$, so that $|v_A^H - v_B^H| = -(v_A^H - v_B^H)$. Thus, near $\sigma = 0$ we have $V^H = \frac{\Phi_H + \Phi_L}{2}[\frac{3}{4} + \sigma\theta - \sigma^2] + 2\beta\frac{\Phi_H + \Phi_L}{2}[-\frac{\theta}{4} + \sigma + 3\sigma^2\theta]$ and $\frac{dV^H}{d\sigma} = \frac{\Phi_H + \Phi_L}{2}[\theta - 2\sigma + 2\beta + 6\beta\sigma\theta]$. Evaluating this at $\sigma = 0$ yields $\frac{dV^H}{d\sigma} = \frac{\Phi_H + \Phi_L}{2}[\theta + 2\beta] > 0$. This says that at the first order condition for σ is never satisfied at $\sigma = 0$. It follows that equal splitting is never optimal when asymmetric types have been discovered. This results holds for all values of $\phi_H > \phi_L$.

Next we consider the model without discovery, where we maximize $U = U_A + U_B$. Using $U_A = u_A - \beta|u_A - u_B|$ and $U_B = u_B - \beta|u_A - u_B|$ this means maximizing $U = (u_A + u_B) - 2\beta|u_A - u_B|$. Using the same reasoning as above we have $v_A^j = \frac{1}{2}\Phi_j w_A^2 + \Phi_L w_A w_B$ and $v_B^j = \frac{1}{2}\Phi_L w_B^2 + \Phi_j w_A w_B$ with $w_A = \frac{1}{2} + \sigma$ and $w_B = \frac{1}{2} - \sigma$, which yields $v_A^j =$

$\frac{1}{2}\Phi_j(\frac{1}{4}+\sigma+\sigma^2)+\Phi_L(\frac{1}{4}-\sigma^2)$ and $v_B^j=\frac{1}{2}\Phi_L(\frac{1}{4}-\sigma+\sigma^2)+\Phi_j(\frac{1}{4}-\sigma^2)$. We obtain a.t. $u_A = q\frac{\Phi_L}{2}[\Phi(\frac{1}{4}+\sigma+\sigma^2)+(\frac{1}{2}-2\sigma^2)] + \overline{q}\frac{\Phi_L}{2}(\frac{3}{4}+\sigma-\sigma^2)$ and $u_B = q\frac{\Phi_L}{2}[(\frac{1}{4}-\sigma+\sigma^2)+\Phi(\frac{1}{2}-2\sigma^2)] + \overline{q}\frac{\Phi_L}{2}(\frac{3}{4}-\sigma-\sigma^2)$. Thus we get a.t. $u_A+u_B=\chi_H[\frac{3}{4}+\sigma\theta-\sigma^2]+\chi_L(\frac{3}{4}-\sigma^2)$ and $u_A-u_B=\chi_H(-\frac{1}{4}\theta+\sigma+3\sigma^2\theta)+\chi_L\sigma$. At $\sigma=0$ we have $u_A-u_B=\chi_H(-\frac{1}{4}\theta)<0$, so that near $\sigma=0$ we have $|u_A-u_B|=-(u_A-u_B)$ and thus $U=\chi_H[\frac{3}{4}+\sigma\theta-\sigma^2]+\chi_L(\frac{3}{4}-\sigma^2)+2\beta(\chi_H(-\frac{1}{4}\theta+\sigma+3\sigma^2\theta)+\chi_L\sigma)$. Thus $\frac{dU}{d\sigma}=\chi_H[\theta-2\sigma]+\chi_L(-2\sigma) + 2\beta[(\chi_H(1+6\sigma\theta)+\chi_L)]$. Evaluating this at $\sigma=0$ we obtain $\frac{dU}{d\sigma}=\chi_H\theta+2\beta(\chi_H+\chi_L)>0$. This says that the first-order condition for the optimal σ is never satisfied at $\sigma=0$. It follows that equal splitting is never optimal in the model without discovery.

Overall we notice that in our base model with only two skill levels $\phi_A \in \{\phi_L, \phi_H\}$, equal splitting occurs with discovery and symmetric skills, but never with discovery and asymmetric skills, nor without discovery. It follows that the UIA model predicts a positive relationship between discovery and equal splitting. This stands in direct contrast to the OIA model, where Proposition 3 predicts a negative relationship between discovery and equal splitting. In the empirical analysis we find evidence for a negative relationship.

The prediction from the UIA model depends on the assumption that there is a positive probability that the two founders are perfectly symmetric with discovery. In the model extension with a continuous skill distribution $\widetilde{\phi} \in [\phi_L, \phi_H]$, we found in the OIA model that the prediction of equal splits was robust to small asymmetries (as reflected in the range $[\phi_L, \widehat{\phi}]$). This is not true in the UIA model. The proof above shows that even a small skill asymmetry immediately leads to unequal splits. It follows that equal splitting never occurs in the continuous skill model (strictly speaking, it is a measure zero event). This prediction is clearly inconsistent with the evidence.

Overall we note that the UIA and OIA models generate different predictions, and that the UIA predictions are easily rejected by the data. From a theoretical perspective it is also worth noting that all the differences between the UIA and OIA models are driven by founder asymmetries. Much of the prior economic theory literature on fairness assumes ex-ante symmetry, which means that it is not set up to uncover differences between UIA and OIA.

5.4 Asymmetric OIA preferences

The analysis so far assumes the each team has a single OIA parameter α (even though different teams can have different values of α). In this section we show how to relax this assumption by introducing two related extensions. First, the two founders may have different degrees of OIA; using obvious notation, this means $\alpha_A \neq \alpha_B$. Second, any individual founder may have a different aversion to relative gains and losses. Using obvious notation, we assume that α_i^+ pertains to the aversion of getting more (think 'guilt') and α_i^- pertains to the aversion of getting less (think 'resentment').

Let us focus on the choice between $\sigma = 0$ and $\sigma > 0$, in which case the two parameters that matter are α_A^+ and α_B^-, generating payoffs $w_A = (\frac{1}{2} + \sigma) - \alpha_A^+ \sigma$ and $w_B = (\frac{1}{2} - \sigma) - \alpha_B^- \sigma$. The model behaves the same with discovery and symmetric skills. For the case of discovery with asymmetric skills we obtain the optimal choice of σ from maximizing $v_H = v_A^H + v_B^H = \frac{1}{2}\Phi_H w_A^2 + \Phi_L w_A w_B + \frac{1}{2}\Phi_L w_B^2 + \Phi_H w_A w_B$. From the first order condition we obtain a.t. $\sigma = \frac{(\Phi+1)[\frac{1}{2}\alpha_A^+ + \frac{1}{2}\alpha_B^-] - \frac{1}{2}\Phi(1-\alpha_A^+) + \frac{1}{2}(1+\alpha_B^-)}{[\Phi(1-\alpha_A^+)^2 + (1+\alpha_B^-)^2 - 2(\Phi+1)(1-\alpha_A^+)(1+\alpha_B^-)]}$

Is is useful to rewrite the coefficients as $\alpha_A^+ = \alpha$ and $\alpha_B^- = \alpha + \varepsilon$, where ε could be positive or negative. Setting $\sigma = 0$, we find a.t. $\alpha_1 = \frac{\theta}{3} - \frac{1}{3}\frac{\Phi+2}{\Phi+1}\varepsilon$. Again we note that equal splitting is optimal whenever $\alpha > \alpha_1$. For $\varepsilon = 0$, α_1 reverts to the familiar formula. For $\varepsilon > 0$, B has even stronger OIA than A. In this case α_1 decreases, so that the range of equal splitting increases. The opposite applies to the case of $\varepsilon < 0$. The analysis without discovery follows a very similar logic.

5.5 Transaction costs

We briefly outline an alternative model with transaction costs of unequal splitting. We assume that unequal splitting generates a transaction cost $\tau > 0$ that is jointly incurred by the founders. Let us focus on the asymmetric case ($j = H$) with discovery. The joint utility is given by $v_{\sigma=0}^H = \frac{3}{8}(\Phi_H + \Phi_L)$ for an equal split ($\sigma^* = 0$) and $v_{\sigma\neq 0}^H = \frac{\Phi_H + \Phi_L}{2}[\frac{3}{4} + \frac{1}{4}\frac{(\theta - 3\alpha)^2}{1 + 2\alpha\theta - 3\alpha^2}]$ for an unequal split ($\sigma^* \neq 0$) . In the transaction cost model the unequal split is chosen whenever $v_{\sigma\neq 0}^H > v_{\sigma=0}^H + \tau$ which yields a.t. $z(\alpha) \equiv \frac{\Phi_H + \Phi_L}{8}\frac{(\theta - 3\alpha)^2}{1 + 2\alpha\theta - 3\alpha^2} > \tau$. It is easy to verify that $\frac{dz}{d\alpha}\frac{(\theta - 3\alpha)^2}{1 + 2\alpha\theta - 3\alpha^2} < 0$ and $z(\alpha_1) = 0$. Thus there exists $\widehat{\alpha} < \alpha_1$, so that for all $\alpha < \widehat{\alpha}$ we have $v_{\sigma\neq 0}^H > v_{\sigma=0}^H + \tau$, making $\sigma^* > 0$ optimal; and for all $\alpha > \widehat{\alpha}$ we have $v_{\sigma\neq 0}^H < v_{\sigma=0}^H + \tau$, making $\sigma^* = 0$ optimal. Thus there exists a parameter range $\alpha \in [\widehat{\alpha}, \alpha_1]$ where our base model predicts unequal sharing, but the transaction cost model predicts equal sharing. Recall that the optimal σ^* is decreasing in α, with $\sigma^*(\alpha_1) = 0$, thus $\sigma(\widehat{\alpha}) > 0$. It follows that there exists a range of equity premiums $\sigma \in (0, \sigma(\widehat{\alpha}))$ (we can think of these are 'near-equal' splits) that occur in our base model, but that never occur in the model with transaction costs. Note also that this insight does not depend on the existence of OIA. We can set $\alpha = 0$, and use variation in ϕ_H to obtain the same results.

Finally, in the main text we mention that better teams self-select into unequal splitting. In our base model we normalize returns to 1, standard calculations reveal that if returns are given more generally by Π, then the quadratic effort cost model generates utilities than are multiplied by a factor $\pi = \Pi^2$. We can rewrite the condition for unequal splitting as $\pi v^H_{\sigma \neq 0} > \pi v^H_{\sigma = 0} + \tau \Leftrightarrow v^H_{\sigma \neq 0} > v^H_{\sigma = 0} + \frac{\tau}{\pi}$. This suggests that transaction costs need to be evaluated relative to expected returns of the venture. The threshold for incurring transactions costs of unequal splitting is decreasing in π, so that better teams are more likely chose unequal splits.

6 Variable list

The table below explains the most important variables used in the analysis.

Variable	Meaning
$i = A, B$	Founder A and B
q	Probability that A has high skills is q
$\phi_j, j = \{H, L\}$	A's skills ϕ_j are high ($j = H$) or low ($j = L$)
R	Total founder resources
r	Parameter for asymmetry of founder resources
$k, K(k)$	Discovery costs k with distribution $K(k)$
$d = \{0, 1\}$	Discovery decision: $d = 1$ means discovery undertaken
σ_d^j	A's share premium
c, e_i, ψ	Private effort e_i has costs $c(e_i) = 0.5\psi e_i^2$
p	Probability of success, given by $p = \phi_A e_A + \phi_B e_B$
P_d	Expected probability of success $P_d = qp_d^H + \overline{q}p_d^L$
α, Ω	Outcome inequality aversion (OIA) parameter α, with distribution Ω
w_i	Utility of i at outcome stage
v_i^j	Utility of i after discovery of state $j = \{H, L\}$
u_i	Utility of i before discovery
v^j, u	Without subscript i, utilities denoted joint utilities
$\Phi_j = \frac{\phi_j^2}{\psi}, \Phi = \frac{\Phi_H}{\Phi_L}, \theta = \frac{\Phi - 1}{\Phi + 1}$	Transformed skill parameters
$\chi = \frac{\overline{q}}{q} \frac{2}{1 + \Phi}, \gamma = 1 + \chi$	Useful transformations
μ	Probability of no discovery
$\lambda(\alpha)$	Probability of equal splitting, conditional on α
Λ	Probability of equal splitting (unconditional)
$\delta(\alpha)$	Performance premium of unequal splitting