Multi-Project Collaborations
Online Appendix

Charles Angelucci and Roi Orzach*

August 5, 2026

1 Proof of Proposition 2

Proof of Proposition 2. Recall that after a deviation in period t, players set $P_i^t = \emptyset$ and $b_i^t = 0$ if not already chosen. In subsequent periods, they revert to the static equilibrium with zero transfers and no selected projects.

The proof proceeds in four steps: (i) we show that it is without loss of optimality to restrict attention to relational contracts that are surplus-maximizing following every on-path history h^t; (ii) we provide a necessary and sufficient condition for the existence of a relational contract that implements a given experimentation policy $\hat{\mathbf{P}}(\cdot)$; (iii) we show that this condition is independent of the division of surplus between the players; and (iv) we show that, for any two histories that generate the same beliefs, selecting the same continuation equilibrium is without loss of optimality.

Step 1 We show that it is without loss of optimality to restrict attention to relational contracts that are surplus-maximizing following every on-path history h^t. To see this, suppose that there exists an on-path history h^t such that the continuation equilibrium starting in period t, denoted by e^1, has lower total surplus than an alternative continuation equilibrium e^2. Thus, if we define $\mathcal{C}_i^k$ to be the continuation value to player i in equilibrium e^k, then $\sum_i \mathcal{C}_i^1 < \sum_i \mathcal{C}_i^2$. For the rest of Step 1, we omit the superscript $t-1$ in our notation, as we are solely concentrating on period $t-1$ objects.

Let us modify the players' relational contract such that play in and after period t is dictated by e^2 and the period $t-1$ $b_i(\cdot)$ transfers associated with history h^t (and, thus, corresponding to a specific realizations of $\mathbf{x}^{t-1}, \mathbf{v}^{t-1}$) are adjusted so that: (i) player 2's expected payoff following the realizations of $\mathbf{x}^{t-1}, \mathbf{v}^{t-1}$ is the same as under the original

*Angelucci: MIT Sloan School of Management, 50 Memorial Drive, Cambridge MA 02142 (e-mail: cangeluc@mit. edu); Orzach: Department of Economics, Boston University, 270 Bay State Road, Boston MA 02215 (e-mail: orzach@bu.edu).

equilibrium and (ii) player 1's expected payoff following the realizations of $\mathbf{x}^{t-1}, \mathbf{v}^{t-1}$ increases by $\sum_i \mathcal{C}_i^2 - \sum_i \mathcal{C}_i^1$. Specifically, take the vector of transfers $\mathbf{b}_1 = (b_1^1, b_2^1)$ associated with the original equilibrium and create a new vector of transfers $\mathbf{b}_2 = (b_1^2, b_2^2)$ such that:

$$\mathcal{C}_1^2 + b_2^2 - b_1^2 > \mathcal{C}_1^1 + b_2^1 - b_1^1, \tag{1}$$

$$\mathcal{C}_2^2 + b_1^2 - b_2^2 = \mathcal{C}_2^1 + b_1^1 - b_2^1. \tag{2}$$

Because $\sum_i \mathcal{C}_i^2 - \sum_i \mathcal{C}_i^1 > 0$, finding payments that satisfy $b_1^2 \leq \mathcal{C}_1^2$ and $b_2^2 \leq \mathcal{C}_2^2$ is always feasible.

Note that these changes have no impact on player 1's choices of actions made in any period $t' \leq t-1$ because all actions are observable, and hence choosing a different action from the proposed equilibrium would be labeled a defection. If defections were deterred in the original equilibrium, which had a strictly smaller continuation value for player 1, then they are also deterred in the new equilibrium. The same logic applies to player 2 since they obtain the same expected payoff in period $t-1$ (compared to the original equilibrium), and thus also have the same continuation values in all periods $t' < t-1$. Finally, note that surplus from a date 0 perspective is strictly higher under the new equilibrium.

Step 2 We show that there exists a relational contract that implements an experimentation policy $\hat{\mathbf{P}}(\cdot)$ if and only if the following inequality holds for all t and for all on-path histories $h^t \in \mathcal{H}^t$:

$$\sum_{p \in \hat{\mathbf{P}}^t} \sum_{i=1,2} \max\Big(0, c - \mathbb{E}(v_p \mathbb{1}_{x_p=i} | h^t)\Big) \leq \mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t), \tag{3}$$

where $\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t)$ is the continuation value.

To show that (3) is a necessary and sufficient condition, consider a set of transfers $b_i(\mathbf{x}^t, \mathbf{v}^t) \geq 0$ to be paid on path given a vector of realized values $\mathbf{x}^t, \mathbf{v}^t$.

Given an equilibrium experimentation policy $\mathbf{P}^t$, note that it is without loss of generality to assume that $P_1^t = P_2^t = \mathbf{P}^t$. Thus, for each player and for each $p \in \mathbf{P}^t$, the player must weakly prefer to include p in P_i^t, rather than excluding it. Let $\sigma_i(\mathbf{x}^t, \mathbf{v}^t)$ denote player i's share of $\mathcal{C}(\hat{P}(\cdot), h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)$ as a function of $\mathbf{x}^t, \mathbf{v}^t$. Hence, the condition for selecting $\mathbf{P}^t$ is:

$$\begin{aligned}\sum_{p \in \mathbf{P}^t} \max\left(c - \mathbb{E}(v_p \mathbb{1}_{x_p=i} | h^t), 0\right) \leq \mathbb{E}\Big(&b_{-i}(\mathbf{x}^t, \mathbf{v}^t) - b_i(\mathbf{x}^t, \mathbf{v}^t) \\ &+ \sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)\Big), \quad \forall i,\end{aligned} \tag{4}$$

$$b_i(\mathbf{x}^t, \mathbf{v}^t) \leq \sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t), \quad \forall \mathbf{v}^t, \forall \mathbf{x}^t, \forall i. \tag{5}$$

Expectations are taken over the project valuations realizations $\mathbf{x}^t, \mathbf{v}^t$ and $h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t$ denotes the updated history after observing $\mathbf{x}^t, \mathbf{v}^t$.[1] The first expression states that the promised transfers and the expected share of the total continuation value must be enough to prevent a player from shirking on any subset of the projects. The second expression states that each player is willing to pay the other player the necessary transfer.

To show necessity: Note that since Equation (4) must hold for a fixed i, the inequality also holds summing over all i. Further, all transfers cancel out when summing over i. Finally, by definition, $\mathbb{E}(\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)) = \mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t)$. Hence, we are left with Equation (3).

To show sufficiency: Note that for any set of transfers that satisfies Equation (4) (5), there exists an alternative set of transfers for which, one of b_1 or b_2 is equal to zero for all $\mathbf{x}^t, \mathbf{v}^t$. If both were positive, decreasing both by the same expected amount would leave Equation (4) the same and add slack to Equation (5).

We will show this result in two substeps.

SubStep 1: We show it is necessary and sufficient to replace Equation (5) by its expectation. This new expression is as follows:

$$\mathbb{E}(b_i(\mathbf{x}^t, \mathbf{v}^t)) \leq \mathbb{E}\Big(\sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)\Big) \quad \forall i. \tag{6}$$

We first show that if there is a solution to Equations (6) and (4), then there exists a solution to Equations (5) and (4). By the discussion above, it is without loss to assume that for one of the players $b_i(\cdot) = 0$, so denote by i the player the player making transfers. For player $-i$, Equation (5) holds trivially. Further, for any bonuses satisfying Equation (6) one can construct an alternative set of transfers for player i as follows

Define:

$$b_i'(\mathbf{x}^t, \mathbf{v}^t) = \frac{\mathbb{E}(b_i(\mathbf{x}^t, \mathbf{v}^t))}{\mathbb{E}\Big(\sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t)\Big)}\sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t) \tag{7}$$

Since Equation (6) holds, the term in the expectation of Equation (7) is positive and thus Equation (5) holds for all realizations of $\mathbf{x}^t, \mathbf{v}^t$. Finally, $\mathbb{E}(b_i'(\mathbf{x}^t, \mathbf{v}^t)) = \mathbb{E}(b_i(\mathbf{x}^t, \mathbf{v}^t))$ so Equation (6) continues to hold.

SubStep 2: Using substep 1, it suffices to show that Equation (3) implies a solution to Equations (4) and (6). To simplify all the notation with expectations, Equation (4) can be

[1]The history also includes the project selections, and both the upfront and end-of-period transfers. However, for notational convenience we only include the realized valuations as every other object can be inferred on path from the realized valuations.

re-expressed as:

$$\beta_i - \gamma_i \leq (\tilde{b}_{-i} - \tilde{b}_i), \tag{8}$$

where $\tilde{b}_i$ is the expected transfer from i to $-i$, $\beta_i = \sum_{p \in \mathbf{P}^t} \max\left(0, c - \mathbb{E}(v_p \mathbb{1}_{x_p=i} | h^t)\right)$, and $\gamma_i = \mathbb{E}(\sigma_i(\mathbf{x}^t, \mathbf{v}^t)\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t \sqcup \mathbf{x}^t \sqcup \mathbf{v}^t))$. Equation (6) can thus be re-written as:

$$\tilde{b}_i \leq \gamma_i. \tag{9}$$

Rearranging Equation (3) implies $\sum_i (\beta_i - \gamma_i) \leq 0$. One can now show that $\tilde{b}_i = \max(0, \beta_{-i} - \gamma_{-i})$ satisfies Equation (9). Further, Equation (8) holds because:

$$\begin{aligned} \beta_i - \gamma_i \leq \max(0, \beta_i - \gamma_i) - \max(0, \beta_{-i} - \gamma_{-i}) \qquad (10) \\ \iff \max(0, \beta_{-i} - \gamma_{-i}) - \max(0, \gamma_i - \beta_i) \leq 0 \qquad (11) \\ \Longleftarrow \sum_i (\beta_i - \gamma_i) \leq 0, \qquad (12) \end{aligned}$$

where the final step follows from noting that $\beta_1 - \gamma_1$ and $\beta_2 - \gamma_2$ cannot both be positive and analyzing the remaining three cases based on the signs of $\beta_i - \gamma_i$.

Finally, Equation (9) reduces to

$$\begin{aligned} \max(0, \beta_{-i} - \gamma_{-i}) \leq \gamma_i \Longleftarrow \beta_{-i} - \gamma_{-i} \leq \gamma_i \qquad (13) \\ \Longleftarrow \sum_i (\beta_i - \gamma_i) \leq 0, \qquad (14) \end{aligned}$$

where the final implication is due to β_i being weakly positive.

Step 3: We show that any relational contract that implements a given experimentation policy can be replaced by an alternative relational contract that implements the same experimentation policy and yields no surplus to player 2. First, note that the way the players share their continuation value does not affect Equation (2) from the main text. Hence, for any period t where player 2's expected payoff is positive, $w_{2,1}$ can be increased until player 2's expected payoff is zero. Player 2 is willing to make this transfer because not doing so would be seen as a deviation, resulting in a payoff of 0 for player 2.

Step 4: We now show that, for any two histories h_1^t and $h_2^{t'}$ that generate the same beliefs μ, selecting the same continuation equilibrium is without loss of optimality. Take a relational contract r that is surplus-maximizing at all on-path histories and has two histories h_1^t and $h_2^{t'}$ prescribing different (surplus-maximizing) continuation equilibria under the same beliefs μ. Recall from Step 3 that one can consider relational contracts in which player 2 obtains an expected payoff equal to 0 in every period. In this case, since the two continuation equilibria are both sequentially optimal and both give all the surplus to player 1, switching from one continuation equilibrium to the other does not change the players' incentives as both prescribe the exact same payoffs to the players. Hence, when focusing on relational contracts that specify the same continuation equilibrium following histories that induce the same beliefs, one can replace $\mathcal{C}(\hat{\mathbf{P}}(\cdot), h^t)$ with $\mathcal{C}(\hat{\mathbf{P}}(\cdot), \mu^t)$. $\square$

2 Example where $m^* > 1$

We consider a three-point support of benefits: $\{0, \underline{v}, \bar{v}\}$. We also assume $\mathbb{E}(v_p) = 2c$, $\alpha = 1$, and $m = 2$ when listing sufficient inequalities for $\delta^* = \bar{\delta}$ to hold. The following inequalities ensure that there exists a δ such that $|\mathbf{P}^1| = 1$ is not feasible but $|\mathbf{P}^1| = 2$ is, which is sufficient to show $\delta^* = \bar{\delta}$:

$$\frac{\underline{v} - 2c}{1-\delta} > \frac{\delta \Pr(\bar{v})}{1 - \delta(1 - \Pr(\bar{v}))}(\bar{v} - 2c)\frac{1}{1-\delta} \tag{15}$$

$$2c < \frac{\delta}{1-\delta}(\bar{v} - 2c) \tag{16}$$

$$c > \frac{\delta}{1-\delta}(\underline{v} - 2c) \tag{17}$$

$$2c > \frac{\delta}{1-\delta}(\underline{v} - 2c) + \frac{\delta \Pr(\bar{v})}{1 - \delta(1 - \Pr(\bar{v}))}(\bar{v} - 2c)\frac{1}{1-\delta} \tag{18}$$

$$c > \frac{\delta \Pr(\bar{v})}{1 - \delta(1 - \Pr(\bar{v}))}(\bar{v} - 2c)\frac{1}{1-\delta} + \frac{\delta \Pr(\bar{v})}{1 - \delta(1 - \Pr(\bar{v}))}\mathrm{npv}^0 \tag{19}$$

$$\begin{aligned} 2c < {} & \frac{2\Pr(\bar{v})(1 - \Pr(\bar{v}))}{1 - \delta\big(1 - 2\Pr(\bar{v})(1 - \Pr(\bar{v}))\big)}\Big((\bar{v} - 2c)\frac{\delta}{1-\delta} + \mathrm{npv}^0\Big) \\ & + \frac{\Pr(\bar{v})^2}{1 - \delta\big(1 - \Pr(\bar{v})^2\big)}\frac{\delta}{1-\delta}2(\bar{v} - 2c) \end{aligned} \tag{20}$$

$$\mathrm{npv}^0 = \frac{\delta}{1-\delta}\frac{\Big(\Pr(\bar{v}) + \Pr(\underline{v})\Big)\Big(\frac{\Pr(\bar{v})\bar{v}}{\Pr(\bar{v}) + \Pr(\underline{v})} + \frac{\Pr(\underline{v})\underline{v}}{\Pr(\bar{v}) + \Pr(\underline{v})} - 2c\Big)}{1 - \delta(1 - \Pr(\bar{v}) - \Pr(\underline{v}))} \tag{21}$$

Inequality (15) ensures that $v^0 \leq \underline{v}$. Inequality (16) ensures that $\bar{v}, 0$ satisfies Inequality 5 from the main text. Inequality (17) ensures that the players are unable to exploit a project worth $\underline{v}$ in isolation and, by extension, cannot jointly exploit two such projects. Inequality (18) implies that the players cannot exploit a project worth $\underline{v}$ while exploring in the additional domain. Inequality (19) implies that the players would be unable to begin exploring if the players began their exploration on one domain and, upon finding a project worth $\bar{v}$, started exploring the additional domain.[2] The continuation value under this experimentation policy is bounded above by npv^0: the net-present value of a single domain under the first-best experimentation policy, the value of which is stated in (21). Finally, Inequality (20) is a necessary condition for the initially maximal policy to be feasible at date 1. This condition is necessary, but not sufficient, as the continuation value is computed assuming that if the players discover one project worth $\bar{v}$ before doing so on the other domain, the highest-valued project on the other domain is zero. We use Mathematica to show that these constraints jointly hold strictly.[3]

3 Extensions

In this section we provide the proofs for the extensions outlined in Section 5.4 of the text.

3.1 Risky Collaborations

We make the following simplifications: $\alpha = 1$, $\mathbb{E}(v_p) = 2c$, and $m = 2$. Let the distribution of benefits on domain 1 be as in the main analysis. On domain 2 there exists only one project with value $v_r \in \{0, v\}$, where $\Pr(v_r = v) = q$ and $v \geq 2\tilde{v}$, ensuring that upon discovery of a project with value v the players can implement the first-best policy on domain 1.

Proposition A1 *For an open set of parameter values: (i) the players explore the domain* 2 *project after having first discovered a project with value exceeding $\tilde{v}$ on domain* 1*, (ii) despite being able to explore both domains in period* 1*.*

Proof of Proposition A1. We now provide the necessary conditions for the statement in the proposition to hold:

$$c \leq q\frac{\delta}{1-\delta}(v - 2c) \tag{22}$$

$$c > \frac{\Pr(v_p \geq \tilde{v})\delta}{1 - \delta \cdot \Pr(v_p < \tilde{v})} \frac{1}{1-\delta} \mathbb{E}(v_p - 2c | v_p \geq \tilde{v}) \tag{23}$$

$$c < \frac{\Pr(v_p \geq \tilde{v})\delta}{1 - \delta \cdot \Pr(v_p < \tilde{v})} \Big(\frac{1}{1-\delta} \mathbb{E}(v_p - 2c | v_p \geq \tilde{v}) + qv - 2c + q\frac{\delta}{1-\delta}(v - 2c) \Big) \tag{24}$$

These conditions state, respectively, that domain 2 could be started on its own, that domain 1 cannot be started on its own (or, equivalently, with knowledge that the domain 2 project

[2] Further, it will never be optimal to explore the additional domain upon finding $\underline{v}$, since such a project cannot be exploited.

[3] Code available upon request.

is worth zero), and that domain 1 can be started when anticipating that domain 2 is started after discovering a project with value at least $\tilde{v}$. These conditions imply that the optimal experimentation policy takes one of two forms: (1) Explore projects on both domains, and if the risky project fails, then permanently idle on domain 1 if a project with value $\tilde{v}$ is not discovered in the first period. If the risky project has value v, then the experimentation policy becomes efficient on domain 1. (2) Explore domain 1 and only explore domain 2 once finding a project with value at least $\tilde{v}$. As policies (1) and (2) are both feasible, the players choose the policy with a higher net present value. Exploring both domains implies a surplus of:

$$\begin{aligned} &qv - 2c + q\frac{\delta}{1-\delta}(v - 2c) \\ &+ (1-q)\frac{\delta}{1-\delta}\mathbb{E}((v_p - 2c)\mathbb{1}_{v_p \geq \tilde{v}}) + q\frac{\Pr(v_p > v^0)\delta}{1 - \delta \cdot \Pr(v_p < \tilde{v})}\frac{1}{1-\delta}\mathbb{E}(v_p - 2c | v_p \geq v^0), \end{aligned} \tag{25}$$

where the first line corresponds to the surplus from domain 2 and the second line corresponds to the surplus from domain 1. Further, the surplus from exploring domain 1 and later exploring domain 2 is exactly Equation (24) as the expected surplus in period 1 is equal to zero. Let us hold fixed the surplus from domain 2 and δ, but let us consider varying q while v is pinned down to maintain the same expected surplus. When $q = 1$, there is no uncertainty and the surplus from exploring both domains at once can be shown to be strictly greater than delaying. When domain 2 is explored after domain 1, changing q but holding the expected surplus from domain 2 fixed has no effect on the total expected surplus. In contrast, when exploring both domains at once, a decrease in q increases the probability that domain 1 is never explored. As q continues to decrease, the surplus from exploring both domains at once converges to the expected surplus from domain 2 plus the probability that the first drawn project in domain 1 exceeds $\tilde{v}$, multiplied by the profitability of exploiting such a project. In contrast, the surplus from a sequential approach converges to the surplus obtained with certainty in domain 1, plus the expected profitability of domain 2 when explored after identifying a suitable project in domain 1. One can see that the players may delay domain 2, even if it is more profitable than domain 1, so as to prevent permanently idling on domain 1. □

3.2 Symmetric Benefits

Let us first consider a case where in one domain the benefits are symmetric but in the other domain benefits are asymmetric with $c - (1-\alpha)\mathbb{E}(v_p) > 0$. Namely, for domain 1 all projects equally benefit both parties but domain 2 is asymmetric ex-post (as in our main analysis) and sufficiently asymmetric ex-ante to prevent exploration as a Nash Equilibrium of the stage game.

Proposition A2 *Let δ^* and $\overline{\delta}$ be defined as in the main analysis. In this extension, $\delta^* < \overline{\delta}$.*

Proof of Proposition A2. Observe that $\delta^* = 0$ because an experimentation policy consisting of (i) exploration in the symmetric benefits domain and (ii) expanding scope if a project exceeding $\tilde{v}$ is discovered in the first domain is always feasible. Further, $\bar{\delta} > 0$ because $\mathcal{C}(\cdot) \to 0$ as $\delta \to 0$, implying that the players cannot explore the asymmetric benefits domain in period 1. □

Let us now consider the case where α, the probability that the benefits go to player 1, differs across the domains. Denote by α_j, the corresponding probability for domain j. For simplicity, let us assume $\alpha_1 = 1/2$ and $\alpha_2 = 1$.

Proposition A3 *Let δ^* and $\overline{\delta}$ be defined as in the main analysis. In this extension, $\delta^* < \overline{\delta}$.*

Proof of Proposition A3. Observe that $\delta^* = 0$ because an experimentation policy consisting of (i) exploration in domain 1 and (ii) expanding scope to domain 2 if a project exceeding $2\tilde{v}$ is discovered in domain 1 is always feasible. Further, $\bar{\delta} > 0$ because $\mathcal{C}(\cdot) \to 0$ as $\delta \to 0$, implying that the players cannot explore projects in domain 2 in period 1. □

3.3 Technological Interdependencies

In this extension, we assume that domains are perfectly technologically dependent. Formally, assume there is a bijection between the projects in the two domains, such that for any two projects paired by the bijection, the project values are equal. Such an assumption implies that any project value found in a given domain can be used to produce an identical value on the second domain. Further, for simplicity, we will again restrict attention to $\alpha = 1$ and $m = 2$, but the forces naturally extend.

Proposition A4 *Let δ^* and $\overline{\delta}$ be defined as in the main analysis. In this extension, $\delta^* < \overline{\delta}$, if the distribution of v_p admits a continuous density. Further, scope is terminally maximal with probability* 1 *for any non-empty experimentation.*

Proof of Proposition A4. The latter statement follows directly from technological interdependencies and Bernheim and Whinston (1990): if the players permanently exploit k projects with value $\hat{v}$, then the players can exploit m projects with value $\hat{v}$.

In this analysis, Proposition 1 continues to hold and the optimal experimentation policy only conditions on the best project found across all domains: $\hat{v}$. Denote by $B(\hat{v})$ the bellman equation corresponding to the players' strategy. Therefore, $\bar{\delta}$ is defined as

$$2c = \bar{\delta}\mathbb{E}_{\hat{v}_1,\hat{v}_2}(B(\max\{\hat{v}_1,\hat{v}_2\})|\bar{\delta}). \tag{26}$$

To prove, $\delta^* < \bar{\delta}$, it suffices to show:

$$c < \bar{\delta}\lim_{\delta\uparrow\bar{\delta}}\mathbb{E}_{\hat{v}_1}(B(\hat{v}_1))|\delta) \iff 2\lim_{\delta\uparrow\bar{\delta}}\mathbb{E}_{\hat{v}_1}(B(\hat{v})|\delta) > \mathbb{E}_{\hat{v}_1,\hat{v}_2}(B(\max\{\hat{v}_1,\hat{v}_2\})|\bar{\delta}). \tag{27}$$

$B(\cdot|\bar{\delta})$ is characterized by two thresholds, $\underline{\hat{v}}, \bar{\hat{v}}$: for $\hat{v} \geq \bar{\hat{v}}$ the players exploit two projects with value $\hat{v}$, for $\hat{v} \in [\underline{\hat{v}}, \bar{\hat{v}})$ the players exploit one project with value $\hat{v}$ and explore a new project, and for $\hat{v} < \underline{\hat{v}}$ the players explore two new projects.

Consider an alternative experimentation policy with $\epsilon > 0$: for $\hat{v} \geq \bar{\hat{v}} + \epsilon$ the players exploit two projects with value $\hat{v}$, for $\hat{v} \in [\underline{\hat{v}} + \epsilon, \bar{\hat{v}} + \epsilon)$ the players exploit one project with value v and explore a new project, and for $\hat{v} < \underline{\hat{v}} + \epsilon$ the players explore one new project. As $\epsilon > 0$, this experimentation policy is implementable for $\hat{v} > \underline{\hat{v}} + \epsilon$ as $\delta \uparrow \bar{\delta}$. Hence, it suffices to show that this policy is implementable at period 1.

Under this policy, $\lim_{\delta\uparrow\bar{\delta}} B(0|\delta) > \frac{1}{2}B(0|\bar{\delta})$, as there are no mass points in the distribution. Therefore, the expected continuation value of this policy at date 1 when $\delta \uparrow \bar{\delta}$ is strictly greater than half the expected continuation value of the optimal policy when $\delta = \bar{\delta}$. This assertion is exactly the claim we needed to show Equation (27) holds.

□

3.4 Expanded Analysis of Section 5.1

Recall the setup from Section 5.1 where $p = \Pr(v_p = \bar{v})$ and $n \in \{0, 1, 2\}$ is the number of domains where a $\bar{v}$ project has been found. Let $B^e(n)$ denote the expected joint surplus under policy $e \in \{\mathrm{FB}, \mathrm{GE}, \mathrm{GW}\}$ at state n. Since $\alpha = 1$, the reneging temptation per active domain is c. For all three policies, once $n = 2$, both domains are permanently exploited, yielding $B(2) = \frac{2(\bar{v}-2c)}{1-\delta}$.

3.4.1 First-Best (FB) Threshold δ^{FB}

The FB policy treats domains independently, exploring both starting from $n = 0$. Because domains are symmetric, $B^{\mathrm{FB}}(0)$ is twice the expected surplus of a single domain:

$$B^{\mathrm{FB}}(0) = 2\frac{\mathbb{E}(v_p) - 2c + \frac{\delta p}{1-\delta}(\bar{v} - 2c)}{1 - \delta(1-p)}. \tag{28}$$

The most binding implementability constraint occurs at $n = 0$, where the total deviation temptation is $2c$. The constraint requires $2c \leq B^{\mathrm{FB}}(0) - 2(\mathbb{E}(v_p) - 2c)$. Substituting $B^{\mathrm{FB}}(0)$ yields:

$$c \leq \frac{\mathbb{E}(v_p) - 2c + \frac{\delta p}{1-\delta}(\bar{v} - 2c)}{1 - \delta(1-p)} - (\mathbb{E}(v_p) - 2c). \tag{29}$$

Multiplying by $1 - \delta(1-p)$ isolates the surplus terms:

$$c(1 - \delta(1-p)) \leq \mathbb{E}(v_p) - 2c + \frac{\delta p}{1-\delta}(\bar{v} - 2c) - (\mathbb{E}(v_p) - 2c)(1 - \delta(1-p)). \tag{30}$$

Simplifying the right side and multiplying the entire inequality by $1-\delta$ to clear the remaining fraction gives:

$$c(1-\delta) - c\delta(1-\delta)(1-p) \leq \delta p(\bar{v} - 2c) + \delta(1-\delta)(1-p)(\mathbb{E}(v_p) - 2c). \tag{31}$$

Expanding both sides and grouping by powers of δ defines the quadratic constraint:

$$(1-p)(\mathbb{E}(v_p) - c)\delta^2 - (p(2-p)\bar{v} + (1-p)^2\underline{v} - cp)\delta + c \leq 0. \tag{32}$$

Solving this equation gives the exact closed-form threshold δ^{FB} as its smaller root:

$$\delta^{FB} = \frac{p(\bar{v} - c) + (1-p)\mathbb{E}(v_p) - \sqrt{[p(\bar{v} - c) + (1-p)\mathbb{E}(v_p)]^2 - 4c(1-p)(\mathbb{E}(v_p) - c)}}{2(1-p)(\mathbb{E}(v_p) - c)}. \tag{33}$$

3.4.2 Gradual-Exploit (GE) Threshold δ^{GE}

The GE policy begins with one exploration at $n = 0$, and at $n = 1$, it exploits the single $\bar{v}$ project while exploring the other domain. The value functions are:

$$B^{\mathrm{GE}}(1) = \frac{\bar{v} + \mathbb{E}(v_p) - 4c + \frac{2\delta p}{1-\delta}(\bar{v} - 2c)}{1 - \delta(1-p)}, \quad B^{\mathrm{GE}}(0) = \frac{\mathbb{E}(v_p) - 2c + \delta p B^{\mathrm{GE}}(1)}{1 - \delta(1-p)}. \tag{34}$$

The GE policy must satisfy implementability constraints at both $n = 1$ (temptation $2c$) and $n = 0$ (temptation c). These constraints satisfy

$$2c \leq B^{\mathrm{GE}}(1) - (\bar{v} + \mathbb{E}(v_p) - 4c) \quad \text{and} \quad c \leq B^{\mathrm{GE}}(0) - (\mathbb{E}(v_p) - 2c), \tag{35}$$

and δ^{GE} is the minimum of the two values of δ that bind these two inequalities.

3.4.3 Gradual-Wait (GW) Threshold δ^{GW}

There are three incentive constraints that may be binding. Therefore, the Bellman equations satisfy:

$$B^{\mathrm{GW}}(1) = \frac{\mathbb{E}(v_p) - 2c + \frac{2\delta p}{1-\delta}(\bar{v} - 2c)}{1 - \delta(1-p)}, \quad B^{\mathrm{GW}}(0) = \frac{\mathbb{E}(v_p) - 2c + \delta p B^{\mathrm{GW}}(1)}{1 - \delta(1-p)}. \tag{36}$$

In this case the constraint at $n = 1$ is weaker than that of $n = 0$, thus the only constraints are $n = 0$ and $n = 2$ as shown below

$$c \leq B^{\mathrm{GW}}(0) - (\mathbb{E}(v_p - 2c)) \quad \text{and } 2c \leq \frac{\delta(\bar{v} - 2c)}{1 - \delta}, \tag{37}$$

and δ^{GW} is the minimum of the two values of δ that bind these two inequalities.

4 Further Discussion of the Theoretical Literature

In this section, we revisit the theoretical literature in light of the empirical and qualitative findings on buyer–supplier relationships and persistent productivity differences presented in Section 6. We focus on how our framework compares with alternative theoretical models in generating and rationalizing these patterns.

As discussed in Section 1.1, two frameworks related to ours (with explicit links to buyer–supplier dynamics and productivity differences) are the model of Chassang (2010) and the dynamic-screening literature, especially the strand on gradualism (Watson, 1999, 2002). We therefore focus on these frameworks here.

Another natural benchmark against which to compare our model is a symmetric-learning model in which both parties start uninformed about the match-specific surplus. For instance, Macchiavello and Morjaria (2015) develop a setting where neither the buyer nor the supplier knows the supplier's reliability. Such frameworks replicate many of the stylized facts documented in Section 6.1 and–if the surplus is allowed to evolve stochastically–can also generate the observed expansions and contractions in relationship scope. We do not compare our model with this class of learning models for two reasons. First, the qualitative evidence in Section 6.1 singles out incentive conflicts between buyers and suppliers, and their joint exploration/exploitation choices, as the key drivers of relational dynamics. Second, most patterns produced by symmetric-learning models already arise in single-agent settings with aggregate uncertainty, thereby downplaying the strategic interaction and trust-building that lie at the heart of buyer–supplier relationships. Accordingly, we regard our framework as complementary: it isolates how incentive problems shape relational dynamics, while learning models highlight the additional role of aggregate uncertainty.

In Table A.1 below, we revisit each empirical or qualitative phenomenon discussed in Section 6 and compare our model to dynamic screening frameworks–such as those in (Watson, 1999, 2002)–as well as to the framework developed by Chassang (2010), focusing on their ability to rationalize these patterns and the underlying economic logic they rely on.

Table A1: Comparing Theoretical Frameworks on Section 6.1 Outcomes

Phenomenon	**Angelucci–Orzach (this paper)**	**Dynamic Screening à la Watson (1999, 2002)**	**Chassang (2010)**
Firms engage in lengthy joint exploration (R&D) before exploitation (commercialization)	Players share symmetric uncertainty about project payoffs and explore until the continuation value is high enough to sustain exploitation.	Exploration and exploitation decisions are not modeled.	Agent already knows all project payoffs; the model centers on information elicitation, not joint experimentation.
Increase in scope / scale of collaboration	When δ is low (resp. high), the parties "start small" (resp. "start big"). Full scope in long run not guaranteed to occur.	The players "start small" and increase stakes gradually for all δ. Full scope in long run guaranteed to occur in relationships with good types.	No natural notion of scope. If scope is proxied by the number of actions played with positive probability during non-punishment phases, it weakly rises over time.
Decrease in scope / scale of collaboration	With multiple domains, exploitation in one domain lowers continuation value, which can lead to contraction in other domains.	Absent; scale either rises or falls to zero.	No natural notion of scope. Even under the definition above, scope is monotone in non-punishment phases, so contractions are ruled out.
Path dependence	Random project discoveries alter continuation value, making both the short-run trajectory and long-run outcome history-dependent.	Absent: Relationships with good types evolve identically, depending solely on the passage of time.	Random project discoveries alter continuation value, making both the short-run trajectory and long-run outcome history-dependent.
Persistent productivity gaps among ex-ante similar collaborations	Path dependence yields persistent productivity differences.	Absent: relationships with good types evolve identically.	Non-infinitesimal productivity gaps have not been formally shown to exist: punishment equilibria are only partially characterized, leaving cross-sectional heterogeneity unaddressed.

References

Bernheim, B. Douglas and Michael D. Whinston, "Multimarket contact and collusive behavior," *The RAND Journal of Economics*, 1990, pp. 1–26.

Chassang, Sylvain, "Building routines: Learning, cooperation, and the dynamics of incomplete relational contracts," *American Economic Review*, 2010, *100* (1), 448–65.

Macchiavello, Rocco and Ameet Morjaria, "The value of relationships: evidence from a supply shock to Kenyan rose exports," *American Economic Review*, 2015, *105* (9), 2911–2945.

Watson, Joel, "Starting small and renegotiation," *Journal of Economic Theory*, 1999, *85* (1), 52–90.

_ , "Starting Small and Commitment," *Games and Economic Behavior*, 2002, *38* (1), 176–199.