Searching for Collaboration: The Dynamics of Relationship Building

Charles Angelucci and Roi Orzach*

November 15, 2023

Abstract

Building a successful collaboration is often a time-intensive and gradual process. We model collaborative dynamics with self-enforcing incentives. Two players are presented with infinitely many ex-ante identical projects, each yielding asymmetric benefits. Every period, they collectively explore or exploit multiple projects and make voluntary transfers to each other. After exploring a project, players learn its benefits and choose whether to exploit the project in future periods. We show that lengthy exploration occurs, and that the way the collaboration evolves exhibits significant path dependence. Players temporarily exploit projects, return to previously abandoned projects, and initially explore a limited number of projects.

*Angelucci: MIT Sloan School of Management, 50 Memorial Drive, Cambridge MA 02142 (e-mail: cangeluc@mit. edu); Orzach: Department of Economics, MIT, 50 Memorial Drive, Cambridge MA 02142 (e-mail: orzach@mit.edu).
We thank Abhijit Banerjee, Daniel Barron, Alexander Bloedel, Alessandro Bonatti, David Byrne, Yifan Dai, Joyee Deb, Mert Demirer, Glenn Ellison, Dana Foarta, Guillaume Fréchette, Bob Gibbons, Marina Halac, Amit Khandelwal, Danielle Li, Jin Li, Chiara Margaria, Simone Meraglia, Aroon Narayanan, Juan Ortner, Hazhir Rahmandad, Daniel Rappaport, Patrick Rey, Kareen Rozen, Klaus M. Schmidt, Michael Ting, Jean Tirole, Marta Troya Martinez, Michael Whinston, and participants in both the MIT Organizational Economics and Theory groups, LMU Munich, the TSE Microeconomic Theory Workshop, the 2023 Montreal IO Conference, and the 2023 SIOE Conference.

1 Introduction

Collaborations rely on the promise of sustained cooperation between the parties involved. Managers rely on promises of bonuses to incentivize their subordinates, with the expectation that the subordinates will reciprocate by intensifying their efforts to further management's goals. Firms enter strategic alliances to bolster R&D, and the enduring success of these collaborations relies heavily on the continued willingness of firms to share resources. In supply-chain relationships, buyers pledge to increase future transactions with suppliers who, in turn, pledge to adopt practices that primarily benefit the buyer, like enhanced quality control or cost reductions. In each of these contexts, promises play a key role in providing incentives, steering collaborations towards profitability. However, the efficacy of promises hinges on their credibility. How can those engaged in collaboration build trust in their relationships, and how exactly does it translate into collaborative success?

The credibility of promises in collaborations relates to the value of the relationship. When the collaboration is highly valued, partners are less likely to break promises. Consequently, a collaboration that offers significant present and future benefits not only leads to high gains but also fosters credibility, allowing the parties to make substantial promises and realize a more profitable collaboration overall. Thus, the interplay between collaboration value and credibility establishes a reinforcing relationship, providing a strong impetus for the parties to enhance the value of their collaboration as much and as early as possible. Yet, building a successful collaboration is often a gradual process, especially when the benefits are distributed unequally, requiring credible promises of compensation to ensure all parties cooperate. In these instances, substantial time is invested to define the collaboration's nature. Moreover, rather than expanding all at once, collaborations tend to evolve incrementally. As noted by McKinsey about buyer-supplier relationships: "Building trust takes time and effort. Often this means starting small, with simple collaboration efforts that deliver results quickly, building momentum" (Gutierrez et al., 2020). Similarly, Dwyer et al. (1987) note that "the critical distinction [between the exploration and expansion phases in buyer-supplier relationships] is that the rudiments of trust and joint satisfactions established in the exploration stage now lead to increased risk taking within the dyad. Consequently, the range and depth of mutual dependence increase." In this paper, we present a model that rationalizes the notion that, despite the parties' strong motivation to grow their collaboration quickly, the path to a successful partnership often proceeds at a slow and deliberate pace. This process often involves unavoidable delays and exhibits considerable path dependence, due to the inherent time required to establish credibility.

The key features of our model are as follows. We consider a discrete time framework where two players interact repeatedly over an infinite time horizon, with all actions being publicly observable. Each period provides players with an opportunity to cooperate on multiple projects chosen from an infinite pool of ex ante identical projects. Any project that has been selected before can be chosen again for collaboration, a scenario we call "project exploitation." The benefits of each project are time-invariant but initially uncertain, and they may vary asymmetrically across the players. Cooperation on a new project, or "project exploration," immediately reveals its actual benefits. Moreover, all projects entail a constant fixed cost for the players, both during the exploration and exploitation phases. As a result, players might be reluctant to cooperate in exploring projects if they expect that their individual benefit will not exceed the cost, and they may similarly be reluctant to cooperate in exploiting a project if their realized individual benefit falls below the cost. Finally, players can transfer money to each other, but these transfers are voluntary in nature.

We focus on relational contracts (i.e., subgame-perfect Nash equilibria) that maximize the players' joint surplus. In the main setting, we assume that all projects yield asymmetric benefits across the players, implying that their incentives to collaborate on any given project are not aligned. To highlight the importance of building credibility within such asymmetric contexts, we contrast the dynamics that emerge under the optimal relational contract against those from a benchmark scenario where every project equally benefits both players, eliminating the issue of credibility.

In the first part of our analysis, we assume that project exploration does not need to be motivated. Specifically, cooperating in exploring a project is advantageous for the players because they both expect their per-period individual benefit to outweigh their cost. The central question then becomes the identification of projects worth exploiting. We initially suppose that the players can only explore or exploit at most one project per period. We show that players transition to permanent project exploitation only when they identify a project that offers sufficient aggregate benefits across the two players, with monetary transfers serving to redistribute these benefits. With a high discount factor, players' criteria regarding which projects to exploit coincide with that of the symmetric-benefits benchmark. Conversely, when the discount factor is low, the players become more selective with asymmetric benefits. Their primary focus shifts towards identifying a project valuable enough to make the continued relationship sufficiently beneficial to enable cooperation in project exploitation. In these instances, the resulting exploration time is likely to substantially exceed that observed under the symmetric-benefits benchmark.

Next, we allow the players to explore or exploit multiple projects per period while maintaining the assumption that project exploration does not need motivation. When the discount factor is high, the players' actions again replicate those that arise in the symmetric-benefits benchmark. They treat the search for these projects independently and opt for a project's exploitation when its aggregate benefits exceed a common threshold set for all projects. When the discount factor is low, the players are again likely to engage in lengthy exploration compared to the symmetric-benefits benchmark. Furthermore, their search for exploitable projects becomes interdependent. Specifically, the thresholds used for assessing whether a project is worth exploiting permanently are co-determined and differ across projects. This co-determination arises because selecting a project for exploitation impacts the overall value of the relationship and, consequently, the players' ability to cooperate on exploiting other projects. As a result, in their quest to identify projects for permanent exploitation, the players may opt for the temporary exploitation of certain projects with the understanding that they might abandon them later, or they may bypass some projects, only to return to them when the value of their relationship has grown sufficiently to enable exploitation. In sum, inefficiencies arise not only in terms of the time spent exploring projects, but also in the deviation from the decision rule consisting of either permanently exploiting projects or permanently abandoning them.

In the third part of our analysis, we no longer assume that both players are motivated to explore projects. As a result, the scope of the players' relationship – determined by the number of projects they are exploring or exploiting – may not reach its maximum right from the beginning. Specifically, when the players' discount factor is high, players immediately explore as many projects per period as possible. In contrast, when the players' discount factor is low, the players find themselves compelled to adopt a gradual approach, expanding the scope of their relationship incrementally over time. The gradual approach involves initially exploring a limited number of projects and starting the exploration of additional projects only after identifying projects worthy of exploitation. This gradual process offers two advantages. Firstly, the possibility of exploring and eventually exploiting additional projects in the future serves as an incentive for players to explore the early projects. Secondly, the exploitation of valuable early projects enhances the value of the players' relationship, thereby facilitating collaboration on additional projects.

Lastly, we analyze an extension of the model with projects that yield both symmetric and asymmetric benefits. We find that players set lower thresholds for symmetric projects compared to asymmetric ones and utilize symmetric projects as stepping stones, enabling cooperation in identifying asymmetric, yet more lucrative projects.

The rest of the paper is organized as follows. Section 1.1 discusses the related literature. Section 2 describes the model. Section 3 characterizes the set of optimal relational contracts we focus on and analyzes a benchmark scenario with aligned incentives. Section 4 solves the model, first by focusing on the players' project exploitation choices and then by analyzing the dynamics of the scope of their relationship. Section 5 analyzes extensions and Section 6 concludes.

1.1 Related Literature

This paper contributes to multiple strands of the literature. Firstly, our research connects to the large literature on multi-armed bandit problems, dating back to Weitzman (1979). For a comprehensive review, see Bergemann and Valimaki (2006). Only a small subset of this literature analyzes strategic interactions. For example, Bolton and Harris (1999) consider a setting in which players separately pull arms and free-ride on others' experimentation. In Strulovici (2010), players collectively choose between a safe arm and a risky one, with its asymmetric benefits revealed over time through experimentation. We contribute to this body of literature by introducing a setting characterized by analytically tractable dynamics and a wide array of applications.

Secondly, this work is related to the literature on relational contracts (see e.g., Bull, 1987; Macleod and Malcomson, 1989; Baker et al., 1994, 2002; Levin, 2003, for early contributions).[1] The positive feedback effect, which links the value of players' relationships to incentive strength, is pervasive across relational contracting models.

[1] Also at the intersection of the bandit and the relational contracting literatures, Urgun (2021) examines a scenario where a principal interacts with multiple agents whose publicly-observable types depend on the contracting history. We employ a multi-armed bandit problem to study when a group of players should transition from the exploration of projects to their (infinitely repeated) exploitation.

However, it rarely produces dynamics, because current production is typically influenced only by present actions, and not past choices as in our setting. An exception is Halac (2014), who studies a setting in which the value of the players' relationship increases with the duration of the relationship. The players initially choose to cooperate on low-risk, low-return projects, and they switch to high-risk, high-return projects once their relationship has grown sufficiently valuable.[2] In our setting, it is the discovery of projects worthy of collaboration that increases the value of the relationship. We analyze the implications of this effect on the players' choices between project exploration and exploitation when they are engaged in multiple projects. In contrast, Chassang (2010) analyzes a setting where increases in relationship value diminish the players' motivation to enhance their collaboration. In the model, the agent knows which arms are productive and which are not, while the principal, at the outset, cannot differentiate between the two. Without monetary incentives, incentivizing the agent to choose productive arms is accomplished by the threat of firing the agent following failures. This dynamic makes motivating exploration progressively expensive as more productive arms are identified. Should the relationship endure, it ultimately enters an "exploitation" phase and its value stops growing. In our model, the players are symmetrically informed about their environment, and the presence of transferable utility removes the need for inefficient on-path punishments. These two features lead to the positive feedback effect mentioned above.[3]

Our work also connects to Bernheim and Whinston (1990), who analyze firms operating in multiple markets, showing that maintaining collusion in easier markets can help support collusion in more challenging ones. Similarly, Levin (2002) shows the advantages firms gain by pooling heterogeneous employees' incentives into a "multilateral" relational contract. In our setting, the players' ability to pool relational incentives across multiple projects also enhances their relationship. However, increasing the scope of their relationship gradually might still be optimal.

Finally, we add to the body of research that examines the concept of gradualism. Watson (1999, 2002) examine a setting in which players are uncertain regarding their counterparts' intentions—to either collaborate genuinely or take advantage of the other. The players begin with minimal cooperation to mitigate the losses from defec-

[2]In Halac (2015), a principal leverages this feedback effect by making an upfront and relationship-specific investment prior to her repeated interaction with an agent.

[3]Moreover, introducing transferable utility within Chassang (2010), where information asymmetry plays a central role, would make the value of the players' relationship constant on path.

tion. As the players become more optimistic, the collaboration grows. Collaborations involving trustworthy players achieve optimal cooperation, while those with untrustworthy players eventually end. In our model, no player begins the collaboration with the intent to take advantage of the other. Instead, each is inclined to cooperate as long as it serves their best interest. The relationship develops incrementally, not due to screening intentions, but because credibility is built by the players over time.[4,5]

2 The Setup

Two collaborators interact repeatedly. During each interaction, they select multiple projects to cooperate on and exchange money. These projects entail costs for the parties, and their benefits are initially uncertain and may not be evenly distributed. Examples of such collaborations include a buyer and a supplier, a manager and an employee, or two firms in an R&D alliance.

There are 2 players who have the opportunity to interact at different time periods $t = 0, 1, 2, \ldots$. Each player, denoted by $i = 1, 2$, has a discount factor δ and a per-period outside option equal to zero. The players' interaction spans across m "dimensions," where m is determined exogenously. In each period t, each player chooses a finite set P_i^t of no more than m projects from the set $\mathcal{P} = [0, m)$. For each dimension $j = 1, \ldots, m$ of the relationship, the players may select up to one project from the set $[j-1, j)$. We assume that a project is selected if and only if both players choose it, thus following a unanimity rule. We denote by $\mathbf{P}^t$ the corresponding set of projects selected by the players, where $\mathbf{P}^t = P_1^t \cap P_2^t$. We refer to $|\mathbf{P}^t| \leq m$ as the players' "relationship scope."

Each project in $\mathbf{P}^t$ imposes a cost of $c > 0$ on both players. Further, each

[4] Macchiavello and Morjaria (2015) show that in the context of the Kenyan Rose export market, the duration of buyer-supplier relationships is predictive of the scale of the relationship. This finding is rationalized in a learning model in which neither the buyer nor the supplier know the supplier's reliability. Our framework centers on the need for relationships to be built rather than learning about reliability or productivity. Moreover, the trade literature has documented the strong persistence of buyer-supplier relationships, ascribed to substantial switching or search costs (see e.g., Bernard and Moxnes, 2018; Monarch, 2022). This literature has also noted the tendency for transaction volumes to increase over long periods of time (see e.g., Monarch and Schmidt-Eisenlohr, 2023). Our framework suggests that buyers may exhibit reluctance to switch suppliers when such a transition entails the need to rebuild a relationship from scratch.

[5] Gradualism also arises in Ghosh and Ray (1996) and Kranton (1996), where players are randomly matched and can exit relationships at any time, with new partners unaware of their history. Initial cooperation levels are low and gradually increase to discourage defection from existing relationships.

project $p \in \mathcal{P}$ is associated with a vector of time-invariant individual valuations $(v_{p,1}, v_{p,2}) \in \mathbf{R}^2_+$. A project's associated individual valuations are publicly observed immediately after a project is selected for the first time. We say that a project p is being "explored" in period t if it is selected for the first time, and that it is being "exploited" during period t if it has been selected in some prior period.[6] Note that we place no intertemporal restrictions on the set of projects that are available; for instance, nothing prevents the players from exploring a project, potentially exploiting it for several periods, then temporarily abandoning it, and subsequently returning to it at a later time. We refer to a project's sum of individual valuations as the project's value, and denote it by s_p. We assume that each project p's vector of individual valuations is drawn independently and identically across projects and dimensions according to the cumulative distribution function F. This assumption implies that all dimensions of the relationship are ex ante identical.

The players can exchange money twice during each period. At the beginning of each period t, the players make discretionary transfers to each other, where $w^t_{i,-i} \in \mathbf{R}^+$ denotes such a transfer from player i to player $-i$. At the end of each period t, players again make discretionary transfers to each other, where $b^t_{i,-i} \in \mathbf{R}^+$ denotes such a discretionary transfer from player i to player $-i$.[7]

Player i's period t payoff is equal to:

$$\pi^t_i = w^t_{-i,i} - w^t_{i,-i} + b^t_{-i,i} - b^t_{i,-i} + \sum_{p \in \mathbf{P}^t} (v_{p,i} - c), \text{ where } i, \in \{1, 2\}\,. \qquad (1)$$

Equation (2) implies that one key friction between the players will be about the selection of projects that benefit one player but not the other, and that money will serve the purpose of aligning incentives. The source of this friction stems from the incurred cost c for any selected project.[8]

We conclude the model's description by stating the timing of the stage game. Both

[6] Our analysis remains qualitatively unchanged if we assume that benefits are realized solely during the exploitation phase of a project, which would capture a scenario in which exploration is purely experimental.

[7] We incorporate the option of monetary transfers both before and after the players' project choices, although removing either would not qualitatively affect our results. Without transfers at the beginning of each period, surplus might no longer be fully redistributed across the players without affecting incentives. Without transfers at the end of each period, incentives for the current period would rely on transfers from the subsequent period, complicating the proofs.

[8] Assuming that the cost c is incurred by player i for every project in P^t_i would lead to identical results.

players simultaneously choose their discretionary transfers $w_{i,-i}^t$. Next, both players simultaneously make their project choices P_i^t. For each selected project $p \in \mathbf{P}^t$, the players observe the vector $(v_{p,1}, v_{p,2})$ and pocket their individual valuation. Finally, both players simultaneously choose their discretionary transfers $b_{i,-i}^t$.

Relational Contracts. A relational contract is a complete plan for the relationship. Let $h^t = (\mathbf{w}^0, \mathbf{P}^0, \mathbf{v}^0, \mathbf{b}^0, \ldots, \ldots, \mathbf{w}^{t-1}, \mathbf{P}^{t-1}, \mathbf{v}^{t-1}, \mathbf{b}^{t-1})$ denote the history up to date t and $\mathcal{H}^t$ the set of possible date t histories, where boldface lowercase letters indicate vectors. Then, for each date t and every history $h^t \in \mathcal{H}^t$, a relational contract describes: (i) the $\mathbf{w}^t$ transfers; (ii) the set of projects $\mathbf{P}^t(\mathbf{w}^t)$ to be selected as a function of $\mathbf{w}^t$; and (iii) the $\mathbf{b}^t(\mathbf{w}^t, \mathbf{P}^t, \mathbf{v}^t)$ transfers as a function of $\mathbf{w}^t$, $\mathbf{P}^t$, and the realizations of $\mathbf{v}^t$. Such a relational contract is self-enforcing if it describes a Subgame Perfect Equilibrium of the repeated game. Within the class of Subgame Perfect Equilibria, we analyze pure-strategy equilibria which maximize the players' joint surplus.[9] In the event of a deviation in some period t, the players respond (i) by choosing $P_i^t = \emptyset$ and $b_{i,-i}^t = 0$ if these choices have not been made yet and (ii) by permanently breaking off their relationship (i.e., reverting to the worst equilibrium of the stage game from the next period onward). This punishment is without loss of generality as it occurs only out-of-equilibrium (c.f. Abreu, 1986).[10]

3 Preliminary Analysis

In this section, we characterize the set of surplus-maximizing relational contracts our analysis focuses on. We also analyze the benchmark case where every project equally benefits both players.

[9]Restricting attention to pure strategy equilibria is without loss because (i) mixing on transfers cannot benefit the players as it increases the maximal transfers they promise each other and (ii) mixing on projects either results in inefficiently low relationship scope due to miscoordination or a limited relationship scope that can also be achieved by players not choosing projects on some dimensions.

[10]Equivalently, in the period following a deviation, players could transition to an alternative continuation equilibrium in which everything remains the same, except that the entire surplus is allocated to the player who did not deviate. This punishment provides identical incentives. Because it is Pareto optimal, it is also less prone to renegotiation.

3.1 Characterization of Optimal Relational Contracts

In our setting, surplus-maximizing relational contracts will depend on the players' beliefs about the projects. Denote by $\mu^t(h^t) := \{\Delta(v_{p,1}, v_{p,2})|h^t)\}_{p\in[0,m)}$ the beliefs the players hold about the projects' valuations given all the observed valuations up through period $t-1$. We provide a characterization of one set of surplus-maximizing relational contracts that mirrors the characterization in Levin (2003). We show that there exist surplus-maximizing relational contracts that condition on h^t only through the beliefs $\mu^t(h^t)$. Formally, restricting attention to relational contracts that specify the same continuation equilibrium following any two on-path histories h_1^t and $h_2^{t'}$ that lead to the same beliefs μ is without loss. Further, the continuation equilibrium the relational contract prescribes is surplus-maximizing, in the sense that there does not exist another continuation equilibrium that generates a higher total surplus across the players. The following proposition formalizes this result and provides a necessary and sufficient condition for a given project selection rule (i.e., a mapping from beliefs to projects) to be implemented by a relational contract. The proof for this proposition, along those for any other result not proven within the main text, can be found in the Appendix.

Proposition 1. *The following statements are true:*

- *For any surplus-maximizing relational contract, there exists an alternative surplus-equivalent relational contract such that (i) for all t and for all on-path histories $h^t \in \mathcal{H}^t$ the continuation equilibrium is surplus maximizing and (ii) for any two on-path histories h_1^t and $h_2^{t'}$, if $\mu^t(h_1^t) = \mu^{t'}\left(h_2^{t'}\right)$, then the relational contract specifies the same continuation equilibrium following these histories. We call such relational contracts <u>optimal</u>.*

- *There exists a relational contract that implements a project selection rule $\mathbf{P}(\cdot)$ if and only if the following inequality holds for all t and for all histories $h^t \in \mathcal{H}^t$:*

$$\sum_{p\in\mathbf{P}(\mu^t)} \sum_{i=1}^{2} \max(0, c - \mathbb{E}(v_{p,i}|\mu^t)) \leq \mathcal{C}(\mu^t), \tag{2}$$

  where $\mathcal{C}(\mu^t)$ ("the continuation value") is the expected net present value of the players' relationship starting in $t+1$ given $\mathbf{P}(\cdot)$ and μ^t.

The intuition for the first statement of the proposition is based on the following two observations. First, any surplus-maximizing relational contract is necessarily surplus-maximizing following any on-path history, for otherwise non-surplus-maximizing continuation equilibria could be replaced with surplus-maximizing ones, with transfers appropriately designed to maintain all players' incentives. Second, by confining our attention to surplus-maximizing continuation equilibria, we show that the only history-dependent outcome that can alter the set of optimal continuation equilibria are the players' beliefs μ^t about the projects.

Recall that the main tension for the players is that the project selection rule which maximizes their joint surplus may involve the selection of projects that do not individually benefit each player. Inequality (2) states that for a relational contract to implement a given project selection rule everywhere on the equilibrium path, the continuation value must exceed the total reneging temptation across players and projects in all periods and for all possible histories. The total reneging temptation is the sum across players and across projects of a project's reneging temptation to a player. The sum is across projects because each player can deviate from the relational contract by selecting any subset of $\mathbf{P}^t$. In turn, a project's reneging temptation to a player is either equal to zero, in case the project generates a positive net expected gain to the player, or equal to the magnitude of the net expected loss. That the relational contract creates more continuation value to the players than the sum of their gains from defecting is necessary for the relational contract to constitute an equilibrium. In the proof, we show that the presence of transferable utility also ensures that this condition is sufficient.

Finally, we note that the second statement of the proposition means that characterizing the optimal relational contract can be reduced to characterizing the players' optimal project selection rule, which will thus be the focus of our analysis hereafter. To understand this, observe that all transfers cancel each other in the expression for the joint surplus of the players, as well as on the right-hand side of Equation (2).

3.2 Benchmark with Symmetric Benefits

We now analyze the benchmark case where every project benefits the players equally, ensuring that they have perfectly aligned incentives. Specifically, we suppose that for each project $p \in \mathcal{P}$, $v_{p,1} = v_{p,2}$.[11] The predictions this benchmark analysis produces are identical to those that would result if a single decision-maker, whose payoff is given by the sum of the payoffs of both players, were to make all the decisions. We show that the standard results from single-agent multi-armed bandit problems in stationary settings hold in this setting (e.g., Weitzman, 1979).

Proposition 2. *When projects generate symmetric benefits, all optimal relational contracts specify a project selection rule that is identical and independent across all m dimensions of the players' relationship. Further,*

1. *If the players select any project in any period t for dimension j, then the players select a project for dimension j in all periods.*

2. *There exists a monotone increasing function $s^0(\delta)$ such that the players exploit project p for dimension j if and only if $s_p \geq s^0(\delta)$.*

When projects yield equal benefits for both players, Inequality (2) from Proposition 1 simply states that the net present value of the net payoff resulting from the selection of any project (accounting for the potential abandonment of a project) must be non-negative. Because this feature will always hold under any optimal relational contract, Inequality (2) can be ignored.

The intuition behind the players treating each dimension of their relationship separately and identically follows from our assumptions wherein (i) payoffs are additively separable across projects, meaning there are no interdependencies like economies or diseconomies of scope, and (ii) all projects benefit the players equally.

The intuition for statement (1) is that if the players find it rational to explore project $p \in [j-1, j)$ in some period t, then exploration must exhibit a positive net present value of the players' net payoffs, accounting for the possibility of project abandonment. Since the players have access to an infinity of ex ante identical projects, they would thus always opt for exploration rather than the non-selection of a project and, by extension, a project will be selected in every period.

To gain intuition for statement (2), note that when the players find it optimal to select a project (as opposed to not selecting any project), the players either (i) exploit a previously-explored project p or (ii) explore new projects with the hope of

[11] We assume that every project benefits the players equally for simplicity. The findings of this subsection remain valid for asymmetric benefits as long as each project either positively or negatively affects both players.

ultimately settling on a superior project in the future. The players never return to an abandoned project because we have assumed an infinite supply of ex ante identical projects, which ensures that they always new potentially viable projects to explore. Finally, statement (2) also states that as the discount factor increases, the value of exploring alternative projects rises, since any superior project identified can be used across all future periods.

In sum, in our environment, when benefits are symmetric, the players maximize the scope of their relationship at all times and they switch to permanently exploiting projects based on an independent, identical, and time-invariant threshold. These features will not always be true when benefits are asymmetric, a scenario we now analyze.

4 Analysis

We divide the analysis of the asymmetric benefits case into two subsections. In subsection 4.1, we assume that motivating the players to explore projects is not a concern and instead focus on their decision-making process regarding project exploitation. Initially, we consider the case where they can collaborate on only one project per period, followed by an examination of the case where they can collaborate on multiple projects per period. In subsection 4.2, we shift our focus to settings where the players may require motivation to explore projects. Within this context, we analyze the scope of the players' relationship.

4.1 The Dynamics of Collaborative Project Selection

We specialize the model described in Section 2 by assuming that each project benefits only a single agent. Removing the possibility of projects that benefit both players allows us to focus squarely on the forces of interest: getting players to build the credibility necessary to cope with incentive problems.[12] Moreover, for any given project p, it is randomly determined (i.i.d. across projects) which player among the two receives the entirety of the chosen project's associated benefits. Specifically, we assume that $(v_{p,1} = s_p, v_{p,2} = 0)$ occurs with a probability of $\frac{1}{2}$ and that $(v_{p,1} = 0, v_{p,2} = s_p)$ occurs

[12]In Section 5, we consider the scenario where projects benefiting both players coexist with those benefiting only a single player.

with a probability of $\frac{1}{2}$. This approach assumes an extreme form of asymmetric benefits. However, for our results to hold qualitatively, it suffices that each project yields a net benefit to just one player. Finally, we suppose that the distribution of s_p is subject to the following restrictions: (i) $\infty > \mathbb{E}(s_p) > 2c$, implying that exploring a project selected at random is both Pareto optimal and an equilibrium of the stage game, and (ii) $\text{supp}(s_p)$ is convex, implying that there exists a unique cutoff between exploration and exploitation.[13]

4.1.1 Collaborating on a Single Project

Suppose the dimensionality of the players' relationship, m, is equal to one. We show that when the discount factor is high, the players' behavior in terms of project selection corresponds to their behavior in the symmetric-benefits benchmark. However, as the discount factor decreases, the players are likely to explore a greater number of projects before settling on one for exploitation.

With asymmetric benefits, for any project, one player will not find exploitation in their interest. However, if cooperation in a project's exploitation is socially desirable, then for any such project, there exists a relational contract enabling both players to choose the project for exploitation in case the discount factor is sufficiently high. If the discount factor is not sufficiently high, the players must continue exploring projects. This observation gives the players an additional reason to explore projects. Finding a better project not only increases the players' continuation value because of the project's greater worth, but also because a more valuable project is easier for the players to cooperate on.

Further, in any optimal relational contract, if the players exploit project p with value s in period t, then $\mu_t = \mu_{t+1}$ since the players have not acquired any additional information during period t. It follows that the players also exploit project p in period $t+1$ and in all subsequent periods and, hence, the continuation value in period t is simply $\frac{\delta}{1-\delta}(s - 2c)$. We make the following assumption to ensure Equation (2) holds for a positive measure of projects.

Assumption 1. *There exists an* $s_p \in \text{supp}(F)$ *such that* $c < \frac{\delta}{1-\delta}(s_p - 2c)$.

[13] These assumptions are not crucial, but simplify the proofs and statements of the results.

We denote the minimum value of s that fulfills Assumption 1 by $\tilde{s}(\delta)$, where:

$$\tilde{s}(\delta) = c\frac{1+\delta}{\delta}. \tag{3}$$

Similar to the symmetric-benefits benchmark, we show that the optimal relational contract involves exploring projects until finding a project whose value s_p exceeds some threshold, denoted s^*. Further, the threshold is equal to the maximum between $\tilde{s}$ and s^0 (i.e., the threshold when benefits are symmetric).

Proposition 3. *In any optimal relational contract, there exists a threshold* $s^*(\delta) = \max\{\tilde{s}(\delta), s^0(\delta)\}$ *such that the players explore projects until they find a project* p *with an associated value* $s_p \geq s^*$. *Once they find such a project, the players exploit it in all subsequent periods.*

The intuition is as follows. From Proposition 1, we know that the value of the best project found thus far is the only factor in determining what optimal play is. Also, and as previously mentioned, once players engage in the exploitation of a project, they do so permanently. Consequently, the payoff derived from exploiting a project is increasing in the value of the best project. Conversely, the payoff from exploring a project is independent of the value of the best project. It follows that the players follow a threshold rule when deciding whether to explore or exploit the best project they have found thus far.

Because the joint surplus of the players in the symmetric benefits benchmark represents an upper bound on their joint surplus in scenarios with asymmetric benefits, the players choose to exploit any project with a value exceeding $s^0(\delta)$ if they are able to. Moreover, the players are only able to cooperate in exploiting projects with a value exceeding $\tilde{s}(\delta)$. Therefore, the criterion for project exploitation is that a project's value exceeds $\max\{\tilde{s}(\delta), s^0(\delta)\}$.

Given Proposition 2, $s^0(\delta)$ is monotonically increasing in δ. However, from (3) it follows that $\tilde{s}(\delta)$ is monotonically decreasing in δ. Thus, there exists a value δ^* such that, when $\delta < \delta^*$, s^0 falls below s^*. Moreover, as δ decreases, the gap between s^* and s^0 widens. The intuition for this widening gap is as follows: When benefits are symmetric, a decrease in δ means that players derive lower benefits from exploration, causing s^0 to decrease. Conversely, in the case of asymmetric benefits, players must identify an even more valuable project to maintain cooperation, which leads to an increase in s^*. This intuition is formalized below.

Corollary 1. *The threshold s^* is monotonically increasing (respectively, monotonically decreasing) in δ when $\delta > \delta^*$ (respectively, when $\delta < \delta^*$).*

Figure 1 provides an illustration by plotting the thresholds $\tilde{s}$, s^*, and s^0 as functions of δ, when $c = 1$ and $s_p \sim \text{Exp}(\frac{1}{3})$, where $\text{Exp}(\lambda)$ represents an exponentially distributed random variable with parameter λ. This distribution satisfies our assumptions as a randomly selected project has an expected value equal to 3, which, in turn, guarantees (i) that project exploration is an equilibrium of the stage game and (ii) that, for any δ, there always exists a project with some value v that can be exploited.

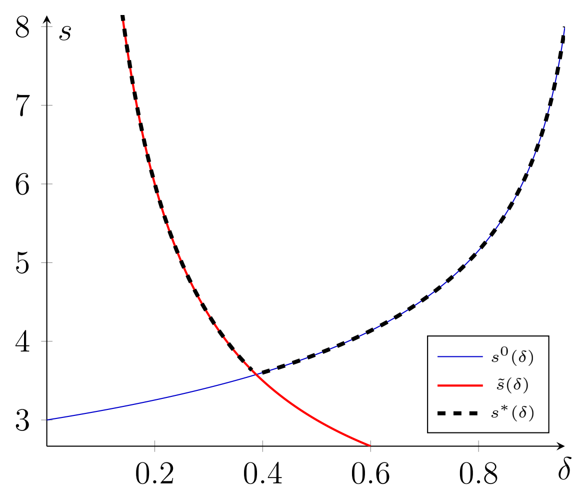


Figure 1: Comparison of Project Exploitation under Symmetric and Asymmetric Benefits

The figure plots the $\tilde{s}$, s^*, and s^0 thresholds when $s_p \sim \text{Exp}(\frac{1}{3})$ and $c = 1$. The threshold $\tilde{s}$ is the minimum value of s such that cooperation in project exploitation is sustainable. The threshold s^0 is the minimum value of s such that the players switch from exploration to exploitation in the benchmark with symmetric benefits. The threshold s^* is the minimum value of s such that the players switch from exploration to exploitation. The closed-form solutions are $s^0 = 3W\left(\frac{\delta}{e(1-\delta)}\right) + 3$, where $W(\cdot)$ is the Lambert W function, and $\tilde{s} = 2 + \frac{1-\delta}{\delta}$. As shown in the figure, s^* is the point-wise maximum of $\tilde{s}(\delta)$ and $s^0(\delta)$.

The threshold s^0 is increasing in δ and approaches infinity as $\delta \to 1$. This positive relationship arises because, in the presence of symmetric benefits, agents become more selective as their patience increases. In contrast, the threshold $\tilde{s}$ approaches infinity as $\delta \to 0$ and it decreases as δ increases. This relationship occurs because, in the presence of asymmetric benefits, the players can cooperate in exploiting a wider range of projects as their patience grows. As a result, for low values of δ, the players must be more selective with asymmetric compared to symmetric benefits, whereas when δ is high, they are equally selective in both scenarios.

4.1.2 Collaborating on Multiple Projects

We now extend the analysis from the previous subsection to consider the scenario in which the dimensionality of the players' relationship is equal to $m \geq 1$. When benefits are symmetric, the decision to continue exploration or exploit a specific project is made independently for each project. As a result, the threshold $s^0(\delta)$ continues to determine whether the players choose to permanently exploit a particular project. In contrast, with asymmetric benefits the players continue to explore more than in the case of symmetric benefits. In addition, we show that the independence property across dimensions fails, leading the players to adopt behaviors such as the temporary exploitation of projects and the recalling of previously abandoned projects.

We begin with the observation that equilibrium play under the optimal relational contract in any period t relies solely on the values of the most valuable projects discovered thus far for each of the m dimensions of the players' relationship, denoted as $\hat{s}_1, \ldots, \hat{s}_m$. To see this, note that exploiting any project increases the left-hand side of Equation (12) by an amount equal to one. Thus, if the players can exploit a project in the current period, they can exploit any project in the current period, and hence by Proposition 1, if the players exploit a project, they will exploit the project achieving the highest value of the surplus.

Further, recall the definition of $\tilde{s}$ as established in Equation (3). An implication of Proposition 3 is that the players can (and thus do) follow the project selection rule of the symmetric benefits benchmark from period 0 onwards if $s^0 \geq \tilde{s}$. If, instead, $s^0 \geq \tilde{s}$, the players follow the selection rule of the symmetric benefits benchmark only upon finding a project whose value s is such that $s \geq \tilde{s}$. The next lemma provides the condition under which the players are able to follow the project selection rule of the symmetric benefits benchmark when cooperating on m projects per period.

Lemma 1. *Upon finding projects with values $\hat{s}_1, \ldots, \hat{s}_m$, the players can follow the selection rule of the symmetric benefits benchmark in all subsequent periods if and* *only if:*

$$f(\hat{s}_1, \ldots, \hat{s}_m) := \frac{1}{m} \sum_{j=1}^{m} \max\{\hat{s}_j, s^0\} \geq \tilde{s}. \tag{4}$$

When cooperating across multiple dimensions, the players can pool their relational incentives across all these dimensions to facilitate their collaboration. Consequently, the continuation value of the players' relationship becomes high enough to allow them to follow the project selection rule of the symmetric benefits benchmark (namely, to exploit a project if its value is weakly greater than s^0) whenever condition (4) holds. This condition states that, when considering the average across all dimensions, the maximum value between the best project found thus far in each dimension and the threshold s^0 must exceed the "feasibility" threshold $\tilde{s}$. We note that the function $f(\hat{s}_1, \ldots, \hat{s}_m)$ does not correspond to the arithmetic mean of the values $\hat{s}_1, \ldots, \hat{s}_m$ for two reasons: (i) the players will choose to explore rather than exploit a project with value lower than s^0 and (ii) exploration is valuable to the players and thus contributes to their continuation value. When condition (4) holds, the players can pool their relational incentives across all m dimensions in a way that enables them to follow the project selection rule of the symmetric benefits benchmark.

Observing that no project is guaranteed to be permanently exploited until all best projects within each dimension of cooperation are permanently exploited, the condition provided in Lemma 1 implies a necessary condition for the players to permanently exploit the highest value project found thus far for each dimension of their relationship. This scenario is referred to as "permanent exploitation". Namely, if $s_j \geq s^0$ for all j and Equation (4) in Lemma 1 holds, then the project selection rule from the symmetric benefits benchmark dictates permanent exploitation, and since Equation (4) holds, the players can adopt this rule. In the proposition below we show this is, in fact, a necessary and sufficient condition for permanent exploitation. We also note that permanent exploitation occurs in finite time since, under Assumption 1, there exist project values exceeding $\tilde{s}$.

Proposition 4. *In any optimal relational contract, the players permanently exploit projects with values $\hat{s}_1, \ldots, \hat{s}_m$ if and only if:*

1. *$\hat{s}_j \geq s^0$ for all $j \in \{1, \ldots, m\}$.*

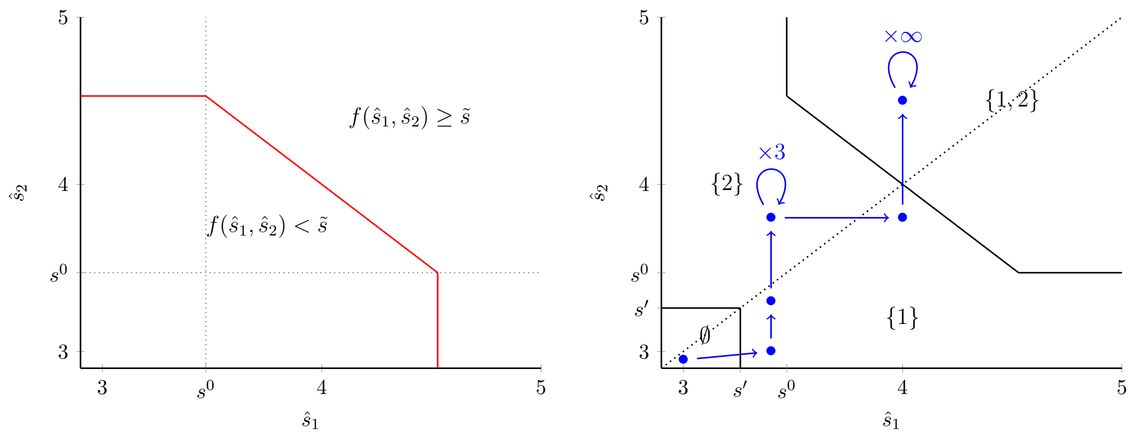
2. *The average of $\hat{s}_1, \ldots, \hat{s}_m$ exceeds $\tilde{s}$.*

We have provided closed-form conditions under which the players can (and thus do) replicate the project selection rule of the symmetric benefits benchmark. Additionally, we have derived the conditions that dictate when projects are selected for permanent exploitation. To delve into the equilibrium dynamics that arise before the players identify a set of m projects suitable for permanent exploitation, we now suppose $m = 2$ and focus on a specific parametric example.

Exponential Distribution Example When $m = 2$. Suppose $c = 1$, $\delta = \frac{1}{3}$, and $s_p \sim 3 - \frac{1}{\lambda} + \text{Exp}(\lambda)$, such that $\mathbb{E}(s_p) = 3$. Further, denote by $\mathcal{C}(\hat{s}_1, \hat{s}_2)$ the continuation value of the players' relationship. We first note that $1 \leq \mathcal{C}(\hat{s}_1, \hat{s}_2)$ for all combinations of $\hat{s}_1$ and $\hat{s}_2$. This inequality holds because the players can always choose to explore two new projects in every period, generating a payoff of $3 - 2$ per project and thus a continuation value of $\frac{\delta}{1-\delta}2$, which reduces to 1 when $\delta = \frac{1}{3}$. Within this example, (i) we utilize Proposition 4 to characterize the players' project selection rule when $f(\hat{s}_1, \hat{s}_2) \geq \tilde{s}$ and (ii) we derive results for the optimal project selection rule when $f(\hat{s}_1, \hat{s}_2) < \tilde{s}$.

Figure 2a displays the threshold s^0 as dotted black lines and the set of $\hat{s}_1$ and $\hat{s}_2$ values that satisfy $f(\hat{s}_1, \hat{s}_2) = \tilde{s}$ in red. This set is the solution to Equation (4) in Lemma 1. Recall that when $f(\hat{s}_1, \hat{s}_2) \geq \tilde{s}$, the players are able to follow the project selection rule of the symmetric benefits benchmark (i.e., the first-best rule). In Figure 2b, we indicate which projects are chosen for exploitation based on the respective values of $\hat{s}_1$ and $\hat{s}_2$. Each region is denoted by the set of dimensions for which the highest-valued project in that dimension is exploited. First, it follows from Figure 2a that both projects are chosen for exploitation when $\tilde{s} \leq f(\hat{s}_1, \hat{s}_2)$ and $\hat{s}_1, \hat{s}_2 \geq s_0$. Outside of this region, we must address two questions: (i) will there be a project selected for exploitation, and (ii) if so, which among the two will be chosen? Evidently, the answer to the second question is the best project. Therefore, in Figure 2b, the selection between $\hat{s}_1$ and $\hat{s}_2$ hinges on which side of the 45 degree line the project values fall. Further, there exists a threshold, s', on the value of the best of the two projects such that, below this threshold, the players choose to explore two new projects rather than exploiting the best of the two projects. Conversely, above this threshold, the players choose to exploit the best of the two projects and explore a new one. Moreover, the threshold s' is independent of the value of the worse of the two projects, because this project will never be exploited. Finally,

Figure 2b also presents a sample path illustrating the evolution of realized project values over time under the optimal relational contract, depicted in blue. In the phase where the players are exploring two projects simultaneously, both $\hat{s}_1$ and $\hat{s}_2$ weakly increase over time. In the phase where the players exploit a project on dimension j, $\hat{s}_j$ remains constant, while $\hat{s}_{-j}$ weakly increases over time. Finally, in the phase of the relationship where the players exploit both projects, $\hat{s}_1, \hat{s}_2$ stay constant because exploitation is permanent. Arrows are used to signify changes in project values when a more valuable project is identified, while self-loops indicate situations where more valuable projects are either not discovered or not pursued.



(a) Feasible Region for First-Best Project Exploitation

(b) Project Exploitation and Sample Path

Figure 2: Optimal Multi-Project Selection Dynamics

In the figure, we assume $c = 1$, $m = 2$, $\delta = \frac{1}{3}$, and $s_p \sim \text{Exp}(\frac{1}{3})$. The values $\hat{s}_1$ and $\hat{s}_2$ represent the values of the best projects discovered by the players to date on cooperative dimensions 1 and 2, respectively. The left figure plots (i) the threshold s^0 such that the players switch from exploration to exploitation in the benchmark case with symmetric benefits and (ii) the set of $\hat{s}_1$ and $\hat{s}_2$ values such that $f(\hat{s}_1, \hat{s}_2) = \tilde{s}$ in red. The right figure plots in Black the project selection behavior of the players under the optimal relational contract. Each region is denoted by the set of dimensions for which the highest-valued project in that dimension is exploited. For example, in the region $\{1, 2\}$, the best project for each dimension is exploited. In region $\{\emptyset\}$, no project is chosen for exploitation. In Blue, we plot one realization of a sample path.

The sample path of realized project values depicted in Figure 2b implies novel dynamics along the equilibrium path, characterized by instances of temporary project exploitation, including of projects with values below s^0. In the second period of their relationship, the players opt to exploit a dimension 1 project (with a value lower than s^0) for two consecutive periods. Subsequently, they abandon it in favor of a dimension 2 project, which they choose to exploit for three consecutive periods before eventually replacing it with a new and superior dimension 1 project.[14]

Building on the previously discussed example, in what follows we assume that $s_p \sim 3 - \frac{1}{\lambda} + \text{Exp}(\lambda)$ without restricting δ to the value $\frac{1}{3}$ or $c = 1$. In this environment, we provide an analytical proof that, beyond the deviations already highlighted in Proposition 4, additional discrepancies from the symmetric benefits benchmark can arise in equilibrium when players decide between project exploitation and exploration.

Proposition 5. *Suppose $s_p \sim 3 - \frac{1}{\lambda} + \text{Exp}(\lambda)$ and $m = 2$. There exist parameter values such that, in equilibrium, the following behaviors occur with positive probability:*

1. *The players exploit a project in period t but not in some period $t' > t$.*
2. *The players exploit a project with value $s_p < s^0$.*
3. *The players choose not to exploit a project in period t, but choose to exploit it in some later period $t' > t$.*

The behaviors described in the proposition do not occur with positive probability across all parameter values, a conclusion that follows when considering that, as δ converges to 1, players follow the project selection rule of the symmetric benefits benchmark. Furthermore, even within the parameter range where these behaviors occur with a positive probability, their occurrence is not guaranteed: for instance, there is always a non-zero probability that the players immediately identify two projects worthy of permanent exploitation.

As argued above, the first statement follows from the second statement combined with the result from Proposition 4, whereby the players only permanently exploit projects whose values exceed s^0. The intuition behind the second statement can be seen by comparing the players' exploration incentives in the presence of symmetric versus asymmetric benefits, and by supposing that s_p is distributed such that the continuation value of the players' relationship exceeds $2c$ with an arbitrarily small

[14] In this scenario, the players never return to previously abandoned projects. However, when $m > 2$, players may opt to exploit a project for several periods, subsequently abandon it, and later revert to its exploitation. Consider, for instance, when $m = 3$. Assume further that players can only cooperate in exploiting a single project, and that they choose to exploit the most valuable one. After several periods, players might identify a project with slightly higher value on another dimension, prompting them to exploit it over the preceding one. Subsequent to additional periods, they may discover a project of such substantial value that they are now able to cooperate in exploiting three projects. At this point, resuming the initial project they exploited might emerge as the optimal choice.

probability. In this case, with a probability approaching 1, the players are able to exploit at most one project, implying the exploration of at least one other project. Consequently, the benefit of exploration stems from the chance to find a more valuable project for exploitation in the next period. However, since the players are already exploring one project, the gain from undertaking a second exploration might be limited: even if two superior projects are identified, they can only exploit one. On the other hand, with projects generating symmetric benefits, the players, upon discovering these two projects, can exploit both, providing them with greater exploration incentives. This higher benefit of exploration, in turn, prompts the players to set a higher exploitation threshold when benefits are symmetric.

Regarding the final statement, consider a value of δ sufficiently small such that the players are unable to cooperate in exploiting projects achieving values slightly exceeding s^0. If in period 0 the players do find two projects with associated values only slightly higher than s^0, the players are compelled to explore two new projects during the next period. However, if one of these new projects happens to achieve a high value, the continuation value of the players' relationship may exceed 2 and the players may wish to return to one of the two period 0 projects.

4.2 The Dynamics of Relationship Scope

We now delve into the analysis of how the two players gradually broaden their relationship scope over time. In Section 4.1, we made an assumption regarding parameter values that ensured that it was always optimal for the players' relationship scope to be maximal at all times.[15] To analyze relationship scope, we relax this assumption. Specifically, we specialize the model presented in Section 2 as follows. First, we assume that $s_p \in \{0, v\}$, where $s_p = v > \frac{2c}{q}$ with probability q, for all projects (i.i.d. across projects). This assumption streamlines the analysis: players opt for exploration upon finding projects with value 0, and always decide to permanently exploit those worth v. Second, we assume $v_{2,p} = 0$. In other words, for any project selected by the players, all output is pocketed by player 1. This approach can be illustrative of, say, an employment setting, where player 1 plays the role of

[15]Specifically, we assumed that $\mathbb{E}(v_{p,i}) > c$, which, in turn, implied that exploring a project selected at random constituted an equilibrium of the stage game. As a result, any optimal relational contract prescribed the exploration (and, eventually, the exploitation) of as many projects as feasible from period 0 onward.

the employer and player 2 assumes the role of the employee. Likewise, this approach can also be applied to a buyer-supplier relationship in which the buyer receives the revenues from the sale of the final product. Because $v_{2,p} = 0$ always, player 2 must now be incentivized not only for project exploitation but also for project exploration. We show that gradualism – that is, delaying the exploration of a project within one dimension of cooperation until previous projects with value v have been found in other cooperation dimensions – can be sustained in equilibrium for a larger range of discount factors, even though it generates less joint surplus compared to when exploring projects along all dimensions at once from the start.[16]

We begin by showing that any (non-empty) optimal relational contract exhibits (weak) growth in the number of projects selected by the players along the equilibrium path. By Proposition 1, we know the optimal relational contract conditions only on the number of projects suitable for exploitation thus far. Denote this number of projects by n and denote by $f(n)$ the number of projects the players cooperate on currently. In equilibrium, the players exploit the n projects, and explore $f(n) - n$ additional projects. If, for example, one of the $f(n) - n$ projects explored in a given period is found to have value v, then in the next period, the players exploit $n + 1$ projects and explore $f(n+1) - (n+1)$ additional projects. The following proposition (i) formalizes why the dynamics that arise under the optimal relational contract depend solely on $f(n)$ when on path and (ii) provides key properties of this function.

Proposition 6. *To any (non-empty) optimal relational contract corresponds a function $f(n)$ specifying that, in any period in which n projects suitable for exploitation have been found, the players exploit these n projects and explore $f(n) - n$ additional projects. Further, $f(n)$ satisfies the following two conditions:*

1. *$f(n)$ is monotonically increasing in n for all n.*
2. *$f(n) \geq n$ for all n, with equality if and only if $n = m$.*

The underlying logic of the first statement is that when players discover high-value projects, the continuation value of their relationship increases since it solely depends on the number of projects with value v identified so far. This raised continuation value

[16] Unlike the analyses in Watson (1999) and Watson (2002) in which the prospect of future increases in relationship scope makes it harder to separate good and bad partners, in our model it makes cooperation easier to sustain.

subsequently enables the players to engage in a weakly larger number of projects, which they choose to do in any optimal relational contract.

The proof for the second statement follows a simple induction argument. Given the assumption of a non-empty relational contract, $f(0) > 0$. Next, assume the statement is true for $n-1$ (i.e., $f(n-1) > n-1$) but fails for n (i.e., $f(n) \leq n$). Combining these inequalities and the first statement of the proposition implies that $f(n) = f(n-1) = n$. Further, we know (i) that the players cannot motivate cooperation on $n+1$ projects given n projects suitable for exploitation and (ii) that they can motivate cooperation on n projects given $n-1$ projects suitable for exploitation:

$$(n+1) \cdot c > n \cdot \mathcal{C}(\text{exploitation}) + \mathcal{C}(\text{exploration}), \tag{5}$$

$$n \cdot c \leq (n-1) \cdot \mathcal{C}(\text{exploitation}) + \mathcal{C}(\text{exploration}). \tag{6}$$

where $\mathcal{C}$ (exploration) (respectively, $\mathcal{C}$(exploitation)) is the per-dimension continuation value when players have not (respectively, have) found a project suitable for exploitation for that dimension. However, since $\mathcal{C}$ (exploration) $< \mathcal{C}$(exploitation), Inequality (5) and Inequality (6) cannot hold jointly.[17]

With symmetric benefits, the scope of the players' relationship is immediately maximal (see Proposition 2). By contrast, Proposition 6 implies that, unless players begin with the broadest scope, they will expand it over time until they reach its maximum extent. In particular, for any $m \geq 2$, we can show that there exists a value of δ low enough such that $f(0) < m$, implying that the relationship "starts small." This is stated below.

Proposition 7. *Define by δ^* the lowest value of δ such that the optimal relational contract is non-empty and by $\bar{\delta}$ the lowest value of δ such that the scope of the players' relationship is maximal at date 0. For any $m \geq 2$, $\delta^* < \bar{\delta}$. Thus, when $\delta \in [\delta^*, \bar{\delta})$, the players' scope strictly increases over time.*

The proof for this proposition is as follows. If there was no increase in the scope of the players' relationship over time, the players would begin with m projects immediately. This project selection rule constitutes an equilibrium if and only if:

$$m \cdot c \leq m \cdot \mathcal{C}(\text{exploration}). \tag{7}$$

[17]One can see this by dividing Inequality (6) by n and Inequality (5) by $n+1$ and noting that the right-hand sides of both resulting inequalities are affine combinations of these two terms.

As δ diminishes, $\mathcal{C}$(exploration) decreases to a point where Inequality (7) fails to hold. Further, such an inequality is independent of m.[18] Instead, if, say, the players begin with a single project and, upon finding a project suitable for exploitation, begin $m-1$ additional projects, the feasibility of such a project selection rule is determined by:

$$m \cdot c \leq (m-1)\mathcal{C}(\text{exploration}) + \mathcal{C}(\text{exploitation}) \tag{8}$$

$$c \leq \mathcal{C}(\text{exploration}) + (m-1) \cdot \mathcal{C}(\text{delayed exploration}) \tag{9}$$

The first constraint ensures that cooperation on m projects after identifying one project suitable for exploitation is feasible. The second constraint ensures that cooperating in exploring one project anticipating the exploration of $m-1$ projects immediately after identifying a first project suitable for exploitation is feasible. Either expression is strictly easier to satisfy than Inequality (7). Hence, when δ is high enough such that a gradual contract (although not necessarily the one described above) is feasible, (i.e., exceeds δ^*), but low enough such that maximal scope is not (i.e., less than $\bar{\delta}$), the players' scope will gradually increase over time.

Finally, we note that the specific project selection rule used for the intuition in the proof of Proposition 7 is one of many potential "gradual" options. For instance, there exists a "very gradual" project selection rule where $f(n) = \min(m, n+1)$ in which the players explore at most one new project per-period. Such a project selection rule is not always optimal. Proposition 1 implies that the players maximize their continuation value at each point on the equilibrium path. Hence, if upon finding n projects suitable for exploitation, the players are able to explore two or more new projects, then they will do so. While fully characterizing $f(n)$ is beyond the scope of the paper, we can provide a result in support of this intuition.[19] Namely, if q is low, the increase in the continuation value of the players' relationship that occurs upon finding a project suitable for exploitation can be large enough to enable the players to start cooperating on several additional projects simultaneously.

---

[18]Throughout, we only consider the most binding constraints, because these are the only relevant ones.

[19]Fully characterizing $f(n)$ involves solving a Bellman Equation as a function of n projects suitable for exploitation. For a given $f(n)$, $B(n) = n(v-2) + (f(n)-n)(q \cdot v - 2) + \delta \sum_{j=0}^{f(n)-n} B(n + \text{bin}(f(n)-n, q)\,[j])$. Next, one must check that the Bellman Equation's associated continuation value, $\mathcal{C}(n) = \delta \sum_{j=0}^{f(n)-n} B(\text{bin}(f(n)-n), q)\,[j] + n)$, satisfies $c \cdot f(n) \leq \mathcal{C}(n)$. Finally, one analyzes all possible $f(n)$ functions satisfying such a constraint and chooses the one that maximizes the payoffs of the players.

Proposition 8. *For any $\alpha \in \mathbf{N}$, there exists a q^* such that, for all $q < q^*$, any (non-empty) optimal relational contract satisfies $f(n) \geq \min(f(n-1) + \alpha, m)$ for all $n \in \{1, \ldots, m\}$.*

In words, when projects suitable for exploitation are extremely valuable—particularly when q is low and v has to be large to maintain $q \cdot v > 2 \cdot c$, upon finding a project worth exploiting, the players can increase their relationship scope by at least α dimensions in the subsequent period. The intuition for this proposition is simply that, when q is small, finding a project suitable for exploitation generates a large increase in the continuation value of the players' relationship. The players, in turn, leverage this increase in continuation value by widening the scope of their relationship. When $\alpha \geq 2$, early successes compound as each additional project suitable for exploitation allows the players to explore two or more additional projects.

5 Extensions and Applications

In this section, we begin by providing additional results regarding the role played by the dimensionality m of the players' relationship. Next, we explore two extensions of the main model, incorporating projects with both symmetric and asymmetric benefits. We show that players might prefer symmetric projects due to their straightforward implementation, and that symmetric projects can help players in the initial phases of their relationship while exploring asymmetric projects.[20] Finally, we discuss how our results shed light on various concrete settings characterized by informal collaborative dynamics.

5.1 On the Benefits of Scope

In this subsection, we provide results regarding the role played by the dimensionality m of the players' relationship. To this end, denote by $\frac{\pi(m)}{m}$ the average joint surplus of the relationship per dimension of the relationship. Similarly, denote by $\delta^*(m)$ the minimum discount factor for which the optimal relational contract is non-empty. Throughout the analysis, the following two weak inequalities hold: $\frac{\pi(m \cdot k)}{m \cdot k} \geq \frac{\pi(m)}{m}$ and

[20]These results bear resemblance to Acharya and Ortner (2022), where two players involved in collective search prioritize projects that benefit both players.

$\delta^*(m \cdot k) \leq \delta^*(m)$. In words, raising the dimensionality of the relationship by a factor of k can only increase the average surplus of the relationship per dimension and the range of discount factors over which the players' optimal relational contract is non-empty (namely, there exists a history h^t such that $\mathbf{P}^t \neq \emptyset$). The intuition behind these weak inequalities is simple: the players can always engage in k independent relationships, each characterized by a dimensionality of m and replicating the optimal relational contract for m dimensions.

In the environment considered in Section 4.1.2, where exploration can always be sustained in equilibrium, $\delta^*(m) = 0$. However, one can show that $\frac{\pi(m \cdot k)}{m \cdot k} \geq \frac{\pi(m)}{m}$ holds as a strict inequality in the range of parameter values such that the players are unable to replicate the project selection rule of the symmetric benefits benchmark. In these instances, pooling relational incentives across dimensions and projects is valuable to the players. This is most easily seen when we consider that for m projects to be selected for permanent exploitation, only their average value needs to exceed the threshold $\tilde{s}$, instead of each individual project having a value exceeding $\tilde{s}$.

In the environment considered in Section 4.2, where exploration is not an equilibrium of the stage game, $\delta^*(m \cdot k) < \delta^*(m)$. To see why this inequality holds strictly, recall that, when $\delta = \delta^*(m)$, the players' optimal relational contract is necessarily gradual. The players could adopt k separate gradual relational contracts. However, doing so would be inefficient because it would condition further project exploration exclusively on the number of projects suitable for exploitation found within an (artificially segmented) separate relational contract. If instead players pool their relational incentives across all dimensions, they are able to sustain a gradual relational contract for a larger range of parameter values. By the exact same argument, $\frac{\pi(m \cdot k)}{m \cdot k} > \frac{\pi(m)}{m}$ whenever the optimal relational contract is gradual (i.e., δ is intermediate).[21]

Note that the benefits of greater scope are not due to asymmetries across dimensions as in Bernheim and Whinston (1990) or from an improved ability to design relational contracts in the face of moral hazard as in Levin (2002), but rather from the players' ability (i) to utilize gradualism more effectively and (ii) to pool incentives across ex ante identical but ex-post asymmetric dimensions of cooperation.

[21]While $\pi(m)$ is monotonically increasing in m, $\frac{\pi(m)}{m}$ may not. For example, if s_p comes from a three-point support (low, medium, and high), and if a high-valued project and a single medium-valued project can be jointly exploited by the players, but a high-valued project and two medium-valued projects cannot, then $\pi(m)$ would depend on the parity of m and monotonicity would break.

5.2 Favoring Symmetric over Asymmetric Projects

We now modify the setting to allow for projects with both symmetric and asymmetric benefits across the players. We show that players exhibit less selectivity for projects with symmetric benefits compared to asymmetric ones. Additionally, the existence of asymmetric projects diminishes selectivity for symmetric projects compared to when only symmetric projects are available. As before, we assume that $s_p \sim F$. However, we also suppose that: with probability $1-q$, both players have valuations $v_{p,1} = v_{p,2} = \frac{s_p}{2}$; with probability $\frac{q}{2}$, player 1 values the project at $v_{p,1} = s_p$ and player 2 at $v_{p,2} = 0$; and with the same probability $\frac{q}{2}$, player 2 values the project at $v_{p,2} = s_p$ and player 1 at $v_{p,1} = 0$. We assume that the distribution of benefits is i.i.d. across projects. Our assumptions regarding the distribution of project values imply that project exploration is an equilibrium of the stage game, as in Section 4.1. For simplicity, we also suppose that $m = 1$. Setting $q = 1$ thus corresponds to the version of model considered in Section 4.1.1. By contrast, setting $q = 0$ corresponds to a special case of the benchmark model with symmetric benefits analyzed in Section 3.2. In the case where $q \in (0, 1)$, two distinct exploration/exploitation thresholds exist, s_s^* and s_a^*, where the first threshold applies to projects with symmetric benefits, while the second one applies those with asymmetric benefits. The key finding that emerges from this analysis is that $s_a^* \geq s_s^*$, indicating that the players are less selective when it comes to projects with symmetric benefits, as compared to those without. Additionally, we show that: $s_a^* \geq s^0 \geq s_s^*$. The intuition behind $s_a^* \geq s^0$ is the same intuition as before: a project that benefits just a player must be valuable enough to enable cooperation in exploitation. The intuition behind $s^0 \geq s_s^*$ is as follows. If the players identify a project with value $s_p \in (s^0, s_a^*)$, they can exploit it only if it yields symmetric benefits. As a result, the overall value of exploration is lower for the players, leading to lower exploration/exploitation thresholds for projects with symmetric benefits.

5.3 Symmetric Projects as Stepping Stones in Relationship Building

We now explore the dynamics of the scope of the players' relationship when projects with both symmetric and asymmetric benefits are available. Suppose that the distribution of project benefits remains as in Section 4.2, namely $v_{p,2} = 0 \, \forall p \in \mathcal{P}$.

However, the players now also have access to m projects (each associated with a distinct dimension) with guaranteed benefits of v_c for both players. We assume that $2c < 2v_c < \mathbb{E}(s_p)$ to ensure that these projects are profitable, but less so (in expectation) compared to those chosen from the set $\mathcal{P}$. With a proof almost identical to that for Proposition 4, one can show that the players' relationship is characterized by a function $f(n)$ where n denotes the number of asymmetric projects identified as suitable for exploitation. In each period, the players explore $f(n) - n$ asymmetric projects, they exploit n asymmetric projects, and they exploit $m - f(n)$ symmetric projects, which implies scope is always maximal. Exactly as before, the function $f(n)$ is increasing, meaning that the players choose fewer projects with symmetric benefits as time passes. Finally, the presence of projects with symmetric benefits increases the value of the players' relationship for two related reasons: (i) the players' relationship is maximal from the very beginning and (ii) the presence of valuable projects with symmetric benefits allows the players to begin exploring and subsequently exploiting the more profitable asymmetric projects earlier. In this sense, the symmetric projects act as stepping stones in the building of the players' relationship.

5.4 Applications

We have shown that when asymmetric benefits are present among collaborating parties and mutual trust is initially low, collaborative parties may find themselves unable to cooperate across all potential collaborative domains at the outset. As a result, they may find it necessary to gradually expand the scope of their collaboration, leveraging early successes as stepping stones to broaden their relationship. Moreover, our analysis indicates that, on average across various collaborative domains, these parties will engage in a prolonged period of experimentation to define the exact nature of their collaboration. This protracted exploration phase is essential for the parties to identify collaborations with substantial value, so that no party will find it in their interest to withdraw their cooperation in the long run.

In buyer-supplier dynamics, suppliers often undertake non-contractible investments that tend to benefit the buyers. Toyota's relationship with its suppliers is a well-documented example of a gradual and experimental approach, where Toyota encouraged its suppliers to incrementally adopt the Toyota Production System—a strategy that initially yielded benefits primarily for Toyota through improvements in quality and efficiency (Dyer and Nobeoka, 2000). Beyond the Toyota example, and as discussed in Section 1, buyer-supplier relationships often exhibit gradualism and lengthy experimentation. For example, analyzing a buyer-supplier relationship involving soy bean products, Vanpoucke et al. (2014) observe, "the buyer decided to purchase from this supplier. This was the start of the exploration stage. ... it took about 10 years before the two parties started up a first integration initiative and entered the expansion stage. It then took a long time for both partners to get to know each other, build enough trust and see the benefits of learning from each other's expertise and further building up the relationship."[22]

Gradualism and extended experimentation manifest in contexts beyond buyer-supplier relationships. Consider, for instance, the dynamics of employee task allocations. Although an employee may be hired with the intention of working on multiple tasks, a manager may initially need to allocate only a limited number of tasks to the new employee. This allows the employee to dedicate non-contractible effort to exploring various methods for each initial task, experimenting until a satisfactory routine is identified, before being assigned additional responsibilities. As Carucci (2018) underlines, onboarding employees is challenging, because of the need to establish trust and convey organizational knowledge. Carucci advises managers to "Start with targets you are confident your new hires can meet. If all goes well, gradually increase the level of responsibility associated with each task."[23] Similarly, the gradual nature of collaboration is often exemplified in the formation of political unions, in which countries decide to cooperate across multiple domains by sharing resources, with the freedom to withdraw at any given moment. For example, the establishment of the European Union was characterized by significant step-by-step policy implementation and a gradual deepening of integration among its member states (see, e.g., Spolaore, 2015, and references therein). Initially, the European Union prioritized projects that conferred mutual advantages to all member states, subsequently transitioning to adopt more ambitious policies where costs and benefits were more unevenly distributed. Another noteworthy example of prolonged experimentation pertains to climate change mitigation policies. In such policies, the net benefits are not uniformly distributed across countries. To address this imbalance, funds and mechanisms are often established to

[22]See Vanpoucke et al. (2014) for additional examples of slow and gradual relationship building in buyer-supplier collaborations.

[23]For further illustrations of the advantages of adopting a gradual approach in manager-employee relationships, refer to Ye et al. (2020) and the cited references therein.

compensate countries that benefit the least from these policies (Pickering et al., 2015). Moreover, climate agreements are continuously modified and involve significant trial-and-error, serving the dual purpose of enhancing policy-making and ensuring ongoing cooperation among nations (Falkner, 2016).[24]

We have also shown that collaborative parties are motivated to pool their relational incentives across all dimensions of their collaboration. These "multilateral" relational incentives, in turn, lead (i) to inter-dependencies across dimensions of co-operation and (ii) the players to revisit previously abandoned projects and to engage in temporary project exploitation. To illustrate these dynamics, consider two firms that form an R&D alliance to share their resources for different research areas. Their partnership involves making repeated investments throughout the life cycle of their joint projects, which are too complex or unverifiable to be fully written out in a contract. The two firms, having experienced significant success in a few key collaborations, build up enough trust to facilitate cooperation on projects in other areas that may be less profitable and with an unequal distribution of benefits, but also demand less development time. These opportunities may be actively sought out, or they may have been identified previously and can now be effectively capitalized upon. Similarly, when there is limited mutual trust, making it challenging to collaborate on multiple projects that mostly benefit one firm, the firms might settle on a specific collaboration within one area for an extended period. However, they may later opt to discontinue it in favor of pursuing a more lucrative collaboration in a different domain. Concurrently, they may resume searching for superior collaboration opportunities in the first domain and eventually reach a point where they can cooperate across both areas.

6 Concluding Remarks

This paper has introduced a framework for analyzing collaborative dynamics, shedding light on the process through which parties build the trust needed for collaborative success and on how this trust, in turn, contributes to the achievement of successful collaborations. The model generates three key insights. First, as trust is

[24] Gradualism has also been documented in tacit price collusion, with firms slowly softening price competition through experimentation and incremental price increases (Byrne and De Roos, 2019). Relatedly, Chilet (2018) shows that pharmacy chains in Chile gradually expanded the scope of their collusion across multiple product spaces.

intricately tied to the value of the players' relationship, collaborating parties invest a substantial amount of time in seeking collaborations that are sufficiently valuable to foster trust-based cooperation. Secondly, when collaborations involve multiple parallel projects, parties combine their relational incentives across projects, resulting in interdependencies and significant path dependence. In this context, while the average project must provide enough value to support trust-based cooperation, particularly valuable projects can be leveraged to reduce inefficiencies in the selection of other projects. Moreover, "relational inter-dependencies" between projects can lead to seemingly erratic behaviors, such as prolonged cooperation on projects that are ultimately discontinued, or the revival of previously abandoned projects, all driven by the time required for the parties to establish trust in their relationship. Third, in situations where initial credibility is low, the parties cannot immediately realize the full potential scope of their relationship. To build the trust needed for sustaining cooperation across multiple projects, the parties may begin by collaborating on a limited number of projects. They can then leverage early successes in these projects to embark on additional ones.

Our model serves as a valuable lens for examining dynamics occurring within firms. In particular, our work contributes to the ongoing debate on persistent performance differences among seemingly similar enterprises. Numerous empirical studies have documented these enduring disparities in performance across a range of industries and countries, and these gaps have proven surprisingly robust against plausible explanations such as market competition, local geographical and demand conditions, or access to human capital (see Syverson, 2011; Gibbons and Henderson, 2013, and references therein).

According to Gibbons and Henderson (2013), and the body of evidence they review, variations in managerial practices are key in creating productivity disparities across firms. We adapt for our purposes their categorization of explanations: (i) managers might either be unaware of their shortcomings, or, even if conscious of them, believe that the best practices from other firms are not suitable for their context; (ii) managers are aware of their lag and are able to seek superior managerial practices suitable to their context, but opt not to; and (iii) managers are "striving mightily" to adopt superior practices but face hurdles during the implementation phase. The first explanation underscores information barriers, prompting questions about why such information does not diffuse more readily (c.f. Bloom et al., 2013; Atkin et al., 2017).

The second explanation aligns with the framework developed by Chassang (2010) and discussed in Section 1.1, in which players are aware of the existence of more efficient practices but choose not to pursue them.

Our model offers insight into the third explanation presented by Gibbons and Henderson (2013). Within our framework, players not only are aware of the potential for improvements in their collaborative endeavors but are also actively in pursuit of them. However, they must identify practices of sufficient value such that all stakeholders find it in their best interest to participate in their implementation. As a result, adopting superior practices is a time-intensive endeavor and may prompt players to focus on refining specific facets before exploring additional areas of improvement. Thus, elements of luck and path dependence can lead to large performance differences, even across collaborations that started under similar circumstances. We illustrate this phenomenon with an example from the environment described in Section 4.2. Imagine a scenario where players have the capacity to undertake up to 20 projects (in the language of the model, their relationship has 20 dimensions). Each project carries a 50% chance of having a value of \$4, and a 50% chance of holding no value and entails a cost of 1 to both parties. With a discount factor equal to .357, the players optimally begin their collaboration by working on two projects. As the players identify valuable projects, the value of their relationship increases and they are able to collaborate on additional projects. Early successes have a compounding effect: identifying valuable projects may allow players to expand their collaborative efforts by more than just a single project. Small differences early on can thus yield large persistent differences in the medium-term, before differences reduce as collaborations approach their maximum dimensionality of 20 projects. Figure 1 reports the expected joint surplus of a collaboration as time passes. It contrasts a collaboration that identified one valuable project in period 0 (in blue) with another that identified none (in red). Emphasizing the persistent and strong impact small initial differences can wield, the figure shows that the blue collaboration is expected to outperform the red collaboration, with the gap widening for an extended duration. Moreover, it is anticipated to take the red collaboration upwards of 14 periods to recover and match the performance of the blue one.

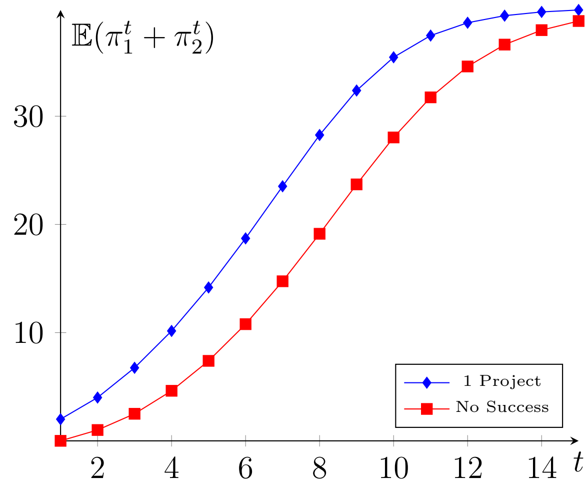


Figure 3: Persistent Performance Differences

The figure plots the expected joint collaborative surplus in period t under the optimal equilibrium described in Proposition 6 when $m = 20, v = 4, q = \frac{1}{2}, c = 1$ and $\delta = .357$. The blue curve assumes that the players have identified 1 valuable project in period 0. The red curve assumes that the players have identified 0 valuable project in period 0.

Further, Gibbons and Henderson (2013) highlight that synergies between managerial practices, such as those underlying Lean Manufacturing and Total Quality Management (Powell, 1995; Shah and Ward, 2003), can complicate a firm's effort to integrate new methods (c.f. Milgrom and Roberts, 1990; Rivkin, 2000). This complexity arises from the potential need—and consequent prohibitive cost—of adopting an entire bundle of practices concurrently, leading to significant productivity disparities among firms. Our findings indicate that managerial practices can display strong complementarities, even in the absence of clear interdependencies, when based on relational contracts involving the same employees. We show that practices with such relational interdependencies are more easily implemented in a gradual manner, and, for this reason, can also lead to lasting performance gaps among firms.

References

Abreu, Dilip, "Extremal equilibria of oligopolistic supergames," *Journal of Economic Theory*, 1986, *39* (1), 191–225.

Acharya, Avidit and Juan Ortner, "Paths to the Frontier," *American Economic Journal: Microeconomics*, 2022, *14* (1), 39–69.

Atkin, David, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen, "Organizational Barriers to Technology Adoption: Evidence from Soccer-Ball Producers in Pakistan*," *The Quarterly Journal of Economics*, 03 2017, *132* (3), 1101–1164.

Baker, George, Robert Gibbons, and Kevin J. Murphy, "Subjective Performance Measures in Optimal Incentive Contracts," *The Quarterly Journal of Economics*, 1994, *109* (4), 1125–1156.

_ , _ , and _ , "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics*, 2002, *117* (1), 39–84.

Bergemann, Dirk and Juuso Valimaki, "Bandit problems," 2006.

Bernard, Andrew B. and Andreas Moxnes, "Networks and Trade," *Annual Review of Economics*, 2018, *10* (1), 65–85.

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

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts, " Does Management Matter? Evidence from India *," *The Quarterly Journal of Economics*, 11 2013, *128* (1), 1–51.

Bolton, Patrick and Christopher Harris, "Strategic experimentation," *Econometrica*, 1999, *67* (2), 349–374.

Bull, Clive, "The Existence of Self-Enforcing Implicit Contracts," *The Quarterly Journal of Economics*, 1987, *102* (1), 147–159.

Byrne, David P and Nicolas De Roos, "Learning to coordinate: A study in retail gasoline," *American Economic Review*, 2019, *109* (2), 591–619.

Carucci, Ron, "To retain new hires, spend more time onboarding them," *Harvard Business Review*, 2018, pp. 1–5.

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

Chilet, Jorge Alé, "Gradually rebuilding a relationship: The emergence of collusion in retail pharmacies in Chile," 2018.

Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh, "Developing Buyer-Seller Relationships," *Journal of Marketing*, 1987, *51* (2), 11–27.

Dyer, Jeffrey H and Kentaro Nobeoka, "Creating and managing a high-performance knowledge-sharing network: The Toyota case," *Strategic Management Journal*, 2000, *21* (3), 345–367.

Falkner, Robert, "The Paris Agreement and the new logic of international climate politics," *International Affairs*, 2016, *92* (5), 1107–1125.

Ghosh, Parikshit and Debraj Ray, "Cooperation in Community Interaction without Information Flows," *The Review of Economic Studies*, 1996, *63* (3), 491–519.

Gibbons, Roberts and Rebecca Henderson, "What Do Managers Do? Exploring Persistent Performance Differences Among Seemingly Similar Enterprises," *Handbook of Organizational Economics*, 2013, pp. 342–372.

Gutierrez, Agustin, Ashish Kothari, Carolina Mazuera, Tobias Schoenherr, and McKinsey Company, "Taking supplier collaboration to the next level," *McKinsey Company*, 2020.

Halac, Marina, "Relationship building: conflict and project choice over time," *The Journal of Law, Economics, & Organization*, 2014, *30* (4), 683–708.

_ , "Investing in a Relationship," *The RAND Journal of Economics*, 2015, *46* (1), 165–185.

Kranton, Rachel E., "The Formation of Cooperative Relationships," *Journal of Law, Economics, Organization*, 1996, *12* (1), 214–233.

Levin, Jonathan, "Multilateral contracting and the employment relationship," *The Quarterly Journal of Economics*, 2002, *117* (3), 1075–1103.

— , "Relational Incentive Contracts," *American Economic Review*, June 2003, *93* (3), 835–857.

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.

Macleod, W. Bentley and James Malcomson, "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment," *Econometrica*, 1989, *57* (2), 447–80.

Milgrom, Paul and John Roberts, "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica: Journal of the Econometric Society*, 1990, pp. 1255–1277.

Monarch, Ryan, "It's Not You, It's Me: Prices, Quality, and Switching in U.S.-China Trade Relationships," *The Review of Economics and Statistics*, 09 2022, *104* (5), 909–928.

— **and Tim Schmidt-Eisenlohr**, "Longevity and the Value of Trade Relationships," 2023. Working Paper.

Pickering, Jonathan, Frank Jotzo, and Peter J Wood, "Structuring international financial support for climate change mitigation in developing countries," *Nature Climate Change*, 2015, *5* (11), 935–941.

Powell, Thomas C., "Total quality management as competitive advantage: A review and empirical study," *Strategic Management Journal*, 1995, *16* (1), 15–37.

Rivkin, Jan W., "Imitation of Complex Strategies," *Management Science*, 2000, *46* (6), 824–844.

Shah, Rachna and Peter T Ward, "Lean manufacturing: context, practice bundles, and performance," *Journal of Operations Management*, 2003, *21* (2), 129–149.

Spolaore, Enrico, "The Political Economy of European Integration," NBER Working Paper 21250, National Bureau of Economic Research, June 2015.

Strulovici, Bruno, "Learning while voting: Determinants of collective experimentation," *Econometrica*, 2010, *78* (3), 933–971.

Syverson, Chad, "What Determines Productivity?," *Journal of Economic Literature*, June 2011, *49* (2), 326–65.

Urgun, Can, "Restless Contracting," 2021. Working Paper.

Vanpoucke, Evelyne, Ann Vereecke, and Kenneth K. Boyer, "Triggers and patterns of integration initiatives in successful buyer–supplier relationships," *Journal of Operations Management*, 2014, *32* (1), 15–33.

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.

Weitzman, Martin, "Optimal Search for the Best Alternative," *Econometrica*, 1979, *47* (3), 641–654.

Ye, Maoliang, Jie Zheng, Plamen Nikolov, and Sam Asher, "One Step at a Time: Does Gradualism Build Coordination?," *Management Science*, 2020, *66* (1), 113–129.

7 Appendix A

Proof of Proposition 1. Recall that after a deviation in period t, players set $P_i^t = \emptyset$ and $b_{i,-i}^t = 0$ 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 project selection rule $\mathbf{p}(\cdot)$; (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 $t-1$ $b_{i,j}(\cdot)$ transfers associated with history h^t (and, thus, corresponding to a specific realization of $\mathbf{v}^{t-1}$) are adjusted so that: (i) player 2's expected payoff following the realization of $\mathbf{v}^{t-1}$ is the same as under the original equilibrium and (ii) player 1's expected payoff following the realization of $\mathbf{v}^{t-1}$ increases by $\sum_i \mathcal{C}_i^2 - \sum_i \mathcal{C}_i^1$. Specifically, take the vector of transfers $\mathbf{b}_1 = (b_{1,2}^1, b_{2,1}^1)$ associated with the original equilibrium and create a new vector of transfers $\mathbf{b}_2 = (b_{1,2}^2, b_{2,1}^2)$ such that:

$$\mathcal{C}_1^2 + b_{2,1}^2 - b_{1,2}^2 > \mathcal{C}_1^1 + b_{2,1}^1 - b_{1,2}^1, \tag{10}$$

$$\mathcal{C}_2^2 + b_{1,2}^2 - b_{2,1}^2 = \mathcal{C}_2^1 + b_{1,2}^1 - b_{2,1}^1. \tag{11}$$

Because $\sum_i \mathcal{C}_i^2 - \sum_i \mathcal{C}_i^1 > 0$, finding payments that satisfy $b_{1,2}^2 \leq \mathcal{C}_1^2$ and $b_{2,1}^2 \leq \mathcal{C}_2^2$ 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 a project selection rule $\mathbf{P}(\cdot)$ if and only if the following inequality holds for all t and for all histories $h^t \in \mathcal{H}^t$:

$$\sum_{p \in \mathbf{P}^t} \sum_{i=1,2} \max \Big(0, c - \mathbb{E}(v_{i,p}|h^t)\Big) \leq \mathcal{C}(h^t), \tag{12}$$

where $\mathcal{C}(h^t)$ is the continuation value.

To show that (12) is a necessary and sufficient condition, consider a set of transfers $b_{i,-i}(\mathbf{v}^t) \geq 0$ to be paid on path given a vector of realized values $\mathbf{v}^t$.

Given an equilibrium project selection $\mathbf{P}^t$, 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 $\alpha_i(\mathbf{v}^t)$ denote player i's share of $\mathcal{C}(h^t \sqcup \mathbf{v}^t)$ as a function of $\mathbf{v}^t$. Hence, the condition for selecting $\mathbf{P}^t$ is:

$$\sum_{p \in \mathbf{P}^t} \max \big(c - \mathbb{E}(v_{i,p}|h^t), 0\big) \leq \mathbb{E}\Big(b_{-i,i}(\mathbf{v}^t) - b_{i,-i}(\mathbf{v}^t) + \alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t)\Big), \quad \forall i, \tag{13}$$

$$b_{i,-i}(\mathbf{v}^t) \leq \alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t), \quad \forall \mathbf{v}^t, \forall i. \tag{14}$$

Expectations are taken over the project valuations realizations $\mathbf{v}^t$ and $h^t \sqcup \mathbf{v}^t$ denotes the players' updated beliefs after observing $\mathbf{v}^t$.[25] 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 the each player is willing to pay the other player the necessary transfer.

To show necessity: Note that since Equation (13) 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, $\mathbb{E}(\mathcal{C}(h^t \sqcup \mathbf{v}^t)) = \mathcal{C}(h^t)$. Hence, we are left with Equation (12).

To show sufficiency: We will show this result in two substeps.

SubStep 1: We show it is necessary and sufficient to replace Equation (14) by

[25] 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 values as every other object can be inferred on path from the realized values.

its expectation. This new expression is as follows:

$$\mathbb{E}(b_{i,-i}(\mathbf{v}^t)) \leq \mathbb{E}\Big(\alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t)\Big) \quad \forall i. \tag{15}$$

We first show that if there is a solution to equations (15) and (13) then there exists a solution to equations (14) and(13).

Take a set of transfers $b_{i,-i}(\mathbf{v}^t)$ that satisfy equations (15) and (13). Define:

$$b'_{i,-i}(\mathbf{v}^t) = \alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t) - \Bigg(\mathbb{E}\Big(\alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t) - b_{i,-i}(\mathbf{v}^t)\Big)\Bigg). \tag{16}$$

Since Equation (15) holds, the term in the expectation of equation (16) is positive and thus Equation (14) holds for all realizations of $\mathbf{v}^t$ under the set of transfers $b'_{i,-i}(\mathbf{v}^t)$. Finally, $\mathbb{E}(b'_{i,-i}(\mathbf{v}^t)) = \mathbb{E}(b_{i,-i}(\mathbf{v}^t))$ so Equation (15) continues to hold.

SubStep 2: Using substep 1, it suffices to show that Equation (12) implies a solution to Equations (13) and (15). To simplify all the notation with expectations, Equation (13) can be re-expressed as:

$$\beta_i - \gamma_i \leq (\tilde{b}_{-i,i} - \tilde{b}_{i,-i}), \tag{17}$$

where $\tilde{b}_{i,-i}$ is the expected transfer from i to $-i$, $\beta_i = \sum_{p \in \mathbf{P}^t} \max\big(0, c - \mathbb{E}(v_{i,p}|h^t)\big)$, and $\gamma_i = \mathbb{E}(\alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t))$. Equation (15) can thus be re-written as:

$$\tilde{b}_{i,-i} \leq \gamma_i. \tag{18}$$

Rearranging Equation (12) implies $\sum_i(\beta_i - \gamma_i) \leq 0$. One can now show that $\tilde{b}_{i,-i} = \max(0, \beta_{-i} - \gamma_{-i})$ satisfies Equations (18). Further, Equation (17) holds because:

$$\beta_i - \gamma_i \leq \max(0, \beta_i - \gamma_i) - \max(0, \beta_{-i} - \gamma_{-i}) \tag{19}$$

$$\iff \max(0, \beta_{-i} - \gamma_{-i}) - \min(0, \gamma_i - \beta_i) \leq 0 \tag{20}$$

$$\Longleftarrow \sum_i (\beta_i - \gamma_i) \leq 0, \tag{21}$$

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

Finally, Equation (18) reduces to

$$\max(0, \beta_{-i} - \gamma_{-i}) \leq \gamma_i \impliedby \beta_{-i} - \gamma_{-i} \leq \gamma_i \tag{22}$$

$$\impliedby \sum_i (\beta_i - \gamma_i) \leq 0, \tag{23}$$

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

Step 3: We show that any relational contract that implements a given project selection rule can be replaced by an alternative relational contract that implements the same project selection rule and yields no surplus to player 2.[26] First, note that the way the players share their continuation value does not affect Equation (2). 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 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 $\mathcal{C}(h^t)$ with $\mathcal{C}(\mu^t)$. □

Proof of Proposition 2.

Recall from the text that one can ignore Equation (2) when analyzing the equilibrium that maximizes the joint surplus of the players. As such, since there are no interdepenencies across the dimensions and the players maximize joint surplus, the players will treat each dimension identically and symmetrically.

Statement 1: Note that all projects within the interval $[j, j + 1)$ are ex ante

[26]Of course, one could take the relational contract derived from Step 3 and 4 and choose to redistribute the surplus by an up-front payment every period from player one combined with reducing the expected payment from player one at the end of each period.

identical. Hence, if the players ever find it optimal to explore a project, then in every period the players will select a project from dimension j. This implies that for each dimension the players either choose no projects in all periods or a project in every period.

Statement 2: Given that there exists a continuum of ex-ante identical projects, the optimal relational contract conditions only on the project with the highest value amongst all previously explored projects, whose value we denote $\hat{s}$. In particular, one can write the Bellman Equation:

$$B^0(\hat{s}) = \max_{\text{explore, exploit } \hat{s}} \{\mathbb{E}(s') - 2c + \delta \mathbb{E}(B^0(\max(\hat{s}, s'))), \hat{s} - 2c + \delta B^0(\hat{s})\}. \tag{24}$$

The first term in the maximum operator corresponds to the players' expected surplus when exploring one more project (chosen at random, since all unexplored projects are ex ante identical) and the second term is their surplus when exploiting the best project found thus far. Next, one can show there exists a threshold s^0, wherein the players explore if $\hat{s} < s^0$ and exploit if $\hat{s} \geq s^0$. Finally, for any $\delta < 1$, one can use Blackwell's Sufficient Conditions to show that there exists a unique solution to the Bellman Equation, and hence the threshold rule dictated by s^0 is a solution. □

Proof of Figure 1. When $s_p \sim \text{Exp}(\lambda)$ and benefits are symmetric, we can compute the threshold s^0 such that the players are indifferent between exploitation and exploration:

$$\frac{s^0}{1-\delta} = \mathbb{E}(s_p) + \frac{\delta}{1-\delta}\mathbb{E}\left(\max\{s_p, s^0\}\right) \tag{25}$$

$$\iff \frac{s^0}{1-\delta} = \frac{1}{\lambda} + \frac{\delta}{1-\delta}\left(e^{-\lambda s^0}\left(s^0 + \frac{1}{\lambda}\right) + \left(1 - e^{-\lambda s^0}\right)s^0\right). \tag{26}$$

The left-hand side corresponds to the exploitation surplus. The right-hand side corresponds to the expected surplus when exploring one more time and subsequently exploiting the best project found until then. The second step utilizes the expected value of the exponential and computes the expected value of the maximum operator conditional on whether $s^0 < s_p$ or $s_p < s^0$, respectively.

Solving this expression for s^0 when $\lambda = \frac{1}{3}$ yields the equation for s^0 provided in the text. Finally, solving for $\tilde{s}$ was done in the text. □

Proof of Lemma 1. Recall that, under any optimal relational contract, continuation play depends only on the values of the best project found for each dimension to date.

When the players have identified projects with values $\hat{s}_1, \dots, \hat{s}_m$ at history h, the condition for the players being able to replicate the project selection rule of the symmetric benefits benchmark in all subsequent periods is that, for all histories h' occurring after h and with associated project values $\hat{s}'_1, \dots, \hat{s}'_m$, the players exploit $\hat{s}'_j$ if and only if $\hat{s}'_j \geq s^0$. This condition is as follows:

$$c \sum_{j=1}^{m} \mathbf{1}_{\hat{s}'_j \geq s^0} \leq \delta \left(\sum_{j=1}^{m} \mathbf{1}_{\hat{s}'_j \geq s^0} \frac{1}{1-\delta} (\hat{s}'_j - 2) + \mathbf{1}_{\hat{s}'_j < s^0} \left(\frac{1}{1-\delta} (\hat{s}^0 - 2) + \frac{s^0 - \mathbb{E}(s)}{\delta} \right) \right), \quad \forall (\hat{s}'_1, \dots, \hat{s}'_m) \geq (\hat{s}_1, \dots, \hat{s}_m), \tag{27}$$

which corresponds to Equation (2) when the players follow the symmetric benefits benchmark. The right-hand side is the sum of the continuation values due to exploitation and exploration. The value of exploration is derived from noting that the players are indifferent between exploring and exploiting a project with value s^0. Further, as the value of the best project cannot decrease, the condition must hold for all values of $(\hat{s}'_1, \dots, \hat{s}'_m)$ where $\hat{s}'_i \geq \hat{s}_i$. Finally, one can note that setting $\hat{s}'_i = \max\{\hat{s}_i, s^0\}$ both minimizes the right-hand side and maximizes the left-hand side of Equation (27). Thus, an equivalent condition is:

$$m \cdot c \leq \delta \Big(\sum_{j=1}^{m} \frac{1}{1-\delta} (\max\{\hat{s}_j, s^0\} - 2) \Big), \tag{28}$$

which corresponds to the expression stated in the Lemma. □

Proof of Propositions 3 and 4. We start by proving Proposition 4 since we will use it as a basis for proving Proposition 3.

Step 1: Proof of Proposition 4. Necessity. First note that, if the players permanently exploit projects with values $\hat{s}_1, \dots, \hat{s}_m$, then Equation (2) implies:

$$m \cdot c \leq \sum_{j=1}^{m} \frac{\delta}{1-\delta} (\hat{s}_j - 2). \tag{29}$$

However, Equation (29) implies Equation (4) in Lemma 1. Thus, the players replicate the project selection rule of the symmetric benefits benchmark, which implies $\hat{s}_j \geq$

$s^0 \, \forall \, j$.

Further, given that $\hat{s}_j \geq s^0 \, \forall \, j$, we can re-write Equation (4) as $\frac{\sum_{j=1}^{m} \hat{s}_j}{m} \geq \tilde{s}$.

Step 2: Proof of Proposition 4. Sufficiency: Condition 2 in Proposition 4 implies that Equation (4) holds, and thus that the project selection rule is that of the symmetric benefits benchmark. Further, Condition 1 in Proposition 4 implies that the project selection rule in the symmetric benefits benchmark specifies permanent exploitation.

Step 3: Proof of Proposition 3. Proposition 4 gives a necessary and sufficient condition for permanent exploitation when $m \geq 1$. Further, upon exploiting project p in period t, $\mu^t = \mu^{t+1}$, and thus the players also exploit project p in period $t+1$. Therefore, exploitation is always permanent. Hence, the players exploit a project p if and only if $s_p \geq \max\{s^0, \tilde{s}\}$. □

Proof of Proposition 5. Suppose $s_p \sim 3 - \frac{1}{\lambda} + \text{Exp}(\frac{1}{\lambda})$, which implies that $\mathbb{E}(s_p) = 3 > 2$. Moreover, the support of this distribution is convex for any λ. Hence, this distribution satisfies the assumptions made in the text.

Recall that the optimal relational contract conditions only on the best project found thus far for each dimension.

Statement 3 Note that there exists a sufficiently small value of δ such that the players are unable to exploit a project worth $s^0 + \epsilon$. Consider such a δ. With positive probability, in period 0 the players identify two projects with values belonging to an arbitrarily small range around $s^0 + \epsilon$ and $s^0 - \epsilon$. The players are unable to exploit either project in period 1 and, thus, must explore two new projects. Because the distribution of s_p is unbounded, for any δ, there exists a realization of s_p large enough such that $f(s^0 + \epsilon, s_p) > \tilde{s}$. Finally, in this region (i.e., $f(s^0 + \epsilon, s_p) > \tilde{s}$), the players follow the project selection rule of the symmetric benefits benchmark and thus permanently exploit both projects. Therefore, with positive probability, the players exploit a project they have previously chosen not to exploit.

Statement 1 Statement 2 implies Statement 1.

Statement 2 Suppose $\delta \geq \frac{1}{3}$ and $c = 1$, which ensures that $\mathcal{C}(\hat{s}_1, \hat{s}_2) \geq 1 = c \, \forall \hat{s}_1, \hat{s}_2$. When $\mathcal{C}(\hat{s}_1, \hat{s}_2) \geq 1$ but $f(\hat{s}_1, \hat{s}_2) < \tilde{s}$, the value of the second best project is irrelevant because within this range, the second best project will never be exploited since at most one project can be exploited. As argued in the text, one can write the Bellman equation for the players $B(\hat{s}_1, \hat{s}_2) = B(\max\{\hat{s}_1, \hat{s}_2\})$ when $f(\hat{s}_1, \hat{s}_2) < \tilde{s}$.

The indifference condition defining s^0 in the symmetric benefits benchmark is:

$$s^0 - 2 + \delta B^0(s^0) = 3 - 2 + \delta \mathbb{E}(B^0(\max\{s, s^0\})). \tag{30}$$

The left-hand side corresponds to the players' surplus when exploiting a project with value s^0. The right-hand side corresponds to the players' surplus when exploring one more project.

Suppose by contradiction that the players never exploit a project with value less than s^0. In other words, suppose that the players weakly prefer to explore two new projects when the best project found so far is worth s^0:

$$\begin{aligned} &s^0 - 2 + (3-2) + \delta \mathbb{E}(B(\max\{s, s^0\}) + \delta \epsilon_1(\lambda) \\ \leq &2(3-2) + \delta \mathbb{E}(B(\max\{s, s', s^0\}) + \delta \epsilon_2(\lambda). \end{aligned} \tag{31}$$

The first line corresponds to the value of exploiting a project worth s^0 and doing a singular exploration. Under such a project selection rule, the first two terms correspond to the players' expected surplus in the current period and the latter two terms correspond to the continuation value. The term ϵ_1 corresponds to the change in continuation value upon finding a project valuable enough that $f(\hat{s}_1, \hat{s}_2) \geq \tilde{s}$. Specifically, $\epsilon_1(\lambda)$ corresponds to the probability that the new project's value, s', is sufficiently large such that $f(\hat{s}_1, \hat{s}_2) \geq \tilde{s}$, multiplied by the difference in continuation value in this region, as opposed to the continuation value when the continuation value is less than 2. The second line corresponds to the players' surplus following two explorations, where $\epsilon_2(\lambda)$ is defined analogously. Both ϵ_1, ϵ_2 approach 0 uniformly as $\lambda \to \infty$. These convergences happen because, to reach $f(\hat{s}_1, \hat{s}_2) \geq \tilde{s}$, the players must draw a project with value equal to at least 4. Because (i) drawing such a project occurs with probability approaching 0 as $\lambda \to \infty$ and (ii) surpluses decrease and remain bounded as $\lambda \to \infty$, the ϵ_1 and ϵ_2 terms uniformly decrease. One can then subtract Equation (30) from Inequality (31) and simplify using the closed-form solution of B^0 to derive:

$$\frac{1}{\lambda(1-\delta)} \leq \mathbb{E}(B(\max\{s, s', s^0\} - B(\max\{s, s^0\})) + \epsilon_2(\lambda) - \epsilon_1(\lambda) \tag{32}$$

Next, we show that $B(x) - B(y) \leq \frac{x-y}{1-\delta} + \epsilon_3(\lambda)$ when $x > y$, where ϵ_3 is exponentially decreasing in λ. Except for the exponentially decreasing probability that the players' continuation value exceeds two (which is accounted for by $\epsilon_3(\lambda)$), the players will be able to exploit at most one project per period. As such, the largest possible benefit from exploiting a better project occurs from exploiting the better project in every period. Utilizing such a bound, we can derive the following inequality:

$$\frac{1}{\lambda(1-\delta)} \leq \mathbb{E}\left(\frac{\max\{s, s', s^0\} - \max\{s, s^0\}}{1-\delta}\right) + \epsilon_2(\lambda) - \epsilon_1(\lambda) + \epsilon_3(\lambda) \tag{33}$$

$$\iff \frac{1}{\lambda(1-\delta)} \leq \frac{1}{1-\delta}\left(\frac{2}{\lambda} - \frac{1}{2\lambda} - \frac{1}{\lambda}\right) + \epsilon_2(\lambda) - \epsilon_1(\lambda) + \epsilon_3(\lambda). \tag{34}$$

Finally, because $\epsilon_1(\lambda), \epsilon_2(\lambda), \epsilon_3(\lambda)$ are all exponentially decreasing in λ, we can ignore these terms in the limit. Thus, one can further simplify to derive:

$$1 \leq \left(2 - \frac{1}{2} - 1\right), \tag{35}$$

which is a contradiction. □

Proof of Proposition 8. Under the optimal relational contract, equilibrium plays depends solely on some function $f(\cdot)$ satisfying the properties listed in Proposition 6. Moreover, the following conditions are sufficient conditions for a relational contract with associated function $f(\cdot)$ to constitute an optimal relational contract:

$$\begin{aligned} c \cdot f(n) \leq\ & n\mathcal{C}(\text{exploitation}) + (f(n) - n)\mathcal{C}(\text{exploration}) \\ & + \sum_{j=f(n)+1}^{m} \mathbb{E}\Big(\delta^{\tau_j} | n, f(\cdot)\Big)\mathcal{C}(\text{exploration}), \quad \forall n. \end{aligned} \tag{36}$$

$$\begin{aligned} c \cdot \big(f(n) + 1\big) >\ & n\mathcal{C}(\text{exploitation}) + (f(n) - n + 1)\mathcal{C}(\text{exploration}) \\ & + \sum_{j=f(n)+2}^{m} \mathbb{E}\Big(\delta^{\tau_j} | n, f'(\cdot)\Big)\mathcal{C}(\text{exploration}). \quad \forall n. \end{aligned} \tag{37}$$

Inequality (36) states that the total reneging temptation in the current period is lower than or equal to the continuation value. In turn, the continuation value stems from three components which we enumerate in order: (i) the continuation value from the n projects that will be exploited in all future periods; (ii) the continuation value from the $f(n) - n$ projects that will be explored in the current period; and (iii) the continuation value from the projects that will be explored in future periods multiplied by the expected discount factor $\mathbb{E}\Big(\delta^{\tau_j}|n, f'(\cdot)\Big)$. Next, Inequality (37) states that the players cannot cooperate on an additional project in the current period, since, if they could, doing so would increase surplus and imply that $f(n)$ is not optimal. In Inequality (37), $f'(m) = f(m) \, \forall m \neq n$ and $f'(n) = f(n) + 1$. Given (36)-(37), we can bound $f(n) - f(n-1)$:

$$\begin{aligned} c \cdot \big(f(n) - f(n-1)\big) \geq \mathcal{C}(\text{exploitation}) - 1 + \mathcal{C}(\text{exploration}) \Bigg(& f(n) + 1 - f(n-1) \\ & - \sum_{j=f(n-1)}^{f(n)+1} \mathbb{E}\Big(\delta^{\tau_j}|n-1, f(\cdot)\Big) \\ & + \sum_{j=m}^{f(n)+2} \mathbb{E}\Big(\delta^{\tau_j}|n, f'(\cdot)\Big) - \mathbb{E}\Big(\delta^{\tau_j}|n-1, f(\cdot)\Big) \Bigg). \end{aligned} \tag{38}$$

Without solving for τ_j, we know $\tau_j \geq 0$, which implies that each term in the second line of the inequality is bounded by one.[27] Also, the number of projects suitable for exploitation is FOSD increasing in $f(n)$, because (i) there are more projects being explored and (ii) the number of projects suitable for exploitation is FOSD increasing in n, since $f(n)$ is itself increasing in n by Proposition 6. These observations imply that each term in the third line of the inequality are weakly positive. Utilizing these bounds implies:

$$f(n) - f(n-1) \geq \frac{1}{c} \left(\mathcal{C}(\text{exploitation}) - 1 - \mathcal{C}(\text{exploration})\right). \tag{39}$$

Next, note that (i) $\mathcal{C}(\text{exploitation}) - \mathcal{C}(\text{exploration}) \geq (1-q)v$ and (ii) $qv \geq 2$ by assumption, implying that:

$$f(n) - f(n-1) \geq \frac{1}{c} \left(\frac{2(1-q)}{q} - 1\right), \tag{40}$$

which concludes the proof. □

[27] To numerically solve for τ_j, one would need to solve a Markov Chain where each transition probability is dictated by a binomial random variable representing the number of projects suitable for exploitation discovered to date.

8 Appendix B

In this section, we provide proofs for the results discussed in Section 5.2.

Proposition 9. *In any optimal relational contract, two thresholds, denoted as s_s^* and s_a^*, exist. These thresholds determine the players' project selection rule: they explore a project if both (i) the highest-valued symmetric benefits project found so far has a value less than s_s^*, and (ii) the highest-valued asymmetric benefits project found so far has a value less than s_a^*. Furthermore, they permanently exploit the first symmetric or asymmetric benefits project with a valuation greater than s_s^* or s_a^*, respectively.*

Proof of Proposition 9. Note first that the optimal relational contract conditions only on the highest-valued symmetric and asymmetric benefits projects found to date. By Proposition 1, these are the only projects that may ever be exploited. Denote the values associated with the highest-valued symmetric and asymmetric benefits projects by $\hat{s}_s$ and $\hat{s}_a$, respectively. In any optimal relational contract, the project selection rule of the players can then be summarized as a function mapping $\hat{s}_s, \hat{s}_a$ into one of three choices: (1) exploiting the symmetric benefits project, (ii) exploiting the asymmetric benefits project, and (iii) exploration.

Next, note that after exploiting a project, the players' beliefs about the projects do not change, and, hence, if the players exploit a project once, they will permanently exploit that project. Therefore, the continuation value of the players' relationship associated with the permanent exploitation of a project with value s (if the exploitation of a project with value s is feasible) is equal to $\frac{\delta}{1-\delta}(s-2c)$. More specifically, for symmetric benefits project, exploitation is an equilibrium of the stage game, and thus the continuation value from exploiting a symmetric benefits project with value s is always equal to $\frac{\delta}{1-\delta}(s-2c)$. In contrast, the continuation value from exploiting an asymmetric benefits project with value s is $\frac{\delta}{1-\delta}(s-2c)\mathbf{1}_{c\leq\frac{\delta}{1-\delta}(s-2c)}$, where the condition in the indicator function corresponds to the condition under which the players are able to cooperate in exploiting the project.

Finally, the players never choose to exploit a project p they previously chose not to exploit. To see this, note that the players cannot exploit p in the future even if p is the highest-valued project (since, by assumption, they have chosen not to exploit it in the past). However, by Proposition 1, the players cannot exploit p when it is not the highest-valued project either. Hence, the continuation value from exploration is some constant, which we denote B.

Finally, suppose the highest-valued project found to date has value $\hat{s}$. If this project is a symmetric benefits project, the players exploit it if and only if $\frac{1}{1-\delta}(\hat{s}-2c) \geq B$. If this project is an asymmetric benefits project, the players exploit it if and only if $\frac{1}{1-\delta}(\hat{s}-2c) \geq B$ and $\frac{\delta}{1-\delta}(\hat{s}-2c) \geq c$. It follows from these expressions that the thresholds s_s^* and s_a^* stated in the proposition exist. □

Recall that $\tilde{s} = \frac{c(1+\delta)}{\delta}$. We now characterize the thresholds s_a^* and s_s^*.

Proposition 10. *The thresholds s_s^*, s_a^* exhibit the following properties:*

1. $s_a^* = \max\left(s^0, \tilde{s}\right) \geq s^0 \geq s_s^*$.
2. *s_s^* is monotone increasing in δ and monotone decreasing in q.*
3. *s_a^* is independent of q and U-shaped in δ.*

To prove Proposition 10, proving the following Lemma first is helpful.

Lemma 2. *Define by $\mathcal{C}(\delta, q)$ the continuation value of the players' relationship following exploration. Then, $\mathcal{C}(\delta, q)$ is decreasing in q and $\mathcal{C}(\delta, q)(1-\delta)$ is increasing in δ.*

Proof of Lemma 2. First, recall from the proof of Proposition 9 that the continuation value following exploration in the current period is independent of the values of the projects explored by the players up until and including the previous period. To prove that $\mathcal{C}(\delta, q)$ is decreasing in q, note that, as q decreases, the players are strictly more likely to encounter a symmetric benefits project. Because the players are always able to exploit symmetric benefits projects, the continuation value of their relationship weakly increases as q decreases.

Next, consider any two values $\delta_1 < \delta_2$. Note that any project selection rule implementable by an optimal relational contract when the players have discount factor δ_1 must also be implementable in equilibrium when the players have discount factor δ_2, because Equation (2) is relaxed as δ increases. Thus, given an optimal project selection rule for discount factor δ_1, $\mathbf{P}$, the players' expected continuation value is simply:

$$\mathcal{C}(\delta_1, q) = \delta_1 \pi(t+1, \mathbf{P}) + \delta_1^2 \pi(t+2, \mathbf{P}) + \ldots, \tag{41}$$

where $\pi(\cdot)$ denotes the expected joint surplus in a given period under the project selection rule. Further, because this project selection rule is also feasible with δ_2,

$$\mathcal{C}(\delta_2, q) \geq \delta_2 \pi(t+1, \mathbf{P}) + \delta_2^2 \pi(t+2, \mathbf{P}) + \dots \tag{42}$$

Combining these observations implies $\mathcal{C}(\delta, q)(1-\delta)$ is increasing in δ. □

Proof of Proposition 10. **Statement 1:** We first show that $s_a^* = \max\left(s^0, \tilde{s}\right)$. This result was shown in Proposition 3 for the case when $q = 1$. By Lemma 2, for any $q < 1$, the continuation value following exploration weakly increases compared to the case when $q = 1$. Because s_a^* is defined by the players' indifference between exploration and exploitation, the increased continuation value following exploration implies that $s_a^* \geq \max\left(s^0, \tilde{s}\right)$. Finally, s_a^* is not necessarily strictly greater than $\max\left(s^0, \tilde{s}\right)$, because (i) $s_a^* \geq s^0$ implies that, when benefits are symmetric, the players would exploit such a project and (ii) $s_a^* \geq \tilde{s}$ implies that the players are able to replicate the project selection rule of the symmetric benefits case.

What is left to show is $s^0 \geq s_s^*$. Note that the continuation value following exploitation is the same in this case and the case of symmetric benefits. However, the surplus following exploration is weakly higher under symmetric benefits. Thus, for any value s where exploitation is preferred in the symmetric benefits benchmark, exploitation is also preferred with asymmetric benefits. Thus, the threshold must be weakly higher compared to the symmetric benefits benchmark.

Statement 2: Note that the joint surplus associated with the exploitation of a symmetric benefits project with value s is equal to $\frac{s-2c}{1-\delta}$. Further, s_s^* represents the value a project must achieve for the players to be indifferent between exploiting the project and exploring. Thus:

$$\frac{s_s^* - 2c}{1-\delta} = \mathcal{C}(\delta, q) \iff s_s^* - 2c = (1-\delta)\mathcal{C}\left(\delta, q\right). \tag{43}$$

The statement now immediately follows given the results stated in Lemma 2.

Statement 3: This is precisely Corollary 1. □