Searching for Collaboration: The Dynamics of Relationship Building

Charles Angelucci and Roi Orzach*

September 18, 2023

Abstract

Building a successful relationship 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, revisit previously abandoned projects, and may 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, Alessandro Bonatti, David Byrne, Joyee Deb, Mert Demirer, Glenn Ellison, Dana Foarta, Guillaume Fréchette, Bob Gibbons, Marina Halac, Danielle Li, Jin Li, Chiara Margaria, Simone Meraglia, Juan Ortner, Hazhir Rahmandad, Daniel Rappaport, Kareen Rozen, Jean Tirole, Marta Troya Martinez, Michael Whinston, and participants in both the MIT Organizational Economics and Theory groups, 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. In organizational settings, managers rely on promises of bonuses to incentivize their subordinates, with the expectation that the subordinates will reciprocate by dedicating themselves to their work. Similarly, in inter-organizational dynamics, firms join strategic alliances to enhance R&D efforts, and the enduring success of these collaborations relies heavily on the continued willingness of firms to share resources and information. Likewise, in supply-chain relationships, buyers pledge to increase future transactions with suppliers who, in turn, pledge to consistently deliver the expected goods or services. Even at the political level, countries within a union initiate common projects, with the expectation that all member nations will contribute for the long term. In each of these contexts, promises play a key role in providing incentives, steering collaborations towards profitability. However, the efficacy of promises in fostering collaboration 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 made between collaborators is closely tied to the perceived value of the relationship. When a collaboration holds greater value for the involved parties, they are less likely to jeopardize it by breaking their promises. Consequently, a collaboration that offers significant present and future benefits presents a twofold advantage. Not only does it lead to higher gains, but it also fosters higher levels of credibility, allowing the parties to make more substantial promises to each other. This increased credibility, in turn, creates stronger incentives that lead to a more profitable collaboration overall. The interplay between the value of a collaboration and the credibility of the parties when making promises establishes a reinforcing relationship, providing a strong impetus for the parties to actively enhance the value of their collaboration as much and as early as possible.

Yet, in practice, building a successful collaboration is often a time-intensive and gradual process. Substantial time is invested in the initial stages to carefully define the collaboration's nature. Moreover, rather than expanding all at once, collaborations tend to evolve incrementally, widening their scope over time. As noted by consulting firm McKinsey in the context of 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).[1,2] Similarly, Dwyer et al. (1987) illustrate that buyer-supplier relationships transition through several stages — including from the "exploration" phase to the "expansion" phase, noting that "the critical distinction [between the exploration and expansion phases] 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." The lengthy and gradual nature of collaboration is also exemplified in the formation of political unions, such as the European Union, whose establishment was characterized by significant step-by-step policy implementation and a gradual deepening of integration among its member states (see, e.g., Haas, 1958; Urwin, 1995; Downs et al., 1998). 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 an unavoidably slow and deliberate pace 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 engage repeatedly over an infinite time horizon, with all actions being publicly observable. Each period provides players with an opportunity to cooperate on possibly multiple projects chosen from an infinite pool of ex ante identical projects. Any project can be perpetually selected for collaboration, a scenario we refer to as "project exploitation." The potential 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

[1]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." See Vanpoucke et al. (2014) for further examples of slow and gradual relationship building in buyer-supplier collaborations. Also, Macchiavello and Morjaria (2015) show that in the context of the Kenyan Rose export market, the duration of buyer-supplier relationships serves as a predictive factor for both the value and scope of the relationship.

[2]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.

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 compensate each other and provide incentives, but these transfers are voluntary in nature.

We focus on relational contracts that maximize the players' joint surplus from the collaboration. In our main setting, we assume that all projects yield asymmetric benefits across the two players. This approach emphasizes a key challenge in collaborative endeavors: the potential inclination for partners to withdraw their cooperation. 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, ensuring perfectly aligned incentives.

In the first part of our analysis, we assume that project exploration does not need to be motivated with monetary transfers. 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 that the players choose to exploit. 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 extend the analysis to allow players to explore or exploit multiple projects per period while maintaining the assumption that monetary incentives are not necessary to motivate project exploration. 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. Conversely, when the discount factor is low, the players are again likely to engage in much more 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 can differ across projects. This co-determination is due to the fact that selecting a project for exploitation affects the value of the players' overall relationship and, thus, their 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 summary, when players are involved in multiple collaborations simultaneously, inefficiencies arise not only in terms of the time spent finding suitable projects, but also in the players' deviation from the simple 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. 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 those 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 the collaboration on additional projects.

Lastly, we explore two extensions to our model. First, we generalize the framework to include projects that benefit all players, in addition to projects that benefit only a single player. We find that the players adopt lower (respectively, higher) thresholds for projects that exhibit consensus (respectively, no consensus). Second, we investigate

an environment where projects have correlated values and differ in their risk. We show that the players have incentives to explore the safer projects first, despite their lower upside potential.

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 symmetric benefits benchmark without incentive issues. Further, Section 4 solves the model, first by focusing on the players' exploration/exploitation choices and then by analyzing the dynamics of the scope of their relationship. Section 5 details the extensions including how groups find consensus and prevent relationship collapse. Finally, Section 6 offers our concluding remarks.

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 (1978). 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 their own arms and have incentives to free-ride on others' experimentation.[3] 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 with strategic interactions, which offers both 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). The positive feedback effect, which links the value of players' relationships to incentive strength, is pervasive across relational contracting models. However, it seldom generates dynamics, because the current period's production is typically assumed to be influenced only by the players' current actions and not by their 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

[3]For more recent work along these lines, see Keller et al. (2005), Heidhues et al. (2015), Ide (2023), and Hörner et al. (2022).

the relationship. The players initially choose to cooperate on low-risk, low-return projects, and they switch to cooperating on high-risk, high-return projects once their relationship has grown sufficiently valuable.[4] 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 collaborative projects. Our work also connects closely to Bernheim and Whinston (1990), who analyze firms operating in multiple markets, emphasizing that maintaining collusion in easier markets can help support collusion in more challenging ones. In a similar vein, Levin (2002) show the benefits firms can reap from pooling multiple employees' incentives into a "multilateral" relational contract, enabling cross-subsidization of incentives. In our setting, the players' ability to cooperate on multiple projects also enhances their cooperation. However, learning dynamics can drive agents to grow their relationship incrementally.

Our work is closely related to Chassang (2010), as both papers lie at the intersection of the multi-armed bandits and relational contracting literatures. Chassang (2010) considers an infinitely-repeated principal-agent setting in which an agent receives an exogenous fixed wage each period he is employed and decides whether to pull a costly arm that might benefit the principal or a costless arm guaranteed to fail. While the agent knows which arms are productive, the principal is initially unable to discern between the two types of arms. As a result, the main challenge she faces consists of motivating the agent to select the productive arms. During the "exploration" phase of their relationship, the principal must fire the agent following failures. This dynamic makes motivating exploration progressively expensive as more productive arms are identified and the relationship grows in value. 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, where growth in the relationship's value amplifies the players' ability to enhance their relationship.[5]

Other papers analyzing group search dynamics include Bonatti and Rantakari (2016), focusing on two heterogeneous agents seeking competing projects. In our

[4] 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.

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

model, project exploration is collective, but their exploitation demands cooperation. Our work is also related to research contrasting group and individual decision-making regarding experimentation (Reshidi et al., 2021; Chan et al., 2018).

Finally, we add to the existing body of research that examines the concept of gradualism in collaborations. In particular, Watson (1999, 2002) examine a setting marked by asymmetric information, where 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 defection and learn about each other's types. 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.[6]

2 The Setup

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. In each period t, each player i chooses a finite set, p_i^t, of no more than m projects from the set $\mathcal{P} = [0, 1]$. We assume that a project is selected if and only if both players choose it. We denote by $\mathbf{p}^t$ the corresponding set of projects selected by the players, where $\mathbf{p}^t = p_1^t \cap p_2^t$.[7]

Each project in $\mathbf{p}^t$ imposes a normalized cost of 1 on both players. Further, each project 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 each time the players select the project. 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

[6]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. See McAdams (2011) for a more recent contribution to this literature.

[7]We note that the number of projects the players are able to select, m, is exogenous. This could represent capacity limitations or the presence of profitable collaborative opportunities.

period t if it has been selected in some prior period. 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 according to the cumulative distribution function F.

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 $-i$.[8]

Player i's period t payoff is equal to:

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

We conclude the model's description by stating the timing of the stage game. Both players simultaneously choose their publicly observable discretionary transfers $w^t_{i,-i}$. Next, both players simultaneously make their project choices p^t_i. For each selected project $p \in \mathbf{p}^t$, the players observe and pocket the vector of benefits $(v_{p,1}, v_{p,2})$. Finally, both players simultaneously choose their publicly observable discretionary transfers $b^t_{i,-i}$.[9]

Relational Contracts. A relational contract is a complete plan for the relationship. Let $\mathbf{h}^t = (\mathbf{w}^0, \mathbf{p}^0, \mathbf{v}^0, \mathbf{b}^0, \ldots, \ldots, \mathbf{w}^{t-1}, \mathbf{p}^{t-1}, \mathbf{v}^{t-1}, \mathbf{b}^{t-1})$ 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 $\mathbf{h}^t \in \mathcal{H}^t$, a relational contract describes the bonuses and projects in each period: (i) the up-front transfers made by all players; (ii) the set of projects to be selected; and (iii) the discretionary payments to be paid as a function of the projects selected and the realizations of $\mathbf{v}^t$. Such a relational contract is self-enforcing if it describes a Subgame Perfect Equilibrium

[8] We incorporate the option of monetary transfers both before and after the players' project choices to simplify some of the proofs. Removing transfers at the beginning of each period would not impact the level of total surplus the players are able to achieve. However, allowing for such transfers allows for any split of the total surplus between the two players, with the sole restriction being that neither player's payoff falls below their outside option.

[9] Players lack private histories since all of their choices are publicly observable. This modeling approach allows us to streamline the analysis by focusing specifically on the effects of variations in the value of the players' relationship on the optimal relational contract.

of the repeated game. Within the class of Subgame Perfect Equilibria, we analyze pure-strategy equilibria which maximize the players' joint surplus.[10] In the event of a deviation, we assume that the players respond (i) by choosing $p_i^t = \emptyset$ if this choice has not been made and setting their remaining payments within the current period equal to zero 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).

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.

3.1 Characterization of Optimal Relational Contracts

In our setting, surplus-maximizing relational contracts will depend on the players' beliefs about the value of the projects, which we denote $\mu^t(h^t) := \{\Delta(s_p)|h^t)\}_{p\in[0,1]}$. 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 relational contract satisfying these features to be self-enforcing. 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:*

[10] Maximizing joint surplus is without loss of generality because the players have the ability to redistribute surplus while keeping the total surplus fixed due to the assumption of transferable utility. Further, restricting attention to pure strategy equilibria is without loss as players cannot gain from mixing on payments or project selections.

- *For any surplus-maximizing relational contract, there exists an alternative payoff-equivalent relational contract such that, on path, (i) for all t and for all histories $h^t \in \mathcal{H}^t$ the continuation equilibrium is surplus maximizing and (ii) for any two histories h_1^t and $h_2^{t'}$, if $\mu_1^t = \mu_2^{t'}$, then the relational contract specifies the same continuation equilibrium following these histories. We call such relational contracts optimal.*

- *Within this class of optimal relational contracts, a necessary and sufficient condition for a relational contract to be self-enforcing is that, for all t and for all histories $h^t \in \mathcal{H}^t$, the following inequality holds:*

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

  where $\mathcal{C}(\mu^t)$ is the discounted expected continuation value of the players' relationship under the relational contract and given beliefs μ^t.

The intuition for the first statement of the proposition is based on the following two observations. First, on the equilibrium path it is always without loss to play the surplus-maximizing continuation equilibrium and rely on money to provide incentives (rather than by playing non-optimal continuation equilibria). Any continuation equilibrium that is not surplus-maximizing can be replaced by a surplus-maximizing continuation equilibrium specifying appropriate monetary transfers. 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 value of the projects.

Inequality (2) states that for a relational contract to be self-enforcing at any history h^t, the discounted expected continuation value (hereafter, continuation value) of the relational contract across players and projects must exceed the total reneging temptation across players and projects. The total reneging temptation is the sum across players and across projects of a project's reneging temptation to a player. 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.

We conclude by noting 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 see this, observe that in the right-hand side of Equation (2), transfers cancel out (and the same is true for the players' total surplus). Similarly, the left-hand side solely depends on the players' net intrinsic gain from the selected projects.

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 $p \in \mathcal{P}$, $v_{p,1} = v_{p,2}$. We show that the standard results from multi-armed bandit problems hold in our setting (e.g., Weitzman, 1978). The predictions this benchmark analysis produces are identical to those that would arise if players behaved cooperatively or had access to formal contracts.

Proposition 2. *When projects generate symmetric benefits, all optimal relational contracts exhibit the following properties:*

1. *If the players select any project in any period t, then the number of projects selected in all periods is equal to the maximum number of projects the players can select, m.*

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

When projects yield equal benefits for both players, an equivalent condition to Inequality (2) from Proposition 1 states that actions at all dates t and histories h^t must induce a weakly positive net present value of the collaboration. Hence, we can characterize the players' behavior ignoring Inequality (2).

The intuition for statement (1) is that since all projects are ex-ante identical, if the players find it rational to select project $p \in [0, 1]$ in some period t (either to explore or exploit it), selecting as many projects as possible as early as possible must also be optimal. To gain intuition for statement (2), note that the players treat each of the m projects independently. 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 alternative projects with the hope of ultimately settling on a superior project in the future. As the discount factor increases, the value of exploring alternative projects rises, since any superior project identified can be used across all future periods.

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 forego private gains for the collective benefit.[11] Moreover, for any given project p, it is randomly determined which player among the two receives the entirety of the chosen project's associated benefits. Specifically, we assume that $v_{p,1} = s_p$ occurs with a probability of $\frac{1}{2}$ and that $v_{p,1} = 0$ occurs with a probability of $\frac{1}{2}$. Finally, the distribution of s_p is subject to the following restrictions: (i) $\infty > \mathbb{E}(s_p) > 2$, 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.[12]

[11] This assumption is relaxed in Section 5.1.

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

4.1.1 Collaborating on a Single Project

Suppose the players are limited to selecting up to one project per period (i.e., $m = 1$). 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 in case the discount factor is sufficiently high. Specifically, cooperation in project exploitation is achieved by sustaining a relational contract where the player benefiting from the project provides appropriate transfers to compensate the other player for their cost. 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 better project is easier for the players to cooperate on.

Similar to the symmetric-benefits benchmark, we can show that the optimal relational contract involves exploring projects until finding a project whose value s_p exceeds some threshold, denoted s^*. The following proposition formalizes this result.

Proposition 3. *In any optimal relational contract, a threshold value s^* exists 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 the value of the best project found thus far is the only factor in determining what optimal play is. Further, if the players find it optimal to stop exploring projects upon finding a project with value s, the same holds true for a project with value $s' > s$. In both cases, the expected payoff from exploring another project is equal, but the total expected payoff associated with exploitation is higher under the project with value s'. This implication explains why players adopt a threshold strategy, where the threshold remains constant over time because the value of exploration is time-invariant.

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. It follows that the players exploit project p in period $t+1$ and in all subsequent periods and, hence, the continuation value in period t is simply $\frac{\delta}{1-\delta}(s-2)$. 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* $1 < \frac{\delta}{1-\delta}(s_p - 2)$.

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

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

Given Proposition 3, characterizing the optimal relational contract reduces to characterizing s^*, which we now do.

Proposition 4. *In any optimal relational contract,* $s^*(\delta) = \max\{\tilde{s}(\delta), s^0(\delta)\}$. *Further, there exists a critical value* δ^* *such that* $\tilde{s}(\delta) > s^0(\delta)$ *if and only if* $\delta < \delta^*$.

In words, for a high discount factor (i.e., $\delta > \delta^*$), the presence of transferable utility ensures that the players are willing and able to exploit any project whose aggregate benefits are such that the project would be chosen for permanent exploitation under symmetric benefits. Consequently, exploration parallels the case of symmetric benefits. Instead, when the discount factor is low (i.e., $\delta < \delta^*$), the players are not able to exploit all the projects whose aggregate benefits would dictate permanent exploitation when those benefits are shared equally. As a result, they explore projects until they find one suitable for collaborative exploitation, with aggregate benefits such that they would also opt to exploit it under symmetric benefits.

It also follows from Proposition 4 that, when $\delta < \delta^*$, the gap between s^* and s^0 increases as δ decreases. When benefits are symmetric, as δ decreases, the players receive less benefit from exploration, and thus s^0 decreases. By contrast, when benefits are asymmetric, the players must find an even better project to be able to sustain cooperation, thus s^* increases.

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 δ varies for a fixed distribution $F \sim \exp(\frac{1}{3})$. 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. As one can see, s^0 is increasing in δ and approaches infinity as $\delta \to 1$ (for high δ the value of finding a marginally better project is high). Further, $\tilde{s}$ is decreasing in δ and approaches infinity as $\delta \to 0$ (for low δ the players must find a sufficiently good project to be able to sustain cooperation). Finally, as the discount factor decreases, the difference between s^* and s^0 increases.

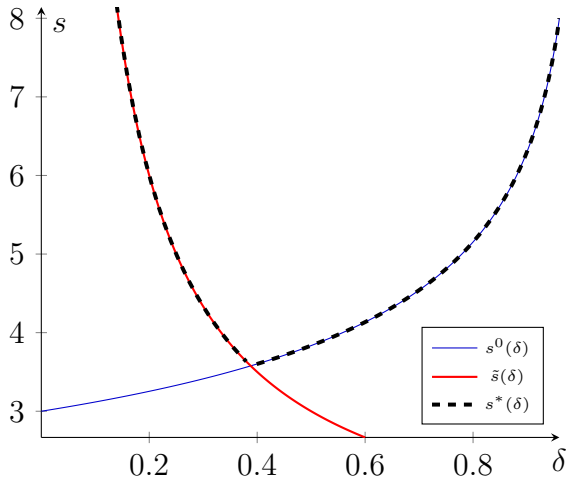


Figure 1: Comparative Statics of Group Search

The figure plots the $\tilde{s}$, s^*, and s^0 thresholds when $s_p \sim \exp(\frac{1}{3})$. The threshold $\tilde{s}$ is the minimum value of s for which cooperation in project exploitation is sustainable. The threshold s^* is the minimum value of s such that the players switch from exploration to exploitation. 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 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 $s^0(\delta)$.

4.1.2 Collaborating on Multiple Projects

We now extend the analysis from the previous subsection to consider the scenario in which the players can select up to $m \geq 1$ projects. 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 asym-

metric benefits the players continue to explore more than in the case of symmetric benefits. In addition, we show that the independence property across projects fails, leading the players to adopt behaviors such as the temporary exploitation of projects and the recalling of previously abandoned projects.

In the Appendix, we show that equilibrium play under the optimal relational contract in any period t relies solely on the values of the m most valuable projects discovered thus far, denoted as $s_1, \ldots, s_m$. Recall the definition of $\tilde{s}$ as established in Equation (3). An implication of Proposition 4 is that the players can (and thus do) mimic the project selection rule of the symmetric benefits benchmark upon finding a project whose value s is such that $\max(s, s^0) \geq \tilde{s}$ (i.e., they permanently exploit it). The next Lemma provides the analogous condition when the players can cooperate on m projects per period.

Lemma 1. *The players mimic the project selection rule of the symmetric benefits benchmark if and only if the set of $s_1, \ldots, s_m$ values satisfies:*

$$f(s_1, \ldots, s_m) := \frac{1}{m} \sum_{j=1}^{m} \max\{s_j, s^0\} \geq \tilde{s}. \tag{4}$$

The function $f(s_1, \ldots, s_m)$ corresponds to the minimal possible average value of the projects chosen for permanent exploitation when following the rule adopted by the players in the case of symmetric benefits. When the average of these values exceeds $\tilde{s}$, the players can effectively pool their incentives across the m projects, allowing the players to exploit these projects.

The condition provided in Lemma 1 implies a necessary condition for the players to permanently exploit the m highest values projects found thus far, a scenario we refer 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.

Proposition 5. *In any optimal relational contract, the players permanently exploit projects with values $s_1, \ldots, s_m$ if and only if:*

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

2. The average of $s_1, \ldots, s_m$ exceeds $\tilde{s}$.

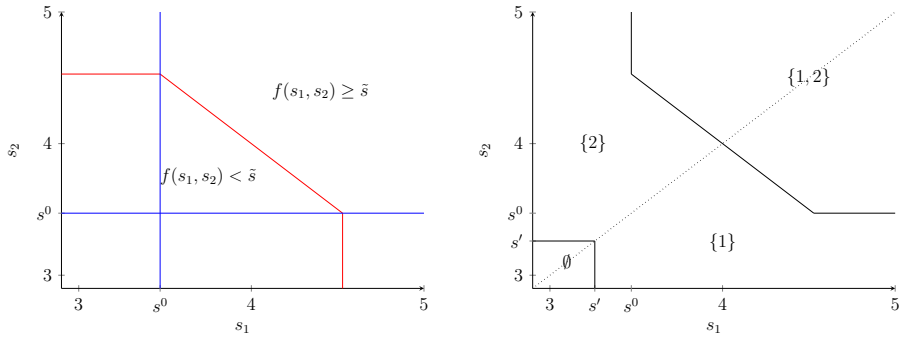
Note that exploration ceases in finite time since, under Assumption 1, there exist project values exceeding $\tilde{s}$. Furthermore, the two conditions stated in the proposition mirror those in Proposition 4. However, instead of stating that each s_j exceeds $\tilde{s}$, the Proposition states that the average value must exceed $\tilde{s}$ because relational incentives are pooled across projects.

We have provided closed-form conditions under which the players can 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 consider $m = 2$ and focus on a specific parametric example.

Exponential Distribution Example When $m = 2$. Suppose that $\delta = \frac{1}{3}$ and also that $s_p \sim \exp(\frac{1}{3})$, such that $\mathbb{E}(s_p) = 3$. Further, denote by $\mathcal{C}(s_1, s_2)$ the expected continuation value of the players when the two best projects have values s_1 and s_2. In this specific case, we first note that $1 \leq \mathcal{C}(s_1, s_2)$ for all combinations of s_1 and 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 an expected continuation value of $\frac{\delta}{1-\delta}2$, which reduces to 1 when $\delta = \frac{1}{3}$. Within this example, (i) we utilize Proposition 5 to characterize the strategy for the players when $f(s_1, s_2) \geq \tilde{s}$ and (ii) derive results for the optimal strategy when $f(s_1, s_2) < \tilde{s}$.

Figure 2a displays s^0 in blue and the set of s_1 and s_2 values where $f(s_1, s_2) \geq \tilde{s}$ in red. This set is the solution to Equation (4) in Lemma 1. In Figure 2b, we indicate which projects are exploited based on the values of the two best projects found thus far, s_1 and s_2. First, using Figure 2a, we can determine the conditions under which both projects are exploited, specifically when both $\tilde{s} \leq f(s_1, s_2)$ and $s_1, s_2 \geq s_0$. This region is denoted by $\{1, 2\}$, indicating that both projects are exploited. 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 s_1 (corresponding to the region denoted by $\{1\}$) and s_2 (corresponding to the region denoted by $\{2\}$) 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 project such

that, below this threshold, the players choose to explore two new projects rather than exploiting the best of the two projects (corresponding to the region denoted by $\emptyset$). Conversely, above this threshold, the players choose to exploit the better of the two projects and explore a new one. Moreover, this threshold remains independent of the second-best project's value. Given that players cannot currently exploit this project, it would only come into contention for exploitation if a project even better than the current top one was found. But in such a scenario, the once second-best project would drop to a third-place ranking, rendering it irrelevant. Figure 2b shows this cutoff as derived from a simulation.[13]



(a) Implementability (b) Optimality

Figure 2: In the figure, we assume $m = 2$, $\delta = \frac{1}{3}$, and $s_p \sim \exp(\frac{1}{3})$. The values s_1 and s_2 represent the values of the best two projects discovered by the players to date. 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 s_1 and s_2 values such that $f(s_1, s_2) = \tilde{s}$ in red. The right figure plots the project selection behavior of the players under the optimal relational contract. The $\{1, 2\}$ region corresponds to all the s_1 and s_2 values such that the players choose to exploit both projects. The $\{1\}$ region (resp. $\{2\}$ region) corresponds to all the s_1 and s_2 values such that the players choose to exploit the project generating value s_1 (resp. s_2) and to explore another project. Lastly, the $\{\emptyset\}$ region corresponds to all the s_1 and s_2 values such that the players choose to explore two projects.

Figure 2b implies novel dynamics along the equilibrium path. For example, in some cases, the threshold determining whether a project is exploited in a given pe-

[13]We employ simulation methods because the continuation value is non-linear. To illustrate a source of non-linearity, suppose that $s_1 > s_2$ and note that an increase in s_1 raises exploitation benefits in the current period and that, with a non-zero probability equal to $1 - F(s_1)$, the project achieving s_1 will be replaced in the next period. The simulation randomly draws values of s_p and initializes a value function $B(v_1, v_2)$. Next, over a grid of v_1, v_2, given the assumed value function, we compute the optimal exploration vs exploitation choices for the two projects and generate an updated value function. The simulation continues iterating until the supremum norm between the old and new value functions is below a threshold. All other features of the figures have analytical expressions.

riod can be lower than s^0. However, we know from Proposition 5 that the projects ultimately chosen for permanent exploitation must have value exceeding s^0, implying that some projects might be exploited only temporarily. Building on the previously discussed example, in what follows we continue to assume that $s_p \sim 3 - \frac{1}{\lambda} + \exp(\lambda)$ without restricting δ to the value $\frac{1}{3}$. In this environment, we provide an analytical proof that, beyond the deviations already highlighted in Proposition 5, additional discrepancies from the symmetric benefits benchmark can arise in equilibrium when players decide between project exploitation and exploration.

Proposition 6. *Suppose $s_p \sim 3 - \frac{1}{\lambda} + \exp(\lambda)$ 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 replicate the actions chosen in the case of symmetric benefits. Furthermore, even within the parameter range where these behaviors occur with a positive probability, their occurrence is not guaranteed: 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 immediately from the second statement combined with the result from Proposition 5, whereby the players only permanently exploit projects whose values exceed s^0.[14] 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 expected continuation value of the players' relationship exceeds 2 with an arbitrarily small probability. In this case, with a probability approaching 1, the

[14]The temporary exploitation of project can arise under other scenarios. For instance, in Figure 2b, if the players identify projects with values $3, 3.5$ in period 0, then in period 1 they will exploit the project with value 3.5 and explore another project. If this exploration leads to finding a project with value 4, then this new project will be exploited rather than the project with value 3.5.

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. For instance, 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, could prompt the players when benefits are symmetric to set a higher exploitation threshold.

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 expected 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

In this subsection, we delve into the analysis of how the two players gradually broaden their relationship scope over time. When moving from Section 4.1.1 to 4.1.2, we already expanded the scope of players' relationship from $m = 1$ to $m \geq 2$. However, we concurrently 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. In turn, we specialize the model presented in Section 2 as follows. First, we assume that $s_p \in \{0, v\}$, where $s_p = v > \frac{2}{q}$ with probability q, for all 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

[15]Specifically, we assumed that $\mathbb{E}(v_{p,i}) > 1$, 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.

by player 1. This approach can be illustrative of, say, an employment setting, where player 1 plays the role of the employer and player 2 assumes the role of the employee. We show that gradualism – that is, delaying the exploration of a project until previous projects with value v have been found – can be sustained in equilibrium for a larger range of discount factors, even though it generates less joint surplus compared to when exploring all projects 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. To do so, we denote by $f(n)$ the number of projects the players cooperate on, where n denotes the number of projects suitable for exploitation found thus far. 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 7. *To any (non-empty) optimal relational contract corresponds a function $f(n)$ specifying that, conditional on having found n projects suitable for exploitation, 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 monotone increasing in n for all n.*

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

The intuition behind the first statement is that when the players identify new high-value projects, there is a strict increase in their expected continuation value. This increase in the expected continuation value subsequently enables the players to select a weakly greater number of projects.

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 $n-1$ but fails for n. This implies $f(n-1) > n-1$ and $f(n) \leq n$, respectively. Combining these inequalities and the first statement of the proposition

[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 it easier to sustain cooperation.

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. These conditions amount to:

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

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

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

One implication of Proposition 7 is that the players increase the scope of their relationship and eventually maximize the scope of their relationship. It is worth noting that Proposition 2 implies that when benefits are symmetric, $f(n) = m$ for all n. However, this is not true when benefits are asymmetric. In fact, we can show that for any $m \geq 2$, there exists a value of δ low enough such that $f(0) < m$, implying that the relationship "starts small." This is stated below.

Proposition 8. *For any $m \geq 2$, there exist two thresholds $\underline{\delta} < \bar{\delta}$, such that when $\delta \in (\underline{\delta}, \bar{\delta})$, 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 strategy constitutes an equilibrium if and only if:

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

As δ diminishes, $\mathcal{C}$(exploration) decreases to a point where Inequality (7) fails to hold. 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

[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.

strategy is determined by:

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

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

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 is feasible, (i.e., exceeds $\underline{\delta}$), but low enough such that maximal scope is not (i.e., less than $\bar{\delta}$), the players' scope will gradually increase over time.

We note that the specific strategy used for the intuition above is one of many potential "gradual" options. For instance, there exists a "very gradual" strategy where $f(n) = \min(m, n+1)$ in which the players explore at most one new project per-period. Such a strategy is not always optimal. Proposition 1 implies that the players maximize their joint 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.[18] Namely, if q is low, the increase in expected 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.

Proposition 9. *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 rare (i.e., q is small), upon finding a project worth exploiting, the players can increase their relationship scope by

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

at least α 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 additional projects.

5 Extensions

In this section, we explore various extensions of the main model. Firstly, we generalize the model to allow for projects with both symmetric and asymmetric benefits across the players, and we examine how their inclusion impacts the players' exploration behavior. Secondly, we relax the assumptions of infinite projects and independent project values, allowing us to investigate situations where relationships may come to an end. We discover that to prevent relationship collapse, the players tend to explore projects with lower variance.

5.1 The Role of Consensus Projects

In Section 4.1.1, the assumptions we made regarding the distributions of individual valuations for the projects resulted in a situation where one player had to compensate the other player to ensure project exploitation. We now generalize the setting to include projects that can be exploited without the need for monetary transfers between the players. As before, we assume that $s_p \sim F$, but in this generalized setting, we consider the following valuation distribution: 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 $v_{p,1} = 0$. Setting $q = 1$ corresponds to the model considered in Section 4.1.1. By contrast, setting $q = 0$ corresponds to the consensus benchmark, as no conflict exists among the players. We define a project as exhibiting "consensus" when both players benefit from exploiting it. In the case where $q \in (0, 1)$, it can be shown that there exist two distinct exploration/exploitation thresholds, s_c^* and s_{nc}^*, where the first threshold applies to projects that exhibit consensus, while the second one applies to the remaining projects. The key finding that

emerges from this analysis is that $s^*_{nc} \geq s^*_c$, indicating that the players are less selective when it comes to projects that exhibit consensus, as compared to those that do not. Additionally, we show that: $s^*_{nc} \geq s^0 \geq s^*_c$. The intuition behind $s^*_{nc} \geq s^0$ is the same intuition as the text: a project that benefits a single player must be valuable enough to enable cooperation in exploitation. The intuition behind $s^0 \geq s^*_c$ is as follows: When benefits are symmetric, the players explore a new project and discover a project with value $s_p \in (s^0, s^*_{nc})$, they can exploit it individually. However, cooperation between the players to exploit this project is not possible with asymmetric benefits. As a result, the value of exploration for the players is lower, leading to lower exploration/exploitation thresholds for projects that exhibit consensus. This result bears resemblance to Acharya and Ortner (2022), where two players involved in collective search prioritize projects that exhibit consensus.

5.2 Relationship Collapse

We investigate the impact of project correlation by considering a scenario with only two projects: 1 and 2. Additionally, we introduce the following correlation structure:

$$v_{1,1} \sim F, v_{2,1} \sim v_{1,1} + \epsilon, \text{ and } v_{p,2} = 0 \quad \forall p. \tag{10}$$

We further impose the following three conditions: (1) Inf(Supp(F)) > 2; (2) $\mathbb{E}(\epsilon) = 0$; (3) $v_1 \perp \epsilon$. These conditions imply that project 1 is guaranteed to be profitable, that projects 1 and 2 have the same expected profitability, and that there is a positive correlation between project valuations. With these conditions in place, we proceed to examine the behavior of the players with symmetric and asymmetric benefits.

Lemma 2. *When benefits are symmetric, in period 0, the players select project* 2*. If* $v_{p,2} > v^*$*, then project 2 is selected in all subsequent periods. If* $v_{p,2} \leq v^*$*, the players select project 1 in period 1, and the better of the two projects in all subsequent periods.*

The behavior described in Lemma 2 is motivated by the fact that with symmetric benefits the players are risk-loving in this particular environment. Additional risk entails the (positive) potential of a project being exploited in all subsequent periods, while the (negative) possibility of a disappointing realization leads to the players not utilizing the project again.

The key tension with asymmetric benefits arises when a disappointing realization of project 2 makes players more pessimistic about project 1. This can make it difficult to motivate the exploration of project 1, and the realization may even be low enough to prevent the players from exploiting project 2. Consequently, in period 0, there must be enough continuation value to incentivize effort, considering the possibility of collapse. Likewise, there is a similar force when exploring project 1 first: after sufficiently disappointing realizations of v_1, the relationship may collapse. We show that there exist distributions of project valuations where sampling project 1 first leads to a lower probability of collapse and is dynamically enforceable for a broader range of parameter values. This reasoning is formalized in Proposition **??** in the Appendix.

This finding provides an additional reason why incremental innovation (exploring a project with lower variance) is easier to incentivize than radical innovation (exploring a project with higher variance). This stands in contrast to traditional reasoning that firms struggle to write formal contracts for radical innovation (e.g., Aghion and Tirole, 1994).

6 Concluding Remarks

While much of our focus in applications and examples has been on interactions between firms, our model also 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). In his model, players are aware of the existence of more efficient practices. However, the process of pinpointing them introduces considerable risk to the ongoing stability of their relationship. As a result, the players might prefer to simply exploit the practices they have identified thus far. 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 an intricate and time-consuming endeavor and might even prompt players to refine certain facets of their partnership before venturing into other areas of improvement. Thus, elements of luck and path dependence can lead to large performance differences, even across collaborations that started under similar circumstances.

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. This complementarity occurs because these individuals might find pooling relational incentives across practices advantageous. However, unlike managerial practices with technological interdependencies, we show that practices without such interdependencies are more easily implemented in a gradual manner. In summary, managerial practices, whether characterized by technological interdependencies or not, can lead to lasting performance gaps among firms. The difference arises from their inherent challenges: practices with technological interdependencies may necessitate bundled adoption, whereas those without such ties but anchored in relational contracts might demand extended periods for effective

implementation.

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.

Aghion, Philippe and Jean Tirole, "The management of innovation," *The Quarterly Journal of Economics*, 1994, *109* (4), 1185–1209.

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.

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.

Bonatti, Alessandro and Heikki Rantakari, "The politics of compromise," *American Economic Review*, 2016, *106* (2), 229–59.

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.

Chan, Jimmy, Alessandro Lizzeri, Wing Suen, and Leeat Yariv, "Deliberating collective decisions," *The Review of Economic Studies*, 2018, *85* (2), 929–963.

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.

Downs, George W., David M. Rocke, and Peter N. Barsoom, "Managing the Evolution of Multilateralism," *International Organization*, 1998, *52* (2), 397–419.

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

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.

Haas, Ernst B., *The Uniting of Europe: Political, Social, and Economic Forces, 1950-1957*, Stanford, CA: Stanford University Press, 1958.

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.

Heidhues, Paul, Sven Rady, and Philipp Strack, "Strategic experimentation with private payoffs," *Journal of Economic Theory*, 2015, *159*, 531–551.

Hörner, Johannes, Nicolas Klein, and Sven Rady, "Overcoming Free-Riding in Bandit Games," *The Review of Economic Studies*, 11 2022, *89* (4), 1948 – 1992.

Ide, Enrique, "Dual Moral Hazard and The Tyranny of Success," Working Paper 2023. Unpublished.

Keller, Godfrey, Sven Rady, and Martin Cripps, "Strategic Experimentation with Exponential Bandits," *Econometrica*, 2005, *73* (1), 39–68.

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.

McAdams, David, "Performance and turnover in a stochastic partnership," *American Economic Journal: Microeconomics*, 2011, *3* (4), 107–42.

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

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

Reshidi, Pëllumb, Alessandro Lizzeri, Leeat Yariv, Jimmy H Chan, and Wing Suen, "Individual and collective information acquisition: An experimental study," Technical Report, National Bureau of Economic Research 2021.

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.

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.

Urwin, Derek W., *The Community of Europe: A History of European Integration Since 1945*, London, UK: Longman, 1995.

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*, Vol. 78, Department of Energy, 1978.

7 Appendix A

Proof or Proposition 1. We first note that there always exists an equilibrium of the stage game in which all players choose $p_i^t = \emptyset$. Indeed, choosing $p_i^t = \emptyset$ is a best-response for player i when anticipating that player j chooses $p_j^t = \emptyset$. In this equilibrium, every player obtains a payoff equal to 0. Because the players' outside option is also equal to 0, this equilibrium yields each player their lowest possible equilibrium payoff. For this reason, in all optimal relational contracts we consider, the players revert to this static equilibrium in all periods that follow a deviation. During the deviation period, if the players have not already chosen p_i^t, they opt for $\emptyset$. They also set their remaining payments equal to zero.

The proof proceeds in four steps: (i) surplus-maximizing relational contracts without loss of optimality are surplus-maximizing at each history on path, (ii) the second statement of the proposition stating a necessary and sufficient condition for selecting a set of projects, (iii) showing such a condition is independent of how the continuation value is divided across the players, and (iv) showing that if two histories generate the same beliefs, then relational contract can prescribe the same continuation equilibria at these histories.

Step 1 We show that any surplus-maximizing relational contract can be replaced by a relational contract that is surplus-maximizing at every on-path history h^t. To see this, suppose this assertion does not hold. Namely, there exists an on-path history h^t such that the continuation equilibrium e^1 has lower total surplus than continuation equilibrium e^2. Thus, if we define c_i^j to be the continuation payoff to player i in equilibrium j, then $\sum_i c_i^1 < \sum_i c_i^2$. For the rest of Step 1 in the proof, we will 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 transfers in period $t-1$ for that specific history (and thus for that specific realization of $\mathbf{v}^{t-1}$) are adjusted so that (i) player 2's expected payoff is the same as in the equilibrium dictated by e^1 and (ii) player 1 pockets the entirety of the extra surplus. Specifically, take the vector of transfers $\mathbf{b}_1 = (b_{1,2}^1, b_{2,1}^1)$ associated with e^1 and the vector of transfers $\mathbf{b}_2 = (b_{1,2}^2, b_{2,1}^2)$ associated with e^2, and create a

third vector of transfers $\mathbf{b}_3 = (b^3_{1,2}, b^3_{2,1})$ such that:

$$c^2_1 + b^3_{2,1} - b^3_{1,2} > c^1_1 + b^1_{2,1} - b^1_{1,2},$$
$$c^2_2 + b^3_{1,2} - b^3_{2,1} = c^1_2 + b^1_{1,2} - b^1_{2,1}.$$

Since the continuation equilibrium e^2 has strictly higher surplus, finding payments that satisfy $b^3_{1,2} \leq c^2_1$ and $b^3_{2,1} \leq c^2_2$ is always feasible, given that $b^1_{1,2} \leq c^1_1$ and $b^1_{2,1} \leq c^1_2$. Note that such a change has no impact on player 1's incentives for choices in any period $t' \leq t-1$. This is because all actions are observable, and hence choosing a different action from the proposed equilibrium would be labeled a defection. If continuation equilibrium e^1 (with its strictly smaller continuation payoff for player 1) was enough to deter these defections, then so too is this equilibrium. The same logic applies for player 2 as they obtain the same continuation payoff, and thus have the exact same expected continuation payoffs in all periods $t' \leq t-1$. Finally, note that changing the continuation equilibrium to e^2 (weakly) increases the surplus from a date 0 perspective.

Step 2 Claim: There exists a relational contract prescribing that the players select a set of projects $\mathbf{p}^t$ at history h^t if and only if:

$$\sum_{p \in \mathbf{p}^t} \sum_{i=1,2} \max\Big(0, 1 - \mathbb{E}(v_{i,p}|h^t)\Big) \leq \mathcal{C}(h^t), \tag{11}$$

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

To show that (11) is a necessary and sufficient condition, consider a set of transfers $b_{i,-i}(\mathbf{v}^t) \geq 0$ from player i to $-i$ to be paid on path conditional on the vector of realized values $\mathbf{v}^t$.

Given an equilibrium project selection $\mathbf{p^t}$, it is without loss of generality to assume $p^t_1 = p^t_2 = \mathbf{p^t}$. Thus, for each player and for each $p \in \mathbf{p^t}$, the player must weakly prefer to include p in p^t_i, rather than excluding it. Let $\alpha_i(\mathbf{v}^t)$ denote player i's share of $\mathcal{C}(h^t)$ as a function of $\mathbf{v}^t$. Hence, the condition for selecting $\mathbf{p^t}$ is:

$$\sum_{p \in \mathbf{p^t}} \max\big(1 - \mathbb{E}(v_{i,p}|h^t), 0\big) \leq \mathbb{E}\Big(b_{-i,i}(\mathbf{v}^t) - b_{i,-i}(\mathbf{v}^t) + \alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t)\Big) \quad \forall i \tag{12}$$

$$b_{i,-i}(\mathbf{v}^t) \leq \alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t) \quad \forall \mathbf{v}^t, i \tag{13}$$

Expectations are taken over the project realizations $\mathbf{v}^t$ and $h^t \sqcup \mathbf{v}^t$ denotes the players' updated beliefs after observing $\mathbf{v}^t$.[19] The first expression states that the promised bonuses 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 (12) 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 (11).

To show sufficiency: We will show this result in multiple steps.

SubStep 1: We show we can replace Equation (13) by its expectation. This new expression is as follows:

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

We first show that if there is a solution to equations (14) and (12) then there exists a solution to equations (13) and(12).

Take a set of bonuses $b_{i,-i}(\mathbf{v}^t)$ that satisfy equations (14) and (12). Define:

$$b'_{i,-i}(\mathbf{v}^t) = \alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t) - \left(\mathbb{E}\Big(\alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t) - b_{i,-i}(\mathbf{v}^t)\Big)\right). \tag{15}$$

Since Equation (14) holds, the term in the expectation of equation (15) is positive and thus Equation (13) holds for all realizations of $\mathbf{v}^t$ under the set of bonuses $b'_{i,-i}(\mathbf{v}^t)$. Finally, $\mathbb{E}(b'_{i,-i}(\mathbf{v}^t)) = \mathbb{E}(b_{i,-i}(\mathbf{v}^t))$ so equation (14) continues to hold.

SubStep 2: Using substep 1, it suffices to show that Equation (11) implies a solution to Equations (12) and (14). To simplify all the notation with expectations, Equation (12) can be re-expressed as:

$$\beta_i - \gamma_i \leq (\tilde{b}_{-i,i} - \tilde{b}_{i,-i}), \tag{16}$$

where $\tilde{b}_{i,-i}$ is the expected payment from i to $-i$, $\beta_i = \sum_{p \in \mathbf{p^t}} \max\big(0, 1 - \mathbb{E}(v_{i,p}|h^t)\big)$,

[19] 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.

and $\gamma_i = \mathbb{E}(\alpha_i(\mathbf{v}^t)\mathcal{C}(h^t \sqcup \mathbf{v}^t))$. Equation (14) can thus be re-written as:

$$\tilde{b}_{i,-i} \leq \gamma_i. \tag{17}$$

Rearranging Equation (11) 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 these equations. Equation (16) holds as

$$\begin{aligned}
\beta_i - \gamma_i &\leq \max(0, \beta_i - \gamma_i) - \max(0, \beta_{-i} - \gamma_{-i}) && (18)\\
\iff \max(0, \beta_{-i} - \gamma_{-i}) &- \min(0, \gamma_i - \beta_i) \leq 0 && (19)\\
&\impliedby \sum_i (\beta_i - \gamma_i) \leq 0. && (20)
\end{aligned}$$

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

Further, Equation (17) reduces to

$$\begin{aligned}
\max(0, \beta_{-i} - \gamma_{-i}) \leq \gamma_i \impliedby \beta_{-i} - \gamma_{-i} &\leq \gamma_i && (21)\\
&\impliedby \sum_i (\beta_i - \gamma_i) \leq 0, && (22)
\end{aligned}$$

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

Step 3: In this step we show that we can make the expected utility for player 2 in all periods t equal to 0.[20] First, note that the expression in Equation (2) does not rely on the distribution of the payoff across players, but rather the total continuation value. Hence, for any period t in which player 2's total surplus is positive we can move this surplus to player 1 by increasing $w_{2,1}$ to be exactly this value. Further, player 2 is willing to make this transfer because, if they did not make the transfer, this would be considered a deviation and would yield player 2 a payoff of 0.

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 generality. 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 but with the same beliefs μ. Recall from Step 3, that it is without loss to

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

consider relational contracts in which player 2 obtains a payoff of 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 equilibria to the other does not change the players' incentives as both prescribe the exact same payoff to all players.

Finally, note that the only decision-relevant information for the continuation equilibrium is the beliefs. Hence, we 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.

Statement 1: Note that all projects within the interval $[0, 1]$ are ex ante identical. Hence, if the players believe that any of these projects exhibits a positive net present value before its exploration, leading to its selection and exploration, then all projects belonging to the interval $[0, 1]$ will have a positive net present value. This implies the players either choose 0 projects in all periods or m projects in all periods.

Statement 2: We first note that the players treat each of the m projects they can choose from the interval $[0, 1]$ as independent. This independence property arises because all of the payoffs are additively separable across projects. Given the stationary nature of the problem, the optimal strategy only conditions on the highest value of a previously explored project, which we denote $\hat{s}$. In particular, one can write the Bellman Equation:

$$B_0(\hat{s}) = \max_{\text{explore, exploit } \hat{s}} \{\mathbb{E}(s') - 2 + \delta\mathbb{E}(B_0(\max(\hat{s}, s'))), \hat{s} - 2 + \delta B_0(\hat{s})\}. \tag{23}$$

The first term in the maximum operator corresponds to the players' expected payoff when exploring one more project (chosen at random, since all unexplored projects are ex ante identical) and the second term is their payoff when exploiting the best project found thus far. Next, one can show there exists a cutoff 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.[21]

[21]This is not to say s^0 is unique. For instance, if s had a Bernoulli distribution, any $s^0 \in (0, 1]$ would generate the same equilibrium strategy and payoffs.

□

Proof of Propositions 3 and 4. Note that Proposition 5 gives a necessary and sufficient condition for permanent exploitation. Further, note that upon exploiting project p in period t, $\mu^t = \mu^{t+1}$, and thus the players also exploit project p in period $t+1$. Thus, all exploitation is permanent exploitation. Hence, the players exploit a project p if and only if $s_p \geq \max\{s^0, \tilde{s}\}$. □

Proof of Figure 1. When $v \sim \exp(\lambda)$ and benefits are symmetric, we can compute the players' indifference threshold s^0:

$$\frac{s^0}{1-\delta} = \mathbb{E}(v) + \frac{\delta}{1-\delta}\mathbb{E}\left(\max\{v, s^0\}\right) \tag{24}$$

$$\iff \frac{s^0}{1-\delta} = \frac{1}{\lambda} + \frac{\delta}{1-\delta}\left(e^{-\lambda s^0}\left(s^0 + \frac{1}{\lambda}\right) + \left(1 - e^{-\lambda s^0}\right)s^0\right). \tag{25}$$

The left-hand side corresponds to the exploitation payoff. The right-hand side corresponds to the expected payoff when exploring one more time and subsequently exploiting the best found 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 $s^0 < v$ or $v < s^0$, 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. We first note that, given two projects p and p' with values $s_p > s_{p'}$, $p' \in \mathbf{p} \implies p \in \mathbf{p}$, where recall that $\mathbf{p}$ denotes the set of all selected projects. Thus, continuation play depends only on the values of the best m projects found thus far. Denote these values by $s_1, \ldots, s_m$ and the corresponding continuation value in the case of symmetric benefits by $\mathcal{C}^0(s_1, \ldots, s_m)$.

Sufficiency: It suffices to show that, if Equation (4) holds in the current period under the project selection rule of the symmetric benefits, then Equation (2) holds forever after. To see why, note the values of $s_1, \ldots, s_m$ weakly increase with time, implying that $f(s_1, \ldots, s_m)$ can only increase with time, and thus if Equation (2) ever failed to hold, there would exist a set of projects with values $s_1, \ldots, s_m$ where Equation $f(s_1, \ldots, s_m) < \tilde{s}$ but Equation (2) did not hold in the current period.

Noting that the left-hand side of Equation (2) is equal to the number of exploited

projects, which cannot exceed m, we can establish the following sufficient condition:

$$m \leq \mathcal{C}^0(s_1, \ldots, s_m). \tag{26}$$

Note that $\mathcal{C}^0(s_1, \ldots, s_m) = \sum_{i=1}^m \mathcal{C}^0(s_i)$, since in the case of symmetric benefits the continuation value is additively separable across projects. Further, $\mathcal{C}^0(s_i) \geq \frac{\delta}{1-\delta}\Big(\max\{s_i, s^0\} - 2\Big)$ because, if $s_i \geq s^0$, then $\mathcal{C}^0(s_i) = \frac{\delta}{1-\delta}(s_i - 2)$. If, instead, $s_i < s^0$, then one can show that $\mathcal{C}^0(s_i) = \frac{\delta}{1-\delta}(s^0 - 2) + (s_i - \mathbb{E}(s))$ by noting (i) that the players are indifferent between exploring and exploiting when the best project has value s^0 and (ii) that the continuation value following exploration is independent of the value associated with the current best project sine this project will never be exploited.[22] Utilizing this bound on $\mathcal{C}^0(s_i)$, a sufficient condition for Equation (26) is:

$$m \leq \sum_{i=1}^m \frac{\delta}{1-\delta}\Big(\max\{s_i, s^0\} - 2\Big) \iff \frac{1+\delta}{\delta} \leq \frac{1}{m}\sum_{i=1}^m \max\{s_i, s^0\} \tag{27}$$

Noting that the left-hand side is equal to $\tilde{s}$ implies that condition (4) is sufficient.

Necessity: Suppose the players follow the project selection rule as in the symmetric benefits benchmark. Hence, given projects with values $s_1, \ldots, s_m$ in period t, the players exploit all projects i such that $s_j \geq s^0$. For all projects j where $s_j < s^0$, a new exploration takes place. If all such explorations result in projects with values s^0, then, in period $t+1$, the players permanently exploit these projects worth s^0, as well as all projects such that $s_i \geq s^0$. Hence, the players must be able to permanently exploit the following set of projects: $\{\max(s_1, s^0), \ldots \max(s_m, s^0)\}$ in period $t+1$. Using Equation (2), one can see this condition is:

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

which is equivalent to Equation (4). □

Proof of Proposition 5. **Necessity:** First note that, if the players permanently ex-

[22] This calculation notes $s^0 - 2 + \frac{\delta}{1-\delta}(s^0 - 2) = E(s) - 2 + \mathcal{C}^0(s^i)$ where the left hand side is the exploration payoff and the right hand side is the exploitation payoff.

ploit projects with values $s_1, \ldots, s_m$, then Equation (2) implies:

$$m \leq \sum_{j=1}^{m} \frac{\delta}{1-\delta}(s_j - 2). \tag{29}$$

However, Equation (29) implies Equation (4) in Lemma 1. Thus, the players mimic the project selection rule of the symmetric benefits benchmark, which implies $s_j \geq s^0 \quad \forall \quad j$.

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

Sufficient: Condition 2 in Proposition 5 implies that Equation (4) holds, and thus that the project selection rule is that of the symmetric benefits benchmark. Further, Condition 1 in Proposition 5 implies that the symmetric benefits benchmark strategy is permanent exploitation.

□

Proof of Proposition 6. Suppose $s_p \sim 3 - \frac{1}{\lambda} + \exp(\frac{1}{\lambda})$. This 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.

Using the proof of Proposition 5, we know the strategy under the optimal equilibrium conditions on only the two best projects found thus far.

Statement 3: Note that there exists a sufficiently small value of δ such that the players are unable to exploit a project worth $s^0 + \epsilon$. Take such a δ. With positive probability, in period 0 the players identify 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. Since 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 the players behave as in the case with symmetric benefits and thus permanently exploit both projects. Therefore, with positive probability, the players exploit a project they have previously chosen not to exploit.

As argued in the text, Statement 2 implies Statement 1.

Statement 2: To simplify the proof we consider δ sufficiently large such that $\mathcal{C}(s_1, s_2) \geq 1 \forall s_1, s_2$. A sufficient condition on δ is that if the players simply explored

new projects every period, $\mathcal{C}(s_1, s_2) \geq 1$. This condition is:

$$1 \leq \frac{\delta}{1-\delta}(2(3-2)) \iff \delta \geq \frac{1}{3}.$$

Hence, the sufficient condition is $\delta \geq \frac{1}{3}$.

When $\mathcal{C}(s_1, s_2) > 1$ but $f(s_1, s_2) < \tilde{s}$, 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. Further, the only way the players can explore both projects is to find a third project better than s_2. However, in this case, s_2 becomes the third best project and is thus excluded. Hence, we can consider only the best project found when the continuation value belongs to this range.

Given that only the best project matters, one can write $B(s_1, s_2) = B(s_1)$ when $f(s_1, s_2) < \tilde{s}$. Further, one can show that the optimal policy follows a similar threshold rule where a project is exploited if and only if its value is above a threshold. To see this, note that there are two options within this range: explore two new projects or explore one new project and exploit the project with value s_1. The total surplus associated with the exploration of two new projects has a smaller slope with respect to s_1 (because the project with value s_1 is not exploited in the current period) compared with the slope of the total surplus when exploring only one new project (because the project with value s_1 is exploited in the current period). Hence, there exists a threshold where the players switch to exploitation.

Note that $\mathbb{E}(s_p) = 3$ and $\text{Var}(s_p) = \frac{1}{\lambda^2}$. Hence, we will consider λ arbitrarily large ensuring that the variance is arbitrarily small. Now, *suppose by contradiction,* s^*, the value such that the players are indifferent between exploiting one project and exploring two projects is always weakly greater than s^0 when $f(s_1, s_2) < \tilde{s}$.

Note that in the case of symmetric benefits (respectively, asymetric benefits) the players are indifferent between exploiting s^0 (respectively, s^*) and continuing to explore. The equalities implied by these indifferences are shown below where p denotes

$\Pr(s_p < s^*)$ and p^0 denotes $\Pr(s_p < s^0)$:

$$s^0 - 2 + \delta B^0(s^0) = 3 - 2 + \delta\Big(p^0 B^0(s^0) + (1-p^0)\mathbb{E}(B^0(s')|s' > s^0)\Big) \tag{30}$$

$$\begin{aligned} s^* - 2 + (3-2) + \delta\Big(pB(s^*) + (1-p)\mathbb{E}(B(s')|s' > s^*)\Big) + \delta\epsilon_1(\lambda, s^*) \\ = 2(3-2) + \delta\Big(p^2 B(s^*) + 2p(1-p)\mathbb{E}(B(s')|s' > s^*) \\ + (1-p)^2\mathbb{E}(B(\max(s', s''))|s', s'' > s^*)\Big) \\ + \delta\epsilon_2(\lambda, s^*) \end{aligned} \tag{31}$$

The left-hand side of Equation (30) corresponds to exploiting a project worth s^0, whereas the right-hand side corresponds to exploring a new project when benefits are symmetric. With probability p^0, the best project remains s^0 and with probability $1 - p^0$ there is a new best project. The second equation is similar. Equation (31) represents the value of exploiting s^* and exploring one more project. The continuation value has two parts: the first part is conditional on remaining in the region where $f(s_1, s_2) < \tilde{s}$. One can see that, with probability p, the best known project is as before and, with probability $1 - p$, there is a new best known project. Additionally, $\epsilon_1(\lambda, s^*)$ corresponds to the change in continuation value upon finding a project valuable enough that $f(s_1, s_2) < \tilde{s}$. Specifically, $\epsilon_1(\lambda, s^*)$ corresponds to the probability that s' is sufficiently large such that $f(s_1, s_2) > \tilde{s}$, multiplied by the difference in continuation value in this region, as opposed to the continuation value when the continuation value is less than 2.

A similar expression is shown to hold for the case corresponding to the exploration of two 2 new projects. Note that, upon exploring 2 new projects, there are three possibilities: neither project has value exceeding s^*, one project has value exceeding s^*, or both projects have values exceeding s^*. Further, ϵ_2 is defined similarly to ϵ_1 and is the error when assuming that the continuation value remains below 2. Both ϵ_1, ϵ_2 approaches 0 uniformly as $\lambda \to \infty$. These convergences happen because, to reach $f(s_1, s_2) \geq \tilde{s}$, the players must draw a project with value equal to at least 4. Since (i) drawing such a project occurs with probability approaching 0 as $\lambda \to \infty$ and (ii) payoffs decrease and remain bounded as λ to ∞, the ϵ_1 and ϵ_2 terms uniformly decrease.

One can simplify these indifference conditions as follows:

$$s^0 = 3 + \delta(1-p^0)\Big(\mathbb{E}(B^0(s')|s' > s^0) - B^0(s^0)\Big) \tag{32}$$

$$\begin{aligned} s^* = 3 &+ \delta(1-p)\Big[- pB(s^*) + (2p-1)\mathbb{E}(B(s')|s' > s^*) \\ &+ (1-p)\mathbb{E}(B(\max(s', s''))|s', s'' > s^*)\Big] \\ &+ \delta(\epsilon_2(\lambda, s^*) - \epsilon_1(\lambda, s^*)). \end{aligned} \tag{33}$$

Since we assumed by contradiction that $s^* \geq s^0$, the right-hand side of Equation (33) is greater than the right-hand side of Equation (32) when both are evaluated at s^0. This observation implies:

$$\begin{aligned} &(1-p^0)\Big(\mathbb{E}(B^0(s')|s' > s^0) - B^0(s^0)\Big) \\ &\leq (1-p)\Big[- pB(s^0) + (2p-1)\mathbb{E}(B(s')|s' > s^0) \\ &\quad + (1-p)\mathbb{E}(B(\max(s', s''))|s', s'' > s^*)\Big] \\ &\quad + \frac{\epsilon_2(\lambda, s^0) - \epsilon_1(\lambda, s^0)}{1-\delta}. \end{aligned} \tag{34}$$

One can now consider the limit as $\lambda \to \infty$. First note that $\epsilon_i(\lambda, s^*)$ decreases with an exponential rate since the probability of getting a sufficiently high valued project is exponentially decreasing. Using the closed-form solution of s^0, we can show that $s^0 = 3 + \frac{W(\frac{\delta}{e(1-\delta)})}{\lambda}$, where W is the product-log function. With this in mind, we can show that p^0 is constant with respect to λ. Further, given (i) the memoryless property of the exponential distribution and (ii) the fact that the projects with values s' and s^0 will be exploited in all subsequent periods, it follows that $\mathbb{E}(B^0(s')|s' > s^0) - B^0(s^0)) = \frac{\mathbb{E}(s'|s'>s^0) - s^0}{1-\delta} = \frac{1}{\lambda(1-\delta)}$. Hence, the left-hand side is linear in $\frac{1}{\lambda}$. Since we know the $\epsilon_i(\lambda, s^*)$ terms decrease at an exponential rate, this term can be ignored in the limit. Therefore, the inequality reduces to:

$$\begin{aligned} &(1-p^0)\Big(\mathbb{E}(B^0(s')|s' > s^0) - B^0(s^0)\Big) \\ &\leq (1-p)\Big[- pB(s^0) + (2p-1)\mathbb{E}(B(s')|s' > s^0) \\ &\quad + (1-p)\mathbb{E}(B(\max(s', s''))|s', s'' > s^*)\Big]. \end{aligned} \tag{35}$$

Moreover, given our assumption that $s^* \geq s^0$, it follows that $p \geq p^0$. Combining this with the evaluated expression for the left-hand side yields the following implication:

$$\frac{1}{\lambda(1-\delta)} \leq -pB(s^0) + (2p-1)\mathbb{E}(B(s')|s' > s^0) + (1-p)\mathbb{E}(B(\max(s', s''))|s', s'' > s^*). \tag{36}$$

Let us consider $s' > s$ and analyze $B(s') - B(s)$ to be able to bound the above equation. If $s' < s^*$, then $B(s') = B(s^0)$ as in both cases the best project found will not be exploited. This equality omits the exponentially decreasing probability of reaching a region where two projects may be exploited. If $s' > s^*$, then $B(s') - B(s)$ is at most $\frac{s'-s}{1-\delta}$ plus terms that are exponentially decreasing in λ. To see why this inequality holds, note that, conditional on only exploiting one project every period, the largest possible benefit from exploiting a better project occurs from exploiting the better project in every period. Further, there is an exponentially decreasing term corresponding to the exponentially decreasing probability that $f(s_1, s_2)\tilde{s}$. A marginally better project lets the continuation value of the players' relationship exceed this threshold with a greater probability, however both probabilities remain arbitrarily small. These two inequalities imply (i) $\mathbb{E}(B(s')|s' > s^0) \leq B(s^0) + \frac{1-p}{1-p^0}\frac{1}{\lambda}$ and (ii) $p \geq p^0 \geq \frac{1}{2}$ and thus we can replace both expectations on the right-hand side of Equation (36) by the upper-bounds. Finally, note that $\mathbb{E}(\max(s', s'')|s', s'' > s^*))$ is the maximum of two exponential variables conditional on both being greater than s^0. Thus, $\mathbb{E}(\max(s', s'')|s', s'' > s^*)) \leq s^0 + \frac{2}{\lambda} - \frac{1}{2\lambda}$. Applying these observations yields:

$$\frac{1}{\lambda(1-\delta)} \leq -pB(s^0) + (2p-1)\left(B(s^0) + \frac{1-p}{\lambda(1-\delta)(1-p^0)}\right) + (1-p)\left(B(s^0) + \frac{2-\frac{1}{2}}{\lambda(1-\delta)}\right) \tag{37}$$

$$\iff \frac{1}{\lambda} \leq (2p-1)\frac{1-p}{\lambda(1-p^0)} + (1-p)\left(\frac{3}{2\lambda}\right) \tag{38}$$

$$\iff 1 \leq (2p-1)\frac{1-p}{1-p^0} + (1-p)\frac{3}{2}, \tag{39}$$

which generates a contradiction for any $p \geq p^0$. To see why this contradiction occurs, note that, if $p = 1$, the right-hand side is 0. Further, if $p < 1$, $\frac{1-p}{1-p^0}$ is at most 1.

Replacing this expression with 1 reduces the inequality to $1 \leq (2p-1) + (1-p)\frac{3}{2}$. This inequality is strictly increasing in p and is only satisfied at $p = 1$. However, we just assumed $p < 1$, generating our contradiction. □

Proof of Proposition 7. Based on Proposition 1, the optimal continuation equilibrium hinges solely on the number of projects with value v discovered to date.

Moreover, in any optimal relational contract, the number of selected projects, $\boldsymbol{p}^t$, increases weakly over time. This is due to the fact that as the number of projects with value v identified grows over time, it leads to a corresponding increase in the continuation value of the players' relationship.

This finishes the proof of statement 1 of the Proposition. The proof of Statement 2 is provided in the text. □

Proof of Proposition 9. Fix any $\alpha \in \mathbf{N}$. Under the optimal relational contract, equilibrium plays depends solely on some function $f(n)$ satisfying the properties listed in Proposition 7. Moreover, the necessary and sufficient condition for a relational contract with associated function $f(n)$ to constitute an equilibrium is:

$$\begin{aligned} f(n) \leq n\mathcal{C}(\text{exploitation}) + (f(n)-n)\mathcal{C}(\text{exploration}) \\ + \sum_{i=f(n)+1}^{m} \mathbb{E}(\text{wait}_i|n)\mathcal{C}(\text{exploration}) \quad \forall n. \end{aligned} \tag{40}$$

The continuation value stems from three components which we enumerate in order: (i) the n projects that will be exploited; (ii) $f(n)-n$ the projects that will be explored; and (iii) the remaining projects that will be explored upon the discovery of a high-valued project. These projects have a value of $\mathcal{C}(\text{exploration})$, however none of these projects are begun in the current period. Each project can be assigned an expected discounted wait time until being started.

However, since exploration increases payoffs a sufficient condition for an optimal equilibrium is that the players cannot incentivize $f(n)+1$ projects:

$$\begin{aligned} f(n)+1 > n\mathcal{C}(\text{exploitation}) + (f(n)+1-n)\mathcal{C}(\text{exploration}) \\ + \sum_{i=f(n)+2}^{m} \mathbb{E}(\text{wait}_i|n)\mathcal{C}(\text{exploration}) \forall \quad n. \end{aligned} \tag{41}$$

This condition is sufficient but not necessary as (i) it does not consider changing $f(n)$ and $f(n')$ (ii) increasing $f(n)$ decreases the expected wait time for each remaining task whereas the above assumes the wait times remain the same. What we want to show is thus if $f(n-1)$ satisfies the necessary and sufficient condition, $f(n-1)+\alpha$ satisfies the sufficient condition with n confirmed projects. Suppose by contradiction, this does not hold. Namely,

$$f(n-1)+\alpha > n\mathcal{C}(\text{exploitation}) + (f(n-1)+\alpha-n)\mathcal{C}(\text{exploration}) \tag{42}$$

$$+ \sum_{i=f(n-1)+\alpha+1}^{m} E(\text{wait}_i|n)\mathcal{C}(\text{exploration}). \tag{43}$$

However,

$$f(n-1) \leq (n-1)\mathcal{C}(\text{exploitation}) + (f(n-1)-n-1)\mathcal{C}(\text{exploration}) \tag{44}$$

$$+ \sum_{i=f(n-1)+1}^{m} E(\text{wait}_i|n)\mathcal{C}(\text{exploration}). \tag{45}$$

Combining these two inequalities implies

$$\alpha + \sum_{i=f(n-1)+1}^{f(n-1)+\alpha} E(\text{wait}_i|n)\mathcal{C}(\text{exploration}) > \mathcal{C}(\text{exploitation}) + (\alpha+1)\mathcal{C}(\text{exploration}) \tag{46}$$

Finally, one can show this inequality can not hold for q sufficiently small as when q decreases, v must be sufficiently large such that all projects remain profitable in expectation. When this occurs $\mathcal{C}(\text{exploitation})$ must be sufficiently large. Finally, the exploration payoffs on the left hand side always remain strictly less than the exploration payoffs on the right hand side due to discounting. □