A Rational Framework on the Causes and Cures of Collaborative Projects Failure

Zhijian Cui[1], Christoph Loch[2]
[1]IE Business School-IE University, Madrid, Spain
[2]Cambridge Judge School of Business, Cambridge University, United Kingdom

Abstract--**This paper takes a rational perspective to study the causes and cures of collaborative projects failure in the organizations. It shows that project cooperation failure may result from two causes: project members' private information (with respect to their preferences and/or outside options) as well as their incentives to misrepresent---explicitly and implicitly--that information; and failure to build trust. Different from the prior studies, which often attribute project failure to poor skills of project management (e.g., miscommunication, trust building, etc), this study shows that in collaborative projects, it is the *structural* conflicts among project stakeholders that handicap the communication and trust building. In addition, this paper also examines two mechanisms of self-enforcement and their effects on cooperation. First, when the costs of implicit communication are strongly asymmetric, one party may have the incentive to signal her private information in a way that goes beyond "cheap talk". Second, if the project payoff is fixed for one party but potentially higher for the other party, the risk of cooperation failure actually increases because the latter party then bargains more aggressively.**

I. INTRODUCTION

In the past decade it has become increasingly common for projects to involve multiple parties (stakeholders). Between organizations, the collaborative partners usually form an alliance whereby each partner firm contributes efforts according to its specific capability. The motives and benefits of collaborative projects---which include the sharing of costs and expertise, mutual learning, and access to the other party's resources--have been well addressed in the literature. However, such projects tend to be unstable and often yield disappointing outcomes [15]; in fact, many are abandoned before a product is even launched [10]. One study found that 14% of such collaborations either abandoned or delayed their innovation projects because of partnership difficulties; that outcome is often referred to as "cooperation failure" [20] or "alliance instability [8]. Reference [29] reports that 34% of interfirm alliances were viewed as failures and 51% experienced an intermediate outcome in the form of contract expiration or unilateral withdrawal by a partner; furthermore, the risk of cooperation failure does not vary much as a function of partner type (public research organizations, suppliers, competitors, etc.) [20]. In the research cited here and in the text of this paper, failure is defined as the unplanned and unwanted termination of a project [29], which we term as "type I error" in collaborative project management. Type I error is different from another scenario where a project is not terminated although it should be (we term it as "type II error"). The organizational causes of "type II error" or "escalation of commitment" have been well studied in the literature of information system [24] and marketing [34]. For example, Reference [34] argues that bad projects are hard to kill due to the inherent conflict when designing managers' incentives.

Prior research sought explanations for such cooperation failure from the perspective of several theoretical approaches. First, one can argue from the viewpoint of transaction cost theory and focus on the pursuit of self-interest at the partner's expense---in addition to the high costs of deterring such opportunistic behavior---as a major cause of partnership instability [22]. Thus it is claimed that such behavior explains the positive correlation between an alliance member's withdrawal decision and its so-called outside option (i.e., payoff from collaborating with other partners) [13]. Yet this angle leaves unanswered the questions of why and how cooperation failure occurs even when project survival would be beneficial to all the involved parties. Second, one can take a resource-based view of the firm and argue that the unequal resources brought to a partnership by the individual firms eventually lead to a power imbalance between partners that explain the premature ending of their collaboration [29]. According to this school of thought, the parties' bargaining power stems only from their respective resources and so their outside options are not considered. However, this approach likewise fails to identify the mechanism by which an imbalance in bargaining power leads to cooperation failure, an outcome that is not desired by either party. Third, one can adopt the game-theoretical perspective and emphasize the importance of stakeholder conflicts and the role of technological uncertainty in predicting the partners' intentions and calculating payoffs [26]. This approach views cooperation in a collaborative project as a Prisoner's Dilemma in which it is more beneficial for each party to cheat unilaterally than to cooperate bilaterally. Yet that classic payoff structure may not always fairly characterize a collaborative project, where the payoff from mutual and even unilateral collaboration is often greater than from unilateral cheating. This viewpoint, too, leaves some questions unanswered. For instance: Why can't the parties use effective communication to resolve uncertainties and clarify ambiguities [1, 6, 14] and thus minimize the risk of cooperation failure due to miscalculating the other party's intentions? We also know---from analogous situations in supply chains--that partners that interact over a long period of time may have sufficient incentive to compromise in one project for the benefit of future projects. We therefore ask:

Why can't parties build trust through repeated interactions and thereby reduce the likelihood of collaborative failure?

This paper is motivated by these issues, and it seeks to answer the unresolved questions from prior research. One reason that the cited work has been unable to explain the causes of unwanted cooperation failure is their failure to address the management of conflict or “internal tensions” [8]. A specific example will illustrate this point. Three consumer service companies set out to create a customer loyalty program, offering customers points to be cashed in for any of the participating companies' services. They commissioned a complex software system to track points, which required feature choices such as “Given input A, does the system choose X, Y, or Z?” The partner company representatives could not agree on the answers---not because of any technical difficulty but rather because they had different preferences as regards the best way to treat customers. Forced to acknowledge deep incompatibilities among their business interests, they ultimately canceled the project. In what became an expensive and litigious failure, one executive commented: “The stakeholders worked under the illusion that everyone was going to get everything that they wanted… They papered over their differences rather than going through conflict resolution in the early stages, and unresolved conflicts led to failure” [23]. In other cases, billions of dollars have been spent on initiatives, such as creating cross-unit incentives and offering teamwork training, that aim to mitigate conflicts between project members. In the words of one practitioner, however: “While such initiatives yield the occasional success story, most of them have only limited impact in fostering collaboration---and many are failure[s]” [40]. Therefore, we seek to understand not only how such conflicts affect the risk of cooperation failure but also, and more importantly, why conflicts between project members persist.

The paper proceeds as follows. In Section 2 we discuss how conflicts of interest arise and the negotiatory aspect of collaborative project. In doing so we provide a general theoretical background and motivate our particular setup. Section 3 addresses the puzzle posed by the observation that cooperation failure is costly for all project members. We use a simple formalization of the bargaining problem faced by conflicted collaboration partners to show that, under broad conditions, conflicts can be resolved and rational parties should prefer continued collaboration to a costly termination of the project. In Section 4 and 5, we respectively discuss the two causes of cooperation failure and the effects of two self-enforcing mechanisms designed to reduce the risk of cooperation failure. Section 6 concludes. Throughout the paper we will illustrate theoretical arguments using simple game-theoretic representations and will illustrate the plausibility of our contentions with managerial examples excerpted from published case studies.

II. CONFLICT OF INTERESTS AND THE NATURE OF COLLABORATIVE PROJECTS

Projects have the potential to upset an organization's status quo. They operate outside permanent structures---given that a temporary group of diverse individuals is assembled for a particular and often innovative purpose---and sometimes violate existing political relationships and established chains of command. The symptoms of conflict among participating functions or firms are especially visible in collaborative projects, where each involved business unit or partner firm may have a unique background, expertise, and level of access to information [28]. Consider the following remarks made by managers in conversations with the authors. At a company that prides itself on its leading-edge technical products, *A* staff member quipped: “This company is dominated by stubborn engineers.” In contrast, engineers in partnership with another market-driven company complained that “most decisions are made by people with a ‘soft’ mind. … Those people only care about customers, they do not always understand technologies.” Such conflicts can lead to wide-ranging problems: the specification development phase may be drawn out unduly, differences in opinion may escalate to senior management, specifications may be changed during later stages, and/or the project may be abandoned because the parties cannot agree. At worst, projects fail. At best, expensive after-the-fact problem solving manages only to “avoid competitive disaster rather than providing competitive advantage” [41].

Some argue that such conflicts are avoidable if new product design can be settled with reference to (future-oriented) customer tastes. For example, in viewing each product as “a bundle of well-defined attributes” [335], prior research from the marketing perspective seeks to establish an “objective” value of product design by making *A* system optimization decision (the product) based on the firm's collective interest [16, 33]; here, the individual interests of project members are considered to be of secondary concern. In reality, however, there are three main challenges to the plausibility of that approach.

First, customer tastes are seldom truly homogenous, even within the narrow customer segment targeted by a particular product version. Hence it is still up to each firm or function to decide which segment to target. Second, the consumers' own evaluations can be affected by their lack of familiarity with the new product, their uncertainty about its benefits and risks, their ability to understand how it operates, and their perceptions of product safety [32]. Third, it is extremely difficult to represent the overall appeal of products, especially those for which aesthetic and other holistic attributes are important [18]. In light of these challenges, any take on market information or on the “best” product design is necessarily subjective. In the resulting absence of objective shared criteria, it is only natural that the parties in *A* collaborative project will seek to exert their departmental or personal influence [28]. Thus any product design is inevitably a compromise, which is reflected by the characterization of product development as a negotiation process among multiple involved parties [2, 368]. The implication is that a collapse in negotiations among project stakeholders will lead to termination of the project even if its outcome initially seemed promising. This problem is evidently a widespread one; according to a previous study, "technical and cost concerns are not as important as organizational factors [in] contributing to project abandonment" [10].

III. THE INEFFICIENCY OF COOPERATION FAILURE

In this section, we formulate a game-theoretical model that captures the problems faced by each party in a collaborative project. Suppose that our focal partnership is between two independent firms in an product alliance and seeks to develop a new product characterized by a critical, one-dimensional design parameter: the feature parameter *f* with feasible range [0,1]. Each party takes one specific and non-overlapping responsibility in the partnership. We refer to these parties as *A* and *B*. For expositional convenience we often use "he" and "she" with reference to *A* and *B*, respectively. The two parties share a fundamental common interest because they are participating in a collaborative project, yet they have different preferences for the focal design feature---preferences that may result from the parties' access to different information [15]. Therefore, *A* and *B* have symmetric utility functions (reflecting a fundamental common interest) but different optimal solutions (reflecting their individual utility functions). Let each party's utility function be quadratic with respect to the product's feature or quality value (although our results do not depend on this assumption)[1].

The payoffs of the two parties can therefore be described as follows:

$$\Pi_A = P_A - a(f - f_A^*)^2,$$
$$\Pi_B = P_B - a(f - f_B^*)^2.$$

Here P_i denotes each party's maximum payoff potential from successfully delivering the project. At this moment, we assume that parties have equal stakes in the project outcome, i.e., $P_A = P_B = P$. We will relax this assumption in Section 5.2. f_i^* denotes each party's perceived "optimal" product feature maximizing that party's expected payoff. Without loss of generality, let the value of party *B's* favored product feature be strictly greater than that of party *A's*; thus, $f_B^* > f_A^*$. In (1) and (2), the term *A* captures the project profit's sensitivity to the focal product feature *f*. For simplicity, we let *A* =1 throughout this study. The two utility functions are plotted in Figure 1.

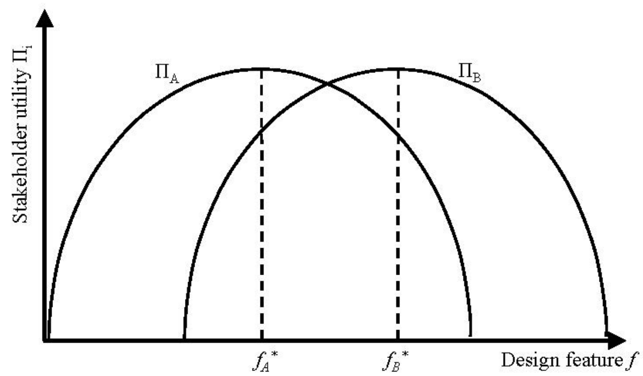

Figure 1: Project Members' Optimal Product Feature Values

Finally, we assume that if cooperation fails then the payoffs for party *A* and party *B* will be their respective outside options or fallback utilities, u_A and u_B. In an alliance, a party's outside option is the "payoff she can get from collaborating with alternative partners" [15]. In practice, the firms' outside options are influenced by their reputation in the industry ---factors that do not change within the short period of a project. Hence the outside option is exogenous to the bargaining process.

Each party I $\in$ {A, B} is described as being of type *t*, which denotes that party's optimal feature value or outside option; we use Ω_i^t to denote the set of product feature values that this party prefers to cooperation failure. For instance, if *t* denotes each party's outside option then $\Omega_i^t = \{f: \prod_i^t(f) \geq u_i^t\}$. Clearly, any feature value within this set yields greater payoffs for the project member than the outside option. Hence the set Ω_i^t establishes a range within which mutual agreement is acceptable for party i; this range is referred to as party *i*'s *cooperation potential*. Figure 2 and Figure 3 illustrate party *A's* cooperation potential Ω_A^t where party *B's* outside option can have only two values (either high or low, H or L). The limit f_A^t is the right endpoint of Ω_A^t, which represents the "best" feature value for party *B* in terms of party *A's* cooperation potential. The figures reveal that, as a party's bargaining power (e.g., value of the outside option) increases, cooperation potential declines and project members take a less compromising stance on the preferred product feature: $f_A^H < f_A^L$.

[1] Quadratic utility functions have been widely used in the design engineering literature (e.g., [42]). We can also show that a shifted quadratic utility representation is general enough to capture both a vertical product feature (i.e., one that drives costs) and a horizontal product feature (one, such as color, that is cost neutral). The detailed discussion is available per request.

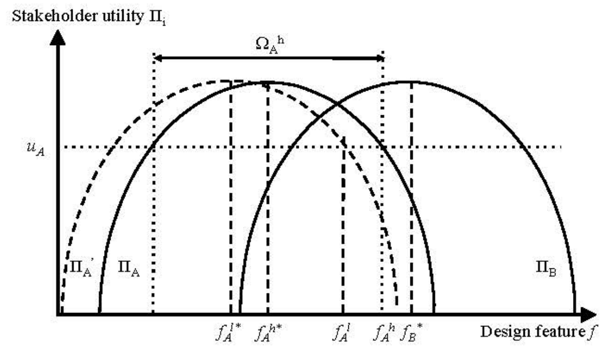

Figure 2 Cooperation Potential for Party A (of Type t) When Preference Varies

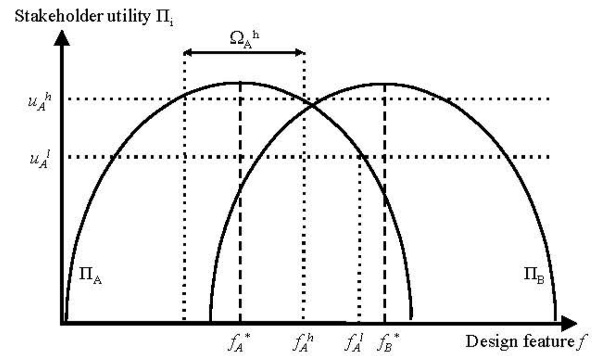

Figure 3 Cooperation Potential for Party A (of Type t) When Outside Option Varies

We can use this formulation to categorize two extreme scenarios. First, when the preferences of two parties are "close" enough (e.g., *B's* optimal feature value f_B* lies within *A's* collaboration potential $\Omega_A{}^t$) or when *A's* outside option is extremely low (and thus corresponds to a relatively wide range of $\Omega_A{}^t$), the parties will agree because otherwise *B* could simply assert her preferred design (f_B*) and know that her partner (party *A*) would have no choice other than to accept. This case corresponds to a region in which there is very little risk of cooperation failure (Scenario 1). The conditions that lead to this scenario holding are empirically plausible. For example, it has also been empirically found that the likelihood of a firm's withdrawal (i.e., a cooperation failure) is positively correlated with the value of that firm's outside option [13]. This finding supports our argument that a low outside option for one party should translate into a low risk of cooperation failure.

In contrast, when parties' preferences are too divergent or if either project member has a high-value outside option, then the payoff from continuing the project may be less than that from pursuing outside option $(P-(f_j^t-f_i^*)-2<u_i)$. This implies that project members would prefer to let the collaboration fail even though that outcome is undesirable from the perspective of the collaboration itself. Here we have a region in which there is a high risk of cooperation failure (Scenario 2). Prior research [6, 26] has suggested that firms in Scenario 2 can take at least two steps to mitigate the high risk of cooperation failure: (i) align the interests of different parties (i.e., reduce the gap between f_A* and f_B*); and (ii) redesign the incentive structure (i.e., change the value of P such that $(P-(f_j^t-f_i^*)-2 > u_i)$..

One might well suppose that carefully designed incentives (e.g., a transfer payment scheme) may help firms significantly reduce the risk of cooperation failure by moving the project from Scenario 2 to Scenario 1. Although this approach has theoretical appeal, from a practical perspective it seems implausible for two reasons. First, any incentive will still be subject to the parties' dynamic preferences and outside options. Therefore, an incentive tailored to Scenario 1 at a given time may prove to be counterproductive later on in the project. Second, it has been observed that, in a complex organization, "an incentive is too blunt an instrument to enable optimal resolution of the hundreds of different trade-offs that need to be made" [40]. For instance, high levels of uncertainty concerning technology and the market will likely prevent firms from designing a contract capable of covering all possible contingencies. We therefore believe that even the most carefully designed incentive may be unable to eliminate all conflicts of interest between parties and hence to sustain Scenario 1.

At this point, the logical next question is: How would project members behave when a project is not characterized by either of these two extreme scenarios? In order to address this question, we make two technical assumptions that rule out those extreme scenarios. First, we assume that party *i*'s most favorable product feature value f_i^{t*} is not acceptable to party *j* and so some bargaining is required; this "no easy win--win" assumption rules out Scenario 1. Second, we assume that, since party *i*'s outside option u_i is structured in a way such that a mutually acceptable solution does exist; this "common ground" assumption rules out Scenario 2. These assumptions may be formally stated as follows.

$$\Pi_A = P_A - a(f - f_A^*)^2, \quad (1)$$

$$\Pi_B = P_B - a(f - f_B^*)^2. \quad (2)$$

The region that satisfies both of these assumptions is referred to as Scenario 3. We then ask whether, in this scenario, the collaborative parties can achieve a mutually beneficial agreement (and let the project continue) and so not suffer from the consequences of the risk of cooperation failure. If they can, then what circumstances lead to that favorable outcome? One general conclusion is immediately evident: there is always a set of product feature values such that both parties prefer any one of these values to cooperation failure (Proposition 1; all propositions are formally stated and proved in Appendix A). Given Assumptions 1 and 2, this first result means that there exists a subset of f such that, for each outcome f in that subset, $\prod_A(f)>u_A$ and $\prod_B(f)>u_B$. In particular, both parties in the project will strictly prefer any agreement in the interval ($f_B^{t'}$, f_A^{t}) to cooperation failure, where *t* and *t'* denote the type of party *A* and party *B*, respectively. In other words: collaboration failure is a negative outcome for both parties and is therefore *inefficient*. So then a puzzle naturally arises: Why would the collaborating parties *not* agree on an outcome and thus avoid the costs of cooperation failure?

IV. CAUSES OF COOPERATION FAILURE IN COLLABORATIVE PROJECTS

Two commonly employed explanations directly address the preceding question. One explanation holds that lacking of information transparency such as outside option or preferences may cause parties in a collaborative project to miscalculate the partner's true intension and willingness to fight over his interest [26]. In our formulation, if *B* is certain about *A's* preference, then she will simply propose f_A^{*t}---a value that maximizes her (i.e., *B's*) own utility and is acceptable to *A*. This school of thought thus highlights the importance of communication. Another explanation focuses on the value of long term interactions between project members and argues that cooperation fails due to lack of trust. In this section we build two arguments. First, while lacking of information transparency point toward a tenable explanation for cooperation failure, it neither goes far enough nor works by itself as it neglect the fact that project members can in principle communicate with each other and so avoid a costly miscalculation of other party's will. Therefore, the cause of cooperation failure cannot be simply lack of information or poor communication skill, but whatever it is that prevents its full disclosure. Our simple formulations show that the fact that project members have (hidden) incentives to misrepresent their true preferences and/or outside options positions is crucial here. Second, we show that due to structural reasons, project members may be unable to build trust even if project members interact over time.

A. Cooperation Failure as a Result of Private Information and Incentives to Misrepresent

In practice, collaborative parties are rarely certain about the true intentions (preferences) and/or interest (outside option) of their partners [1, 28]. Uncertainty with respect to the preferences and/or outside options of others can result from several causes. First, it may be a consequence of parties' different expertise, culture, educational background, and so forth [1], all of which lead parties to access different information [15]. Second, it may be due to a lack of prior collaboration between the parties [6]. Third, uncertainty can arise simply from the bounded rationality of the agents involved. In other words, even though a firm may have known what another firm wants, the firm's employees could easily misinterpret that information and thereby create confusion. In the following, we will examine the effects on collaborative projects of two types of uncertainties, nontransparent preferences and nontransparent outside options.

Case One: Nontransparent Preferences. Let's now suppose party *B* is uncertain about party *A's* optimal product feature value f_A^{t*} but knows that it will be drawn from a cumulative distribution *H(f)* on the nonnegative real numbers with a strictly positive density function h(z) and a nondecreasing hazard rate $\frac{h(f)}{1-H(f)}$[2]. However, only party *A* knows the true value of f_A^{*}. These simple formulations allow us to derive another result: if *B* is uncertain about *A's* preference (f_A^{t*}), then there is a definite likelihood of cooperation failure. Moreover, that probability is always positive when *B's* preference (f_B^{t*}) becomes "too big" for *A* (Proposition 2).

Clearly, party *A* will decline any specification value that yields utility less than his outside option (i.e., any $f > f_A^{t*} + \sqrt{P-u_A}$ will be declined). Therefore, the chance that *A* disagrees with *B's* design proposal f is *Pr* ($f < f_A^{t*} + \sqrt{P-u_A}$) or *H*(f- $\sqrt{P-u_A}$). Hence *B's* expected utility from proposal *f* is *H*(f- $\sqrt{P-u_A}$) u_B + (1- *H*(f- $\sqrt{P-u_A}$)) (P-(f - f_B^{*})^2). Note that the feature value that reduces the probability of disagreement to zero is not optimal from the perspective of party *B*.

Case Two: Nontransparent Outside Options. Suppose that *B* is uncertain about the value of *A's* outside option. In order to derive a tractable and intuitive result, we assume that the outside option for either party can be one of only two values, high or low (H or L). Let $u_i^H > u_i^L > 0$, where $i \in \{A,B\}$. At the outset, party *B* believes with subjective probability q_0 that party *A* has a low outside option, u_A^L; similarly, *A* initially believes with subjective probability p_0 that *B's* outside option is low, u_B^L. Following a proposal concerning f from either party, the other party updates that subjective probability to p_1 or q_1, as applies.

In this setup, party *B* makes an offer f that maximizes

$$\begin{cases} \Pi_B(x) & \text{if } f \in \Omega_A^H, \\ q_0\Pi_B(f) + (1-q_0)u_B^t & \text{if } f \in \Omega_A^L - \Omega_A^H, \\ u_B^t & \text{if } f \notin \Omega_A^L. \end{cases}$$

Under Assumptions 1 and 2, party *B's* proposal f^* should be either f_A^L or f_A^H (anything in between would be inefficient, since it would force the high-outside-option *A* to decline without maximizing the value for *B*). These considerations entail the following equilibrium behavior: when party *B* is uncertain about party *A's* outside option (i.e., about the utility u_A^t), there is a positive probability of cooperation failure. That probability is increasing in party *B's* own outside option (Proposition 3).

Our intuition is that if *B* aggressively proposes the specification f_A^L, which lies closer to her (i.e., to *B's*) own preference, then her proposal will be accepted (resp., rejected)

[2] This condition is satisfied for *A* broad range of distributions; see Reference [12].

by a low-outside-option (resp., high-outside-option) party *A*; in this case, *B's* expected payoff becomes $q_0\prod_B(f_A^L)+(1-q_0)u_B^t$. Therefore, a party *B* of type *t* would aggressively propose f_A^L if and only if *B* has a strong belief ($q_0\prod_B(f_A^L)+(1-q_0)u_B^t > \prod_B(f_A^H)$) that *A* has a *low* outside option. It follows that *B's* threshold probability of acting aggressively depends on her own outside option u_B^t. If *B* has a low outside option, u_B^L, then her threshold probability, q_0^L, will be higher (reflecting a greater willingness to make a utility compromise) than when her outside option is high.

In both cases discussed above, private information seems to be the key driver of cooperation failure. However, a drawback of the simple bargaining model described here is that party *A* cannot communicate his preference or outside option to his project partner. One can imagine that, if *A could* communicate this vital information to his project partner, then doing so might reduce the likelihood of cooperation failure due to information asymmetry---and that this would benefit both parties. In other words, why isn't the private information simply revealed to the other party? To explore this possibility, we next examine the effect of conducting communication via two types of mechanisms between parties: explicit communication (via sending messages) and implicit communication (via making alternative design proposals).

Explicit Communication via Sending Messages. In practice, explicit communication can take the form of sending oral or written messages or having a face-to-face conversation at the meeting of a steering committee. We now modify the baseline game described in *Case One* by allowing party *A* to choose and send (after deciding on his optimal product feature value) a message m from a large but finite set of speeches, *M*. After sending the message, the game proceeds exactly as before with identical payoffs. Party *B* does not know *A's* preference exactly but does know the prior distribution of that preference. Yet it follows that, in any equilibrium where *B* does *not* choose randomly among proposals, she will propose the same product feature value regardless of what *A* says, in which case the risk of cooperation failure remains the same as when there is no explicit communication by *A* (Proposition 4).

To gain an intuition for these results, suppose that party *B* conditioned her behavior on *f*, seeking either more or less as a function of what party *A* communicated. Then, regardless of *A's* true willingness to abandon the collaboration, he does best by communicating his preference in a way that leads to the smallest "grab" by B; that is, *A* has an incentive to misrepresent his actual preference of optimal product feature value. But in that case, *B* learns nothing from *A's* oral communication. In sum, even though both parties prefer to avoid cooperation failure, they also wish to obtain a project outcome that is more favorable from their perspective. This latter desire gives the parties an incentive to exaggerate their true preferences if doing so leads the other side to make concessions. Collaborative parties have a similar incentive to conceal their true outside options if they are concerned that revelation would make them look weak.

Implicit Communication via Proposing Design Alternatives. In the preceding example, we found that when the parties engaged in explicit communication only and so the content of those messages had no direct effect on their respective utility functions, the incentives of stakeholders to misrepresent private information will likely prevent each party from resolving uncertainties about the other party's preferences. A possible explanation for this result is that explicit communication eventually becomes a game of "cheap talk" because the message itself has no binding power. Yet if the message actually had direct consequences for stakeholders, then that dynamic could potentially incentivize project members to reveal their private information[3] .

In organizations, any decision or proposal can be interpreted as a "symbol" that signals the decision maker's concerns, preferences, ability, or even social status [11]. One party can infer the "type" of another party by observing and interpreting the proposal made (the symbol signaled) by the latter. In this way, vital information is not explicitly expressed via messages, but instead implicitly embedded in the project decision process. We refer to such embedded information as *implicit* communication. In this example, we use project members' outside options to explore whether implicit communication gives project members the incentive to reveal their private information.

Suppose that party *B* (resp., party *A*) does not know *A's* (resp., *B's*) true outside option but has a prior probability q_0 (resp., p_0) that *A* (resp., B) has a low-value outside option u_A^L (resp., u_B^L). Suppose that party *A* ---rather than sending *A* message *m* ---proposes *A* specific design alternative *f* to party *B*. If this design is accepted by *B* then both parties are bound by the proposal and so the final product will have feature value *f*. Hence this mechanism has a direct effect on the utility function of both parties. If *A's* proposal is declined, then *B* will make a counter-proposal. The rest of game is identical to the one described in Case Two of Section 4[4]. Our analysis of this scenario reveals that, in equilibrium, party *A* always proposes f_A^H regardless of his true outside option. Party *B's* counter-proposal depends only on her own belief q_0. When *B's* belief q_0 is below a threshold q_0^t, she will accept *A's* initial proposal f_A^H; otherwise, she will decline that proposal and counter-propose f_A^L. Then f_A^L will be accepted (resp., declined) by a low-outside-option (resp., high-outside-option) *A* (Proposition 5).

These results imply that, as in the case of explicit communication, communicating implicitly via proposing design alternatives has no effect on the bargaining outcome.

[3] The key difference between explicit and implicit communications is that the former has no direct effect on either party's payoff whereas the latter does (a design proposal binds both parties if it is accepted).

[4] Technically, this game is related to the broader stream of literature on bargaining with incomplete information [17]. Research in this field has examined "alternative offer" bargaining with an infinite horizon [4, 30,39], where each party can split a "pie" of fixed size (and thus information remains symmetric). Our model extends that approach by considering asymmetric information in a bargaining game with a finite horizon.

In equilibrium, *A* proposes the more advantageous feature f_A^H regardless of his own true outside option, just to "give it a try". Proposing a low product feature value f_A^L initially would mean conceding at the outset, so *A* puts the higher feature f_A^H on the table---just in case *B* is uncertain enough about *A's* outside option to let it pass. As a result, party *B* likewise proposes her preferred feature f_A^L based on her own belief q_0 (i.e., irrespective of what *A* has proposed). Thus, implicit communication does not reduce the risk of cooperation failure. The intuition is that, no matter what *A* thinks about *B's* belief q_0, *A* will not reveal his true outside option because *B* cannot then be prevented from exploiting that information.

The analysis so far establishes that even though project members have good reason to resolve their conflicts and achieve a consensus on product specification value---and even though doing so is more beneficial than allowing the project to fail--there remains the incentive to seek, via bargaining, a product design that most closely matches their own preferences. Because these incentives are at odds, the members of a collaborative project cannot fully communicate their preference or outside option to find a mutually acceptable product specification value. It is the combination of private information (about preferences and/or outside options) and the strategic incentive to misrepresent such information that explains why cooperation can fail despite its potential benefits to both parties.

B. Cooperation Failure Due to Lack of Trust

Trust has been regarded as another important project management principal to coordinate stakeholders in collaborative projects. It has been argued that trust can be a natural constraint on stakeholders' opportunistic behavior and key to resolving conflicts between project partners in a collaborative project [24]. In organization literature, trust is often defined as the common belief among group members that "a particular member will behave in accordance with the commitments, will be honest in the negotiations preceding those commitments, and will refrain from taking undue advantage of another" [7]. Building trust is especially important when there is no third party, such as top management, to ensure that agreements are enforced [9]. In a product alliance, for example, each firm is a legally independent entity and so decides what it says and does according to its own interest.

Building trust is clearly a dynamic process involving project members that will interact repeatedly and consider the future effects of past and current actions. Therefore, we must modify our one-shot bargaining model so that it, too, is dynamic. Toward this end, suppose party *B* is able to propose a specification value for the new product in each of an infinite number of successive periods $t=1,2,\ldots$. In response to *B's* proposal, *A* can either accept and implement that proposal f_t or decline it and pursue his (i.e., *A's*) outside option u_A (then game ends). We discount the cost of receiving a *future* payoff by the factor $\delta\in(0,1)$. Therefore, if one party decides to terminate the collaboration in period t then that period's expected payoff for both parties is $u_i/(1-\delta)$, where $i \in\{A,B\}$.

Let's consider the simple case in which, at the start of collaboration, party *A's* optimal specification value is f_A^{1*}. In the next period, *A* collects new information and his preference changes to f_A^{2*}, where it remains for all subsequent periods. Suppose $f_A^{2*}>f_A^{1*}$. In the unique subgame-perfect equilibrium, party *B* should propose $f_t=f_A^{2*}+\sqrt{P-\frac{u}{1-\delta}}$, the highest value to which *A* will agree. The implication of this proposal is that---during the first period, when *A's* preference is relatively far away from that of *B* and so the latter has the stronger bargaining position---party *A* is choosing between immediate cooperation failure and agreeing to f_1, which yields $P-(f_1-f_A^{1*})^2+\frac{\delta}{1-\delta}[P-(f_2-f_A^{2*})^2]$. Clearly, the best that *B* can do from *A's* perspective in the first period is setting $f_1=f_A^{1*}$. So for *A*, the largest expected payoff from agreeing with *B's* proposal during the first period is $P+\frac{\delta}{1-\delta}[P-(f_2-f_A^{2*})^2]$. Yet one can easily check that under certain conditions (e.g., when $\frac{1-2\delta}{(1-\delta)^2}u_A>P$) it is *beneficial*, for A, to decline *B's* proposal at the beginning and so to allow the collaboration to collapse. Roughly speaking, if party *A's* decline in bargaining power is too much relative to his gain from the project, then the inability of party *B* to refrain from exploiting *A's* future interest makes it reasonable for him to terminate the collaboration.

At this juncture it is worth stressing several points about the difficulty of building trust between project members. First, although we look only at the case where the *preference* of each party changes dynamically, the same result holds for the case where the *outside option* of each party varies from one period to the next. Second, misrepresented private information and failure to build trust are two distinct causes of cooperation failure. In our formulation, project members understand each other's preferences perfectly well and there is no private information. The lack of trust arises from a structure of preferences and payoffs that gives one party an incentive to abandon a mutually beneficial relationship. In a broad range of cases, in particular when $\frac{1-2\delta}{(1-\delta)^2}u_A < P$, it is possible to build trust. For example, as party *A's* outside option u_A decreases, his willingness to continue the collaboration becomes more credible to party *B*.

C. Illustrative Example

An example in the context of an information technology (IT) project failure [25] illustrates the relevance of our insights in reality. In 1998, the procurement manager at Electro Co. proposed setting up "E-PRO" (an electronic procurement system) to help its procurement department develop worldwide sources of cheaper materials. This project received strong support from the top management at Electro Co., who believed that it would expected little opposition: "I could not see why our partners would oppose [such an advanced technology,] which could also benefit them in terms of faster delivery of purchasing orders and payment". The procurement manager was nominated as the manager of this collaborative project.

Once launched, the project was strongly resisted by the system users, namely, Electro Co.'s client firms and suppliers. According to an IT analyst at the firm: "The users were uncooperative... They gave us conflicting information which, in the end, we had to spend a lot of time comparing with the purchasing manual." In addition, several rumors regarding the project quickly spread among the suppliers; the most disturbing of these was that Electro Co. intended to reduce the number of its existing suppliers. In addition, one supplier secretly complained that "another of our customers had earlier requested our participation in their e-procurement system and we did [participate]. We only found out later that there were lots of hidden costs."

The users and suppliers requested that the project to be postponed, giving superficial reasons such as technical incompatibility between their respective companies' systems and Electro Co.'s new system. Even worse, news of the suppliers' objections affected the project members at Electro Co.. The users and suppliers also approached the managing director of Electro Co. with their concerns. Other managers at Electro Co. chose to support the users and suppliers: "I think they needed more time to upgrade their systems to meet the compatibility requirement." In the end, the project was terminated.

Two causes of cooperation failure we established so far (the combination of private information and the incentive to misrepresent together with lack of trust) provide us with different lens through which to examine and explain the procurement IT project failure described in [25]. First, despite project members' private concerns about potential hidden costs and their own continued status as Electro Co. suppliers, they publicly complained about system compatibility, an issue that eventually led to project termination. So in this case the true preferences of all parties were never fully revealed, intensifying the conflicts and leading to failure of the collaboration. Second, neither users nor suppliers trusted Electro Co.'s publicly stated intention. Users were concerned that it might take advantage of shared information if the system were successfully implemented. From a supplier's perspective, the burden of sharing development costs and the risk of future exclusion was inevitable if the system were installed. Thus users and suppliers both believed that they could be relegated to a disadvantageous position once the system was installed, so they preferred to kill it preemptively.

V. STRATEGIES TO COORDINATE COLLABORATION PARTNERS

So far we have developed two arguments to explain the causes of cooperation failure: (i) private information about preferences and/or outside options combined with the incentive to misrepresent them, and (ii) lack of trust. In this section we will discuss some strategies for coordinating partners and thereby dealing with these two concerns. One strategy that immediately comes to mind is that of making all information (e.g., each party's preference and outside option) completely transparent. Indeed, if there is no information asymmetry then project members can no longer benefit from misrepresentation. In that case, one cause of cooperation failure would be eliminated. However, implementing such a strategy is rarely plausible in practice because it requires the support of third party which does not have interest in the project. Hence we need a *self-enforcing* mechanism--that is, one whose enforceability does not require mediation by a third party (e.g., top management). Here we shall examine two such mechanisms that follow naturally from the preceding arguments.

A. Self-Enforcing Mechanism 1: Asymmetric Costs of Proposing Design Alternatives

We have already shown that neither explicit nor implicit communication provides sufficient incentive for project members to reveal their private information, which would be necessary to render information symmetric. A drawback of our previous communication models is that the *cost* of communication is assumed to be symmetric (i.e., normalized to zero for both parties). One might therefore ask whether communication would be more informative if the parties incurred asymmetric communication costs. However, in practice, it is hard to imagine that it costs one party more than another to send messages (e.g., via e-mails, memos, or meetings). It seems more plausible that the cost of proposing design alternative is asymmetrical. Thus, the party with more in-house capabilities or a broader network will need to spend less time (than will its relatively disadvantaged counterpart) collecting information and justifying evidence. That being said, it is not a trivial matter to assess whether such asymmetric costs can solve the problems resulting from information asymmetry. To keep the analysis as simple as possible, we now revise our formulation in Section 4 so that party *B* incurs a positive cost *e* when making a proposal to party *A*. All other aspects of the case are the same as in Section 4 (Implicit Communication via Proposing Design Alternatives). We normalize *A's* cost of making a proposal to zero; hence the parameter *e* reflects the asymmetry between project members with respect to communication costs.

We define $\hat{f}_A^L$ as the feature for which $\prod_B(f_A^L) - e = \prod_B(\hat{f}_A^L)$, and we define $\hat{f}_A^H$ equivalently. Clearly, $\hat{f}_A^L$ is the product feature f_A^L that *B* is willing to accept in order to avoid the proposal cost *e* . Figure 4 demonstrates the effect of an asymmetric cost to proposing a design alternative.

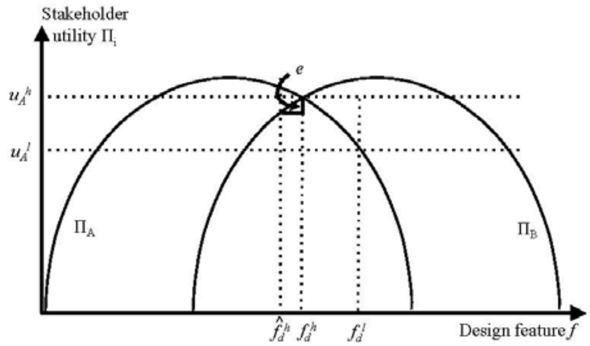

Figure 4. Asymmetric Cost of Proposing a Design Alternative

Weakly Asymmetric Communication Costs: Small e. We first examine the case where *e* takes a small value. It turns out that *A* low cost asymmetry *e* gives *A* the power to propose a slightly lower feature value, but this changes neither the nature of the equilibrium nor the risk of cooperation failure (Proposition 6).

Intuitively, the asymmetric cost of making proposals increases the bargaining power of party *A* in that he can now propose a more aggressive feature value (i.e., one shifted toward his own preference); this is because he knows that party *B* must incur a cost to override that proposal irrespective of her (i.e., party *B's*) outside option. In equilibrium, if the cost asymmetry *e* is small, then *A's* initial proposal is indeed shifted but *B's* counter-proposal is not (because *A's* proposal is still uninformative). More importantly, the risk of cooperation failure does not change: it is still zero if *B* believes that *A* has a high outside option and $1\text{-}q_0$ if *B* is confident (beyond a threshold value) that *A* has a low outside option.

Strongly Asymmetric Communication Costs: Large e. We have just shown that an asymmetric communication cost shifts some of *B's* bargaining power to *A*, pushing the equilibrium feature value toward the latter's preference, but does not change the managerial problem or the risk of cooperation failure. However, if one party faced a *large* cost of communicating then perhaps the risk of cooperation failure could be substantially reduced (or eliminated altogether). By "large" we mean that the shifted design feature $\hat{f}_A^L$, as shown in Figure 4, moves so far leftward that it becomes smaller than f_A^H.

We can show that, if the cost asymmetry is that large, then party *B* is willing to compromise enough (in order to avoid the high cost of making a counter-proposal) that $\hat{f}_A^L$ moves into the feature range acceptable to A_H's--that is, the range for high-outside-option party *A* (Proposition 7). This shift eliminates some of the bargaining conflict, essentially by nullifying Assumption 1 that there is no easy win--win. Party *A* is then willing to signal his true type (L or H) via his proposal. Thus, the surprising effect emerges that a more strongly asymmetrical communication cost may benefit party *B*. On the one hand, it forces her to compromise more by offering a design feature $\hat{f}_A^L$ or $\hat{f}_A^H$ that is closer to *A's* preference; on the other hand, that compromise reduces aggressiveness and thereby reduces the likelihood of cooperation failure.

B. Self-Enforcing Mechanism 2: Asymmetric Stakes in the Outcome

Another potential drawback of our previous models is that the *stakes* of both parties are assumed to be symmetric: the only difference between the two parties' utility functions is the shifted optimal feature value. Yet one might well suppose that, in the case of unequal stakes, the project member with a greater stake in the outcome may have more incentive to avoid cooperation failure. That party may therefore be more inclined to compromise than the party with a lesser stake. We now turn to considering this possibility.

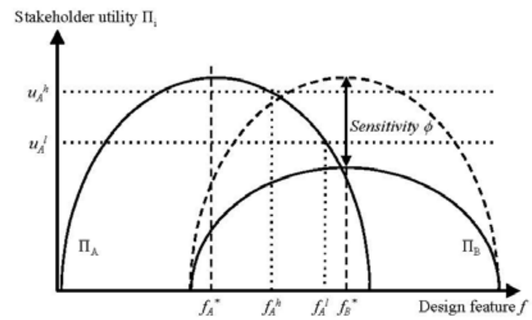

Figure 5. Varying the Relative Stakes of the Two Parties

Figure 5 illustrates the case where party *B's* utility function is scaled by *A* factor of $\varphi > 0$ and party *A's* utility function is normalized to unity. Note that φ denotes *relative* sensitivity and captures whether *B's* sensitivity to the project's payoff increases or decreases in response to organizational incentive policies. We therefore modify equation (2) to read

$$\Pi_B = \phi[P - a(f - f_B^*)^2].$$

When $\varphi > 1$, party *B* has a greater stake in the project outcome and so her utility from obtaining the preferred design decision increases; when $\varphi < 1$, we have the opposite case in which *B* cares less about the outcome. In order to focus the discussion on one phenomenon at a time, we address the basic structure in the absence of communication--that is, the same setup as in Case One and Two of Sections 4.

Surprisingly, we find that increasing party *B's* relative stake in the project outcome (while holding party *A's* stake constant) has the effect of *increasing* the likelihood of cooperation failure (Proposition 8). This is because, when *B* has more at stake, she will insist even more aggressively on her preferred product feature and will be less willing to compromise---which, in turn, increases the risk of rejection by *A*. Conversely, if *B* cares less then she will be more willing to compromise. The implication is that, when $\varphi > 1$, party *B's* utility becomes so steep that even *A* small change in specification value significantly alters her expected utility, which increases the likelihood of cooperation failure. Another way of looking at this result is to allocate an extra payoff (e.g., a bonus); in theory, this should reduce *B's* relative sensitivity to the project's payoff. That is, if *A's* stake is increased then he can more credibly threaten to reject *B's* proposal, inducing her to compromise.

A more extreme conclusion would involve reducing φ to zero: a party who cares not a whit will simply accept any proposal and thus preclude cooperation failure. This is consistent with the organizational rule of thumb that parties who can offer no substantive input or information (and thus have nothing at stake) should not be allowed to influence the outcome of a collaborative decision. However, if we assume that the design preferences f_i of our two parties *A* and *B* reflect relevant considerations, then neither party's preference should be ignored by setting sensitivity at zero. Thus φ serves to balance the inclusion of relevant information (truth-finding aspect) against the danger of a breakdown in consensus (strategic aspect) that could lead to cooperation failure.

VI. CONCLUSION AND DISCUSSIONS

The paper develops three major claims. First, we argue that the management of conflict has not been well addressed in prior research that seeks to explain the high failure rates of collaborative projects. From the perspective that decision making in new product development amounts to a negotiation process, we show that cooperation failure is indeed an unwanted outcome for all project stakeholders. This result implies that there should exist a range of "collaboration potential" in which project members would prefer cooperating over abandoning the project. Second, we offer two causal explanations for why project members are sometimes unable to resolve their conflicts even though doing would benefit all parties: (i) the combination of private information (about true preferences and outside options) and incentives to misrepresent that information, and (ii) the failure to build trust in specific circumstances. We show that although collaborative partners are typically motivated to avoid the negative consequences of cooperation failure, they may also be inclined to misrepresent their private information so that the final outcome of the project is closer to their preferences. Given these mixed incentives, it follows that communication---whether *explicit* (via sending messages to reveal party's preference and/or outside option) or *implicit* (via proposing design alternatives which embeds private information) may actually prevent project members from understanding the other party's intention and from resolving conflicts in a way that does not risk cooperation failure. Contrary to the conventional game theoretical argument that such failure results from miscalculation of others' intentions due to poor communication [26], we argue that it results from strategic dynamics due to the combination of asymmetric information and incentives to misrepresent. In addition, we establish that even repeated interactions may not be enough for project members to build trust and a truly cooperative relationship. If collaborative partners are unable to solve their conflicts, then the project is doomed because at least one of the parties still has an incentive to renege on the terms. Third, we examine the effects of two self-enforcing mechanisms on reducing the risk of cooperation failure: an asymmetric cost to proposing design alternatives and asymmetric stakes in the project outcome. We show that if asymmetry (between parties) in the cost of communication is large enough then parties may be more inclined to make truly informative proposals, which precludes cooperation failure due to information asymmetry. We also find that increasing the project stake of one party relative to another actually heightens the risk of cooperation failure.

It could be argued that there should be a "truth-finding" aspect of communication within collaborative projects [21,35]. For example, prior research in concurrent engineering has shown that communication helps to reduce the negative effect of design rework when an "optimal" communication pattern is followed [21]. Since parties with divergent preferences bring with them different information sets (with respect to markets, manufacturing, etc.), it follows that aggregating such information can increase the overall performance of a collaborative project [1,14]. We concur with this argument. However, the element that we wish to emphasize is the *strategic* aspect of explicit communication, which has been often overlooked in the literature. Note that the existence of truly useful information (e.g., concerns about market access, manufacturing costs, etc.) contained in explicit communication need not alter a stakeholder's incentive to misrepresent some private information, which can lead to the persistence of conflict between the parties. In short, explicit communication between project members is partially informative (the truth-finding aspect) but also partially uninformative (the strategic aspect)[5]. It is the latter that results in cooperation failure.

Alternatively, one could consider Nash bargaining to study the cooperation failure. Yet that approach requires that the parties have complete information and can thereby avoid an equilibrium breakdown [12], and it is precisely the risk of such a breakdown that interests us. Cooperative games are also not applicable to our context because the existence of a core strategy set that is both nonempty and stable requires all involved parties to have complete knowledge of the expected payoffs [27].

The model in this paper is quite stylized: the detailed decision making process between two collaborative parties is simplified as a two parties' bargaining game of product specification value on a one-dimensional space. First, we acknowledge that most products occupy a multidimensional

[5] See [5] for a general discussion of strategic information transmission between two rational agents. It has been shown that information transmission becomes more informative *only* when the two parties' preferences become more congruent.

feature space; however, these multiple dimensions can also be transferred one-dimensional "composite" feature. This approach is frequently taken in the product development literature; see, for example, [19] and [36]. Second, we show in the paper that the results from our simple model are quite robust in different managerial contexts and settings – for example, it can be extended to the case with multiple-period repeated interactions (e.g., failure to build trust), asymmetric information with respect to both preference and outside option, asymmetric cost of proposing design alternatives and asymmetric stakes in project outcomes.

The insights from this study also have broad managerial implications. Prior research often attributes collaboration failure to manager's poor management skills. As a result, a lot of efforts have been spent on the training programs such as trust building, effective communication skills, etc. In contrast to this view, this paper points out the structural and underlying causes which handicap communication and trust building. We show that as long as the conflicts of interests among stakeholders persist, effective communication and trust cannot be achieved, even with excellent project management skills. This then leads to a different set of managerial prescriptions: redesigning project incentive structure according to the decision power (asymmetric stake in project outcome) and selecting a proper partner (with asymmetric cost of proposal design alternative).

In this paper, we implicitly treat project members as rational agents. Even so, we do not mean to rule out other causes of cooperation failure---including both psychological and sociological reasons as well as irrational agents [8, 16, 20, 23]. There is little doubt that such explanations are both important and empirically relevant, but we will not be in a position to explore them until the causal mechanisms for the "rational" case have been rigorously specified. For instance, after clarifying these distinctions we may find that bounded rationality is no less important than private information when it comes to explaining disagreements about optimal product features. Thus a better understanding of how cooperation failure arises among rational agents makes it easier to appreciate the significance of psychology or irrationality in these and similar bargaining games.

REFERENCES

[1] Atuahene-Gima, K., F. Evangelista; "Cross-functional inuence in new product development: An exploratory study of marketing and R&D perspective," *Management Science,* Vol 46, No.10, pp. 1269-1284, 2000.

[2] Bucciarelli, L; "Designing Engineers", MIT Press, Cambridge, MA, 1994.

[3] Chao, R., S. Kavadias; "A theoretical framework for managing the new product development portfolio: When and how to use strategic buckets," *Management Science*, Vol. 54, No.5, pp. 907- 921, 2008.

[4] Chatterjee, K., L. Samuelson; "Bargaining with two-sided incomplete information: An infinite horizon model with alternating offers," *The Review of Economics Studies* Vol.54, No.2, pp.175-192, 1987.

[5] Crawford, V., J. Sobel; "Strategic information transmission," Econometrica, Vol. 50, No.6, pp. 1431-1451, 1982.

[6] Cui, Z., C. Loch, B. Grossmann, R. He; "How the provider selection and management contribute to successful innovation outsourcing: An empirical study at Siemens," Production *and Operations Management,* Vol. 21, No.1, pp. 29-48, 2012.

[7] Cummings, L., P. Bromiley; "The organizational trust inventory (OTI): Development and validation". R. Kramer, T. Tyler, eds., *Trust in organizations: Frontiers of theory and research*. Sage Publications, Thousand Oaks, CA, 302-330. 1996.

[8] Das, T., B. Teng; "Instabilities of strategic alliance: An internal tension perspective," *Organization Science* Vol.11, No.1, pp. 77-101, 2000.

[9] Doz, Y.; "The evolution of cooperation in strategic alliance: Initial conditions of learning processes?" *Strategic Management Journal* ,Vol.17, No.2, pp. 55-83, 1996.

[10] Ewusi-Menshah, K., Z. Przasnyski; "On information systems project abandonment: an exploratory study of organizational practices,". *MIS Quarterly* ,Vol.15, No.1, pp. 67-86, 1991.

[11] Feldman, Martha S., James G. March; "Information in organizations as signal and symbol," *Administrative Science Quarterly* ,Vol.26, No.2, pp. 171-186, 1981.

[12] Fudenberg, D., J. Tirole; *Game Theory*. MIT Press, Cambridge, MA, 1991.

[13] Greve, H., H. Mitsuhashi, J. Baum; "Greener pastures: Outside options and strategic alliance withdrawal," *Organization Science* ,Vol. 24, No.1, pp. 79-98, 2013.

[14] Griffin, A., J. Hauser; "Integrating R&D and marketing: *A* review and analysis of the literature," *Journal of Product Innovation Management* ,Vol.13, No.3, pp.191-215, 1996.

[15] Groves, T; "Incentives in teams," *Econometrica,* Vol.41, No.4, pp. 617-631, 1973.

[16] Gu, X., J. Renaud, L. M. Ashe, S. Batill, A. Budhiraja, L. Krajewski; "Decision based collaborative optimization," *Journal of Mechanical Design* ,Vol.124. No.1, pp. 1-13, 2002.

[16] Hutchison-Krupa, J., R. Chao; "Tolerance for failure and incentives for collaborative innovation," *Production and Operations Management* forthcoming, 2013.

[17] Kennan, J., R. Wilson; "Bargaining with private information," *Journal of Economics Literature* ,Vol.31, No.1, pp. 45-104, 1993.

[18] Krishnan, V., K. Ulrich; "Product development decisions: A review of the literature," *Management Science* ,Vol.47, No.1, pp. 1-21, 2001.

[19] Lacourbe, P., C. Loch, S. Kavadias; "Product positioning in a two-dimensional market space," *Production and Operations Management* ,Vol.18, No.3, pp. 315-332, 2009.

[20] Lhuillery, S., E. Pfister; "R&D cooperations and failures in innovation projects: empirical evidence from French CIS data," *Research Policy* ,Vol.38, No.1, pp. 45-57, 2009.

[21] Loch, C. H., C. Terwiesch; "Communication and uncertainty in concurrent engineering," *Management Science* , Vol.19, No.3, pp. 185-198, 1998.

[22] Lokshin, B., J. Hagedoorn, W. Letterie; "The bumpy road of technology partnership: understanding causes and consequences of partnership mal-functioning," *Research Policy* , Vol.40, No.2, pp. 297- 308, 2011.

[23] May, L.; "Major causes of software project failures," *The Journal of Defense Software Engineering,* Vol.11, No.7, pp. 1-13, 1998.

[24] Newman, M., R. Sabherwal; "Determinants of commitment to information systems development: A longitudinal investigation,". *MIS Quarterly*, Vol.20, No.1, pp. 23-54, 1996.

[25] Pan, G.; "Information systems project abandonment: A stakeholder analysis," *International Journal of Information Management* ,Vol.25, No.2, pp.173-184, 2005.

[26] Parkhe, A.; "Strategic alliance structuring: A game-theory and transaction cost examination of inter-firm cooperation" *Academy of Management Journal* ,Vol.36, No.4, pp. 794-829, 1993.

[27] Peleg, B., P. Sudholter; *Introduction to the theory of cooperative games*. 2nd ed. Springer, Berlin, Germany. 2007.

[28] Pinto, M. B., J. K. Pinto, J. E. Prescott; "Antecedents and consequences of project team cross-functional cooperation," *Management Science* ,Vol. 39, No. 10, pp.1281-1297, 1993.

[29] Reuer, J., M. Zollo; "Termination outcomes of research alliances," *Research Policy,* Vol. 34, No.1, pp.101-115. 2005.

[30] Rubinstein, A.; "A bargaining model with incomplete information about time preferences," *Econometrica* ,Vol.53, No.5, pp.1151-1172, 1985.
[31] Sadowski, B., G. Duysters; "Strategic technology alliance termination: an empirical investigation," *Journal of Engineering Management* ,Vol.25, No.2, pp. 305-320, 2008.
[32] Saunders, M., C. Seepersad, K. Holtta-Otto; "The characteristics of innovative mechanical projects," *Journal of Mechanical Design* ,Vol.133, No.1, pp. 1-9, 2011.
[33] Shi, L., S. Olafsson, Q. Chen; "An optimization framework for product design," *Management Science,* Vol.47, No.12, pp.1681-1692. 2001.
[34] Simester, D., J. Zhang; "Why are bad products so hard to kill?" *Management Science* ,Vol.56, No.7, pp.1161-1179, 2010.
[35] Srinivasan, V., W. S. Lovejoy, D. Beach; "Integrated product design for marketability and manufacturing," *Journal of Marketing Research ,*Vol.34, No. 1, pp. 154-163, 1997.
[36] Terwiesch, C., C. H. Loch; "Collaborative prototyping and the pricing of custom designed products," *Management Science* ,Vol.50, No.2, pp. 145-158, 2004.
[37] Terwiesch, C., C. H. Loch, A. De Meyer; "Exchanging preliminary information in concurrent engineering," *Organization Science ,Vol.*13, No.4, pp. 402-419, 2002.
[38] Ulrich, K. T., S. D. Eppinger; *Product Design and Development*. 3rd ed. McGraw Hill, New York, 2003.
[39] Watson, J.; "Alternative-offer bargaining with two sided incomplete information," *The Review of Economics Studies* ,Vol.65, No.3, pp. 573-594, 1998.
[40] Weiss, J., J. Hughes; "Want collaboration? accept - and actively manage – conflict," *Harvard Business Review* ,Vol 68, No.3, pp.1-10, 2005.
[41] Wheelwright, S., K. Clark; "Creating project plans to focus product development," *Harvard Business Review* ,Vol.70, No.2, pp. 67-83, 1992.
[42] Williams, N., P. Kannan, S. Azarm; "Retail channel structure impact on strategic engineering product design," *Management Science* ,Vol.57, No.5, pp. 897-914, 2006.

APPENDICES

A. Propositions and Proofs

In Propositions 2, 3, and 8, the bargaining between parties *A* and *B* is modeled as a take-it-or-leave-it game. In this game, party *B* moves first, proposing a design feature $f \in [0,1]$. Party *A* then evaluates that proposal and chooses whether or not to accept it. If *A* does not accept, then the game ends and both parties receive their outside option; if *A* does accept, then *A* and *B* receive the utilities defined in equations (1) and (2), respectively.

In Proposition 4, party *A* can send a message before party *B* makes her proposal. The rest of game structure is identical to the game in Propositions 2, 3, and 8.

In Propositions 5, 6, and 7, party *A* can *propose* a *product feature* before party *B* makes her proposal. The rest of game structure is identical to the game in Propositions 2, 3, and 8.

PROPOSITION 1. *There always exists a set of product feature values such that both parties prefer any one of these values to cooperation failure.*

It is sufficient to demonstrate the existence of a nonempty interval $[f_B^t, f_A^t] \subseteq [0, \bar{f}]$. Suppose that $f_B^t > f_A^t$, where f_B^t is the value such that $P - (f_B^t - f_B^{t*})^2 = u_B$ and $f_B^t < f_B^{t*}$. It follows that, if $f_B^t > f_A^t$, then $P - (f_A^t - f_B^{t*})^2 < u_B$; this contradicts Assumption 2 (common ground assumption). □

PROPOSITION 2. *The unique perfect Bayesian equilibrium of the game described in Section 4.1is characterized as follows. Party B proposes f^*, and party A decides to terminate the collaboration if and only if $f_A^{t*} < f - \sqrt{P - u_A}$. Party B's proposal f^* is defined as follows:*

$$f^* = \begin{cases} \sqrt{P - u_A} & \text{if } h(0) > \frac{2(f_B^* - \sqrt{P - u_A})}{2f_B^* \sqrt{P - u_A} - f_B^{*2} + u_A - u_B)}, \\ 1 & \text{if } \frac{h(1 - \sqrt{P - u_A})}{1 - \mathrm{H}(1 - \sqrt{P - u_A})} < \frac{2(f_B^* - 1)}{P - (1 - f_B^*)^2 - u_B}, \\ \text{the unique solution to } \frac{h(f - \sqrt{P - u_A})}{1 - \mathrm{H}(f - \sqrt{P - u_A})} = \frac{2(f_B^* - f)}{P - (f - f_B^*)^2 - u_B} & \text{otherwise.} \end{cases}$$

Moreover, the risk of cooperation failure is always positive for large enough f_B^*.

That party *A* chooses to decline if and only if $f_A^{t*} < f - \sqrt{P - u_A}$ is immediately implied by subgame perfection. In any equilibrium, then, the probability that *A* rejects *B*'s proposal f is $\Pr(f_A^{t*} < f - \sqrt{P - u_A}) = \mathrm{H}(f - \sqrt{P - u_A})$. Thus *B*'s expected utility from proposing f is $u_B(f) \equiv \mathrm{H}(f - \sqrt{P - u_A})u_B + (1 - \mathrm{H}(f - \sqrt{P - u_A})(P - (f - f_B^*)^2)$. Differentiation shows that $u_B'(f)$ is nonnegative only when $\frac{h(f - \sqrt{P - u_A})}{1 - \mathrm{H}(f - \sqrt{P - u_A})} \leq \frac{2(f_B^* - f)}{P - (f - f_B^*)^2 - u_B}$.

We can conclude that: (i) $f^* = \sqrt{P - u_A}$ maximizes $u_B(f)$ if $h(0) > \frac{2(f_B^* - \sqrt{P - u_A})}{2f_B^* \sqrt{P - u_A} - f_B^{*2} + u_A - u_B)}$; (ii) if $\frac{h(1 - \sqrt{P - u_A})}{1 - \mathrm{H}(1 - \sqrt{P - u_A})} < \frac{2(f_B^* - 1)}{P - (1 - f_B^*)^2 - u_B}$, then $f^* = 1$ maximizes $u_B(f)$; and (iii) any $f^* \in [\sqrt{P - u_A}, 1]$ that solves $\frac{h(f - \sqrt{P - u_A})}{1 - \mathrm{H}(f - \sqrt{P - u_A})} = \frac{2(f_B^* - f)}{P - (f - f_B^*)^2 - u_B}$ will be a unique maximum of $u_B(f)$. Since the ex ante probability of cooperation failure is $\mathrm{H}(f - \sqrt{P - u_A})$, it follows that only in case (i) can this probability equal zero—and for large enough f_B^* close to $\sqrt{P - u_A} + \sqrt{P - u_B}$, case (i) cannot obtain. □

PROPOSITION 3. *Define the threshold probability of party B_t with outside option u_B^t as*

$$q_0^t = \frac{\Pi_B(f_A^H) - u_B^t}{\Pi_B(f_A^L) - u_B^t}. \tag{C.1}$$

Then the unique equilibrium of the game described in Section 4.2 is characterized as follows. Party B_t with outside option u_B^t proposes the specification value

$$f = \begin{cases} f_A^L & \text{if } q_0 > q_0^t, \text{ in which case the probability of rejection by } A \text{ is } 1 - q_0; \\ f_A^H & \text{if } q_0 \leq q_0^t, \text{ in which case party } A \text{ will accept.} \end{cases}$$

Let f^* denote the optimal specification value proposed by party B; then f^* must maximize $\Pi_B(f)$ subject to $f \in \Omega_A^t$. By Assumptions 1 and 2, f^* should be either f_A^H or f_A^L. If B proposes f_A^H, then that feature will be accepted by any type of A and so B's expected payoff is $\Pi_B(f_A^H)$. If B proposes f_A^L, then that feature will be accepted by a low-outside-option A but rejected by a high-outside-option A. Party B's expected payoff is then $q_0\Pi_B(f_A^L)+(1-q_0)u_B^t$. Therefore, party B of type t would propose f_A^L if and only if $q_0\Pi_B(f_A^L)+(1-q_0)u_B^t \geq \Pi_B(f_A^H)$, or iff $q_0 \geq \frac{\Pi_B(f_A^H)-u_B^t}{\Pi_B(f_A^L)-u_B^t}$.

Observe that $\frac{U_B(f_A^H)-u_B^t}{U_B(f_A^L)-u_B^t}$ is decreasing in u_B^t. We can therefore define two threshold variables that are critical for the analysis to follow:

$$q_0^H = \frac{\Pi_B(f_A^H)-u_B^H}{\Pi_B(f_A^L)-u_B^H}, \qquad q_0^L = \frac{\Pi_B(f_A^H)]-u_B^L}{\Pi_B(f_A^L)-u_B^L}.$$

It follows that if $q_0 < q_0^H$ then party B would propose f_A^H and if $q_0 > q_0^L$ then B would propose f_A^L. If $q_0^H < q_0 \leq q_0^L$, then party B proposes f_A^L if she has outside option u_B^H but proposes f_A^H if she has outside option u_B^L. □

PROPOSITION 4. *If f^* is the unique equilibrium product feature value in the game without explicit communication then, in any equilibrium of the game (with explicit communication) in which party B uses a pure strategy: (i) B requests f^* regardless of the received message; and (ii) the ex ante risk of cooperation failure is the same as in the game without explicit communication.*

Let nature choose A's type from the set $T=\{f_A^{0*}, f_A^{1*}, f_A^{2*}, \ldots, f_A^{n*}\}$, $n \gg 0$, according to a discrete prior distribution $h(\cdot)$; here $h(f_A^*)=\Pr(f_A^*=f_A^{i*})$ and $h(f_A^{i*})>0$ for all $f_A^{i*}\in T$. The elements of T satisfy $f_A^{i*}<f_A^{j*}$ for all $i<j$, where $i,j\in\mathbb{N}\equiv\{0,1,2,\ldots,n\}$.

In the game with explicit communication, party A has a message strategy that gives a probability distribution on all possible messages in M for each type in the set T. Let $p(m, f_A^{i*})$ be the probability that a party of type f_A^{i*} sends message m in a given equilibrium. Party B's proposal strategy now associates to each m a probability distribution over $[0,1]$. Let $f(m)$ be the proposal made by B after hearing m whenever B does not mix in response to that proposal. In a given equilibrium of a game with explicit communication, B will form posterior beliefs about A's type for each message m. Let these beliefs be denoted $h(f_A^{i*}, m)\equiv\Pr(f_A^*=f_A^{i*}\mid m)$, where $\mathrm{H}(f_A^{i*},m)\equiv\sum_{j=0}^{i-1}h(f_A^{j*},m)$. Let $\mathrm{H}(f_A^{i*})\equiv\sum_{j=0}^{i-1}h(f_A^{j*})$ denote the prior probability that A's optimal product feature is strictly less than f_A^{i*}. Let $\mathrm{H}(f_A^{0*})=0$.

To simplify the analysis we assume that, if party A is indifferent between rejecting and accepting a proposal f, then he will definitely accept it. The proof requires three lemmas as follows.

LEMMA 1. *In any equilibrium, B's payoff is at least $(2\sqrt{P-u_A}f_B^*-(f_B^*)^2+u_A)$, and in no equilibrium will B respond to any message with a proposal that is certain to result in cooperation failure.*

If B sets $f=\sqrt{P-u_A}$ then, in any equilibrium, all types of A will certainly accept (by subgame perfection); hence B can assure herself of $(2\sqrt{P-u_A}f_B^*-(f_B^*)^2+u_A)$ in this way. It is easily shown that $(2\sqrt{P-u_A}f_B^*-(f_B^*)^2+u_A)\geq u_B$ by Assumption 2. If in some equilibrium party B proposes f following a message m such that the project certainly fails, then B receives u_B; in that case, however, B could deviate to $f'=\sqrt{P-u_A}$ and do strictly better. □

LEMMA 2. *In any equilibrium, the request f is in the support of B's proposal strategy after receiving a message m only if there exists an $f_A^{i*} \in T$ such that $f = f_A^{i*} + \sqrt{P - u_A}$.*

If not, then B might (upon hearing m) propose f' such that $f_A^{i*} + \sqrt{P - u_A} < f' < f_A^{i+1*} + \sqrt{P - u_A}$. But in that case B could increase her payoff after hearing m by deviating to $f'' = f_A^{i+1*} + \sqrt{P - u_A}$, since doing so has no effect on the likelihood of cooperation failure. □

LEMMA 3. *Given any equilibrium in which B does not mix, for all $m \in F'$ we have $f(m) = K$ a constant, where F' is the set of messages sent with positive probability in the given equilibrium.*

Let $T_m \equiv \{f_A^{i*} : p(m, f_A^{i*}) > 0\}$. We argue by way of contradiction. So suppose that, in some equilibrium, there exist two distinct messages m and m' such that $f(m) < f(m')$. By Lemma 1, both proposals must be accepted with positive probability, which entails the existence of types $f_A^{i*} \in T_m$ such that $P - (f(m) - f_A^{i*})^2 \geq u_A$ and of types $f_A^{j*} \in T_{m'}$ such that $P - (f(m') - f_A^{j*})^2 \geq u_A$. Yet then any such party A of type $f_A^{j*} \in T_{B'}$ can do better by deviating to m, which gives him $P - (f(m) - f_A^{j*})^2 \geq P - (f(m) - f_A^{j*})^2$. However, this implies that $f(m')$ is certainly rejected—contradicting Lemma 1. □

Proof of Proposition 4. Suppose $f^* = \sqrt{P - u_A} + f_A^{k*}$ is a unique proposal in the game without explicit communication. Then f_A^{k*} is the only element of T such that, for all $j \in \mathbb{N}$,

$$\begin{aligned} &\mathrm{H}(f_A^{k*})u_B + [1 - \mathrm{H}(f_A^{k*})][P - (\sqrt{P - u_A} + f_A^{k*} - f_B^*)^2] \\ &\quad \geq \mathrm{H}(f_A^{j*})u_B + [1 - \mathrm{H}(f_A^{j*})][P - (\sqrt{P - u_A} + f_A^{j*} - f_B^*)^2]. \end{aligned} \tag{C.2}$$

Suppose to the contrary of the proposition that, in the game with explicit communication, there is some other proposal $f' = \sqrt{P - u_A} + f_A^{l*}$, $l \neq k$, such that B proposes f' after hearing any message $m \in F'$ (recall that, by Lemma 3, any equilibrium without mixing by B must have this form). Then, for each $m \in F'$ and for all $j \in \mathbb{N}$,

$$\begin{aligned} &\mathrm{H}(f_A^{l*}, m)u_B + [1 - \mathrm{H}(f_A^{l*}, m)][P - (\sqrt{P - u_A} + f_A^{l*} - f_B^*)^2] \\ &\quad \geq \mathrm{H}(f_A^{j*}, m)u_B + [1 - \mathrm{H}(f_A^{j*}, m)][P - (\sqrt{P - u_A} + f_A^{j*} - f_B^*)^2]. \end{aligned} \tag{C.3}$$

By Bayes' rule, $h(f_A^{j*}, m) = h(f_A^{j*})p(m, f_A^{j*})/\Pr(m)$ and $\mathrm{H}(f_A^j, m) = \frac{1}{\Pr(m)} \sum_{i=0}^{j-1} h(f_A^{i*}, m)$, where $\Pr(m) = \sum_{f_A^* \in T} h(f_A^*)p(m, f_A^*)$ is the probability that party A chooses message m in the equilibrium. After substituting into (C.3) and multiplying both sides by $\Pr(m)$, we obtain

$$\begin{aligned} &\left[\sum_{i=0}^{l-1} h(f_A^{i*})p(m, f_A^{i*})\right]u_B + \left[\Pr(m) - \sum_{i=0}^{l-1} h(f_A^{i*})p(m, f_A^{i*})\right][P - (\sqrt{P - u_A} + f_A^{l*} - f_B^*)^2] \\ &\quad \geq \left[\sum_{i=0}^{j-1} h(f_A^{j*})p(m, f_A^{i*})\right]u_B + \left[\Pr(m) - \sum_{i=0}^{j-1} h(f_A^{i*})p(m, f_A^{i*})\right][P - (\sqrt{P - u_A} + f_A^{j*} - f_B^*)^2] \end{aligned} \tag{C.4}$$

for all $j \in \mathbb{N}$.

Since (C.4) holds for each $m \in F'$, we can sum both sides over all $m \in F'$ and the inequalities still hold. Hence, for all $j \in \mathbb{N}$,

$$\left[\sum_{i=0}^{l-1} h(f_A^{i*}) \sum_{m \in F'} p(m, f_A^{i*})\right]u_B + \left[\sum_{m \in F'} \Pr(m) - \sum_{i=0}^{l-1} h(f_A^{i*}) \sum_{m \in F'} p(m, f_A^{i*})\right][P - (\sqrt{P - u_A} + f_A^{l*} - f_B^*)^2]$$

$$\geq \left[\sum_{i=0}^{j-1} h(f_A^{i*}) \sum_{m\in F'} p(m, f_A^{i*})\right] u_B + \left[\sum_{m\in F'} \Pr(m) - \sum_{i=0}^{j-1} h(f_A^{i*}) \sum_{m\in F'} p(m, f_A^{i*})\right][P - (\sqrt{P-u_A} + f_A^{j*} - f_B^*)^2]. \tag{C.5}$$

Since $\sum_{m\in F'} \Pr(m) = 1$ and $\sum_{m\in F'} p(m, f_A^{i*}) = 1$, it follows that (C.5) can be simplified to yield

$$\mathrm{H}(f_A^{l*})u_B + [1 - \mathrm{H}(f_A^{l*})][P - (\sqrt{P-u_A} + f_A^{kl*} - f_B^*)^2] \geq \mathrm{H}(f_A^{j*})u_B + [1 - \mathrm{H}(f_A^{j*})][P - (\sqrt{P-u_A} + f_A^{j*} - f_B^*)^2] \tag{C.6}$$

for all $j \in \mathbb{N}$. But this contradicts (C.2), the hypothesis that the game without explicit communication has a unique equilibrium: $f^* = \sqrt{P-u_A} + f_A^{k*}$, where $f_A^{k*} \neq f_A^{l*}$. This proves the first part of the proposition. To see the second part, observe that if B proposes f^* after any message sent with positive probability, then all types $f_A^{j*} < f^* - \sqrt{P-u_A}$ will choose to reject the proposal (resulting in cooperation failure) and all types $f_A^{j*} \geq f^* - \sqrt{P-u_A}$ will accept the proposal. Therefore, the ex ante likelihood of cooperation failure is simply $\mathrm{H}(f^* - \sqrt{P-u_A})$, just as in the game without communication. □

PROPOSITION 5. *For any belief that B may have about q_0, the only perfect Bayesian equilibrium of the bargaining game is as follows:*

Step 1: $f^1 = f_A^H$.

Step 2: *If $q_0 < q_0^t$, then B_t accepts and thus $f^2 = f^1 = f_A^H$;*

if $q_0 \geq q_0^t$, then B_t proposes $f^2 = f_A^L$.

Step 3: *If $f^2 = f^1$, then f_A^H is the outcome;*

if $f^2 = f_A^L$, then A_L (party A with low outside option) accepts and A_H rejects.

In order to prove Propositions 5, 6, and 7, we define the following parameters.

- $f^1(p_0) \in \mathbb{R}$: party A's initial specification proposal.
- $a^t(f^1, q_1)$: party B's response to party A's initial proposal; $a^t =$ accept when B accepts f^1 and $a^t = f^2(f^1, q_1) \in \mathbb{R}$ when B rejects and counter-proposes a new value.
- $s^t(f^2, p_1) \in \{0, 1\}$: party A's strategy when party B proposes f^2; $s^t = 1$ indicates "accept", and $s^t = 0$ indicates "reject" (in which case cooperation terminates).
- After observing A's initial proposal, B's updated belief concerning A's type is q_1, which is nonnegative and not greater than 1.

The proof of Proposition 5 requires the following lemma.

LEMMA 4. *If $q_0 \leq q_0^t$ then $f_t^2 = f_A^H$.*

The result immediately follows because, by subgame perfection, $f_t^2 = f_A^H$ when $q_0\Pi_m(f_A^L) + (1-q_0)u_B^t \leq \Pi_B(f_A^H)$. □

Proof of Proposition5. We begin by checking the equilibria under three cases for q_0.

Case 1: *If $q_0 \leq q_0^H$, then the following strategies and belief constitute a perfect Bayesian equilibrium.*

$$f_t^{1*} = f_A^H; \qquad s_t^* = \begin{cases} 1 & \text{if } f^2 \leq f_A^t, \\ 0 & \text{otherwise}; \end{cases} \qquad q_1^* = \begin{cases} q_0 & \text{if } f^1 = f_A^H, \\ 1 & \text{otherwise}; \end{cases}$$

$$a_t^* = \begin{cases} \text{accept} & \text{if } f^1 = f_A^H \text{ or } f^1 \leq f_A^L, \\ \text{propose } f^2 = f_A^L & \text{otherwise}. \end{cases}$$

Proving this case for s_t^* is straightforward. Hence we proceed by checking party B's best response. By Lemma 4, we must have $f_H^2 = f_A^H$ if $q_0 \leq \frac{\Pi_B(f_A^H)-u_B^H}{\Pi_B(f_A^L)-u_B^H} = q_0^H$. Since $f_t^{1*} = f_A^H$ and $\Pi_B(f_A^H) > \Pi_B(f_A^L)$, it follows that a party B of either type will prefer to accept party A's proposal of f_A^L. Following an out-of-equilibrium proposal from party A, the belief $q_1 = 1$ makes $f^2 = f_A^L$ the best response.

We now check to ensure that A will not deviate from proposing $f^{1*} = f_A^H$. A party A of either type will receive $\Pi_A(f_A^H)$ in equilibrium because f_A^H is accepted by a party B of any type. Because A prefers f_A^H to any other proposal that is always accepted, we need only consider deviations that induce B to make a proposal leading to cooperation failure. Given $f_B^* \notin \Omega_A^H$ and B's equilibrium strategy a_t^*, a deviation that triggers cooperation failure yields $\Pi_A(f_A^t)$ for each type, which lowers A's expected utility because $\Pi_A(f_A^H) > \Pi_A(f_A^L)$.

Beliefs regarding whether the equilibrium $q_1^* = q_0$ are determined by Bayes' rule. □

Case 2: *If* $q_0^H < q_0 \leq q_0^L$, *then the following strategies and belief constitute a perfect Bayesian equilibrium.*

$$f_t^{1*} = f_A^H; \qquad s_t^* = \begin{cases} 1 & \text{if } f^2 \leq f_A^t, \\ 0 & \text{otherwise}; \end{cases} \qquad q_1^* = \begin{cases} q_0 & \text{if } f^1 = f_A^H, \\ 1 & \text{otherwise}; \end{cases}$$

$$a_L^* = \begin{cases} \text{accept} & \text{if } f^1 = f_A^H \text{ or } f^1 \geq f_A^L, \\ \text{propose } f^2 = f_A^L & \text{otherwise}; \end{cases} \qquad a_H^* = \begin{cases} \text{accept} & \text{if } f^1 \geq f_A^L, \\ \text{propose } f^2 = f_A^L & \text{otherwise.} \end{cases}$$

First we consider party B's strategy. Given $q_1^* = 1$ following an out-of-equilibrium proposal from party A, f_A^L must be the best response. Hence we focus on B's response to f_A^H. By Lemma 4, $f_L^2 = f_A^H$ when $q_0 \leq \frac{\Pi_B(f_A^H)-u_B^L}{\Pi_B(f_A^L)-u_B^L} = q_0^L$. Therefore, in this equilibrium a low-outside-option B weakly prefers f_A^H to enforcing her own proposal. High-outside-option B parties prefer proposing f_A^L over f_A^H if $q_0\Pi_B(f_A^L)+(1-q_0)u_B^H \geq \Pi_B(f_A^H)$. This implies that $q_0 \geq \frac{\Pi_B(f_A^H)-u_B^H}{\Pi_B(f_A^L)-u_B^H} = q_0^H$, which is guaranteed.

Next we consider party A's proposal f^1. Given that in equilibrium a high-type B always proposes $f^2 = f_A^L$, the expected utilities of low- and high-type A parties are equal to (respectively) $p_0\Pi_A(f_A^H)+(1-p_0)\Pi_A(f_A^L)$ and $p_0\Pi_A(f_A^H)+(1-p_0)u_A^H$. There are two potential deviations that we must consider: party A may wish to propose $f^1 \neq f_A^H < f_A^L$, which would induce an counter-proposal from a high-type B (and then cooperation fails); or A may wish to propose $f^1 \geq f_A^L$, which is always accepted.

Consider the first potential deviation ($f^1 \neq f_A^H < f_A^L$). Given B's best response $f^2 = f_A^L$ and s_t^*, this deviation generates the payoff $\Pi_A(f_A^L)$ for a low-type A who is then clearly worse off; for a high-type A, there is no change in the payoff. So in this case, neither type can benefit from invoking party B to impose take-it-or-leave-it proposal.

Now suppose that party A deviates by proposing $f^1 \geq f_A^L$. Then, for an A of either type, the deviation generates $\Pi_A(f^1) \leq \Pi_A(f_A^L)$ because B always accepts this proposal. Yet by definition of f_A^t, no such proposal is profitable for an A party of either type.

Beliefs regarding whether the equilibrium $q_1^* = q_0$ are determined by Bayes' rule. □

Case 3: *If* $q_0 > q_0^L$, *then the following strategies and belief constitute a perfect Bayesian equilibrium.*

$$f_t^{1*} = f_A^H; \qquad s_t^* = \begin{cases} 1 & \text{if } f^2 \leq f_A^t, \\ 0 & \text{otherwise}; \end{cases} \qquad q_1^* = \begin{cases} q_0 & \text{if } f^1 = f_A^H, \\ 1 & \text{otherwise}; \end{cases}$$

$$a_t^* = \begin{cases} accept & \text{if } f^1 \geq f_A^L, \\ propose\ f^2 = f_A^L & otherwise. \end{cases}$$

The proof is identical to that for Cases 1 and 2 except that $q_0 > q_0^L$ leads to "accept" only if $f^1 \geq f_A^L$ is optimal for both types of party B. □

This completes the proof of Proposition 5. □

PROPOSITION 6. *For $e > 0$, define a new threshold probability of party B_t with outside option u_B^t as*

$$\hat{q}_0^t = \frac{\Pi_B(f_A^H) - u_B^t + e}{\Pi_B(f_A^L) - u_B^t}. \tag{C.7}$$

If the ultimatum cost e is sufficiently small (i.e., if $\hat{f}_A^L \geq f_A^H$) then, for any belief that B may have about q_0, the only perfect Bayesian equilibrium of the bargaining game described in Section 6.1 for weak asymmetry is as follows.

Step 1: $f^1 = z \in [z_t, f_A^H)$.

Step 2: *If $q_0 < \hat{q}_0^t$, then B_t accepts and thus $f^2 = f^1 = z$;*

if $q_0 \geq \hat{q}_0^t$, then B_t proposes $f^2 = f_A^L$.

Step 3: *If $f^2 = f^1$, then z is the outcome;*

if $f^2 = f_A^L$, then A_L (party A with low outside option) accepts and A_H rejects.

Define $z_t = \arg\max \Pi_A(f)$ such that $\Pi_B(f) \geq \max\{\Pi_B(f_A^H), q_0\Pi_B(f_A^L) + (1-q_0)u_B^t\} - e$. We make two observations concerning z_t. First, z_t is a feature value that (if accepted by B) gives B at least the same utility as she would receive from making her own optimal design decision. Second, z_t optimizes A's utility given a specification set that is acceptable to B. If A proposes any $f^1 < z_t$, then B_t will reject and enforce an f^2 that is worse from A's perspective. Therefore, z_t is A's best proposal.

The rest of the proof is based on the following lemma.

LEMMA 5. *We have $\Pi_A(z_t) \geq \Pi_A(f_A^H)$ whenever $q_1 \leq \frac{\Pi_B(f_A^H) - u_B^t + e}{\Pi_B(f_A^L) - u_B^t}$.*

Note that (i) $\Pi_A(f_A^H) = \max \Pi_A(f)$ such that $\Pi_B(f) \geq \Pi_B(f_A^H)$ and (ii) $\Pi_A(z_t) = \max \Pi_A(f)$ such that $\Pi_B(f) \geq \max\{\Pi_B(f_A^H), q_1\Pi_B(f_A^L) + (1-q_1)u_B^t\} - e$. By principles of constrained maximization, $\Pi_A(z_t) \geq \Pi_A(f_A^H)$ if and only if $\Pi_B(f_A^H) \geq q_1\Pi_B(f_A^L) + (1-q_1)u_B^t - e$—that is, iff $q_1 \leq \frac{\Pi_B(f_A^H) - u_B^t + e}{\Pi_B(f_A^L) - u_B^t}$. □

We prove the first case here; proofs for the other cases are the same as in Proposition 5.

Let $z \in [z_H, f_A^H)$. If $q_0 \leq \hat{q}_0^H = \frac{\Pi_B(f_A^H) - u_B^H + e}{\Pi_B(f_A^L) - u_B^H}$ and $\hat{f}_A^L \geq f_A^H$, then the following strategies and belief constitute a perfect Bayesian equilibrium.

$$f_t^{1*} = z; \qquad s_t^* = \begin{cases} 1 & \text{if } f^2 \leq f_A^t, \\ 0 & otherwise; \end{cases} \qquad q_1^* = \begin{cases} q_0 & \text{if } f^1 = z, \\ 1 & otherwise; \end{cases}$$

$$a_t^* = \begin{cases} accept & \text{if } f^1 = z \text{ or } f^1 \geq \hat{f}_A^L, \\ propose\ f^2 = f_A^L & otherwise. \end{cases}$$

Proving this case for s_t^* is straightforward. The rest of the proof is identical to that for Case 1 in Proposition 5. Hence we need only ensure that the range $[z_t, f_A^H)$ is not empty. By Lemma 5, $z_H \leq f_A^H$ whenever $q_1 \leq \frac{\Pi_B(f_A^H) - u_B^H + e}{\Pi_B(f_A^L) - u_B^L}$.

The conditions of Proposition 6 are fulfilled under a focal equilibrium in which $q_0 = q_1$. □

PROPOSITION 7. *If the cost of action e is large enough that $\hat{f}_A^L < f_A^H$ then, for any belief that B may have about q_0, the only perfect Bayesian equilibrium of the bargaining game described in Section 6.1 for strong asymmetry is as follows.*

Step 1: A_t *(party A of type t) proposes* $f^1 = \hat{f}_A^t$).

Step 2: *If* $f^1 \geq \hat{f}_A^L$, *then B accepts and thus* $f^2 = f^1 = \hat{f}_A^L$;

if $f^1 = \hat{f}_A^H$, *then B accepts with probability ϕ, but proposes* $f^2 = f_A^H$ *with probability* $1-\phi$;

otherwise, B proposes $f^2 = f_A^L$.

Step 3: A_t *accepts* f^2.

The probability ϕ is defined as $\phi = \frac{\Pi_A(\hat{f}_A^L)-\Pi_A(f_A^H)}{\Pi_A(\hat{f}_A^H)-\Pi_A(f_A^H)}$.

For party A, any proposal $f^1 \neq \hat{f}_A^t$ will result in $\Pi_A(f_A^t)$, which cannot be a best response because $\Pi_A(\hat{f}_A^L) \geq \Pi_A(f_A^L)$ and ϕ is defined such that $\phi\Pi_A(\hat{f}_A^H) + (1-\phi)\Pi_A(f_A^H) \geq \Pi_A(f_A^H)$. Therefore, to prevent a low-type A party from mimicking a high-type A party, we must have $\Pi_A(\hat{f}_A^H) \geq \phi\Pi_A(\hat{f}_A^H) + (1-\phi)\Pi_A(f_A^H)$. Similarly, to prevent a high-type A from mimicking a low-type A, we must have $\Pi_A(\hat{f}_A^H) \leq \phi\Pi_A(\hat{f}_A^H) + (1-\phi)\Pi_A(f_A^H)$. The only ϕ that satisfies both of these inequalities is $\phi = \frac{\Pi_A(\hat{f}_A^L)-\Pi_A(f_A^H)}{\Pi_A(\hat{f}_A^H)-\Pi_A(f_A^H)}$.

Party B learns party A's type from his proposal f^1 and by the definition of $\hat{f}_A^t$ and f_A^t : $\Pi_B(\hat{f}_A^t) = \Pi_B(f_A^t) - e$. Hence B is indifferent between accepting $\hat{f}_A^t$ and proposing alternative f_A^t while incurring cost e. Therefore, a_t^* is a best response on the equilibrium path. For any other specification value off the equilibrium path, $q_1^* = 1$ implies that enforcing $f_t^2 = f_A^L$ for any such $f^1 > \hat{f}_A^L$ is a best response. It is likewise a best response to accept any $f^1 \leq \hat{f}_A^L$.

In order for there to be at least one feasible probability ϕ, it is necessary that $\Pi_A(\hat{f}_A^L) > \Pi_A(f_A^H)$ or $\hat{f}_A^L < f_A^H$. The necessary condition $\Pi_B(\hat{f}_A^t) = \Pi_B(f_A^t) - e$ and the uniqueness of $\hat{f}_A^t$ and f_A^t exclude any other separating equilibrium. □

PROPOSITION 8. *Increasing (resp., decreasing) the sensitivity ϕ of party B increases (resp., decreases) the risk of cooperation failure.*

Recall from the proof of Proposition 3 that q_0^H and q_0^L mark the respective thresholds for low and high risk of cooperation failure. The equilibria remain identical, and sensitivity analysis is performed by checking how q_0^H and q_0^L vary with ϕ.

Define $X = P - a(f_A^H - f_B^*)^2$ and $Y = P - a(f_A^L - f_B^*)^2$. Clearly, $Y > X$. It follows that

$$q_0^H = \frac{\phi X - u_B^H}{\phi Y - u_B^H} = 1 - \frac{Y - X}{Y - (u_B^H/\phi)}.$$

As ϕ increases, u_B^H/ϕ decreases and also $(Y-X)/(Y-(u_B^H/\phi))$ decreases. Hence q_0^H is increasing in ϕ and, as expected, q_0^L is also increasing in ϕ. □