The Cognitive Foundations of Tacit Commitments: A Virtual Bargaining Model of Dynamic Interactions

Tigran Melkonyan,[1] Hossam Zeitoun[2] & Nick Chater[2]

1: *Department of Economics, Finance and Legal Studies, University of Alabama, Tuscaloosa*

2: *Warwick Business School, University of Warwick*

Author Note

The authors were supported by the British Academy [grant SG170033]. NC was also supported by the ESRC Network for Integrated Behavioural Science [grant number ES/P008976/1], and the Leverhulme Trust [grant number RP2012-V-022]. We would like to thank two anonymous reviewers and the editor for insightful feedback on previous drafts, which has greatly strengthened the paper. Please address correspondence to Tigran Melkonyan, Department of Economics, Finance and Legal Studies, University of Alabama, 361 Stadium Dr Tuscaloosa, AL 35487, USA, tamelkonyan@cba.ua.edu.

The Cognitive Foundations of Tacit Commitments: A Virtual Bargaining Model of Dynamic Interactions

Abstract

People often make, and are held to account for, purely tacit commitments in interactions with other people: commitments that have never been explicitly articulated or agreed. Moreover, unspoken, tacit commitments are often perceived as binding: people often stick to, and are expected to stick to, these commitments, even where it might seem against their interests to do so. If they do not stick to these commitments, they may be punished, and expect to be punished, by others as a result, even if the act of punishment is itself costly for the punisher. These commitments have been widely seen as a crucial underpinning for human collaboration and cooperation. Yet how do such commitments arise, and are they compatible with human rationality? This paper provides a formal, reasoning-based account of tacit commitments based on "virtual bargaining"—a mode of reasoning that joins elements of individualistic and collaborative reasoning. We complement existing accounts by showing that even purely self-interested individuals can, under certain conditions, tacitly commit to punishing counterparts who violate an unenforceable agreement, or to cooperating in dynamic games, including the Centipede game and the finitely repeated Prisoner's Dilemma game.

1. INTRODUCTION

The human ability to make, and stick to, commitments is crucial to human social interactions and relationships, as well as underpinning many economic transactions (Clark, 2006; Michael, Sebanz, & Knoblich, 2016; Nesse, 2001; Tomasello, 2020). Human social interactions are a continual back-and-forth—and the commitments involved in such interactions typically have the same character. One person performs an action, anticipating that the other will respond in a particular way, who in turn may expect a further response from the first—and the entire sequence may be guided by tacit (or sometimes explicitly stated) commitments.

Consider, for example, the tacit commitments that underpin even the smallest and least consequential interactions, such as one person lending another a pen. When the pen is handed over, there is typically a tacit agreement about timeframe and purpose. If, at the bank kiosk, one person asks to borrow a pen from another, there is a tacit agreement that the pen is being borrowed just to sign a particular cheque or form, rather than for a day or a year; and both parties tacitly agree that, unless explicitly agreed otherwise, the borrower should not break the pen in two or otherwise damage it; that, if the pen is inadvertently lost, the borrower should apologize and perhaps offer an equivalent pen; that the borrower cannot normally loan the pen to a third person, and so on. Similarly, the lender cannot subsequently demand payment, unless explicitly agreed beforehand.

But unless people feel under some obligation to stick to the commitment (and to enforce the commitment if the other deviates), such interactions are likely to fail. Having borrowed the pen, the borrower may think it convenient not to give it back; and purely 'forward looking' rational agents will know that it is surely pointless to get involved in a serious confrontation with another person, potentially even escalating into physical violence, over so small a matter as a pen. This seems to imply that rational agents won't return pens if they would rather not; and, consequently, that no rational agent will lend a pen in the first place. Indeed, more broadly, routine social interactions of all kinds will break down unless people can be "trusted" to be guided by the obligation to follow commitment (Dunning & Fetchenhauer, 2011; Tomasello, 2020).

Indeed, the ability to make, enforce, and stick to commitments seems crucial in order to maintain coherence in social interactions of many kinds, from conversational exchanges, jointly cooking a meal, planning an event, bringing in each other's harvest, making a piece of flat-pack furniture or setting up a new business (Bangerter & Clark, 2003; Clark, 1996, 2006).

Recent decades have seen great interest in philosophy and in developmental and social psychology in the idea of joint action (Garrod & Pickering, 2009; Pesquita, Whitwell, & Enns, 2018; Sebanz, Bekkering, & Knoblich, 2006). Most joint actions are, of course, not instantaneous, but involve sequences of sub-actions contributed by the parties concerned. To sustain such a joint action, both parties need to be jointly committed, and to know that they are jointly committed, to playing their respective parts in the joint action. Indeed, successful social interactions seem to require a joint understanding between the parties that they are implicitly 'signed up to.' Thus, the person lending the pen expects the other to be committed to returning it promptly; the person receiving the pen is committed to doing so, and knows that the other knows this, and so on. And if the receiver of the pen were to transgress (e.g., by pocketing the pen, breaking it, putting it in their mouth, etc.), they would anticipate that the other would view such behavior as violating the commitment. The transgressor would expect reprimand (if their transgression were noticed); and the other would feel licensed and perhaps even socially required to provide it.

While fundamental to social life, the ability to form such tacit commitments over sequential interactions is, in many cases, puzzling from the point of view of a rational choice explanation of behavior, as embodied in standard game-theoretic notions that rely on the logic of backward induction. The apparent difficulty is that rational agents will choose their actions based on their likely future consequences, with no direct regard to what they have committed themselves to in the past, unless there is an explicit mechanism that mandates implementing a prior commitment. Thus, a rational agent will follow a commitment just when it is in her future best interests to do so, but not otherwise; and a commitment that one only follows when it is in one's interests to do so is, in many circumstances, scarcely a meaningful commitment at all (e.g., Jensen, 2009). Indeed, what makes commitments powerful in helping coordinate social and economic life is that, to some degree at least, we can rely on people to stick to their commitments even in circumstances in which they would rather not.[1]

[1] While ubiquitous in social interactions of all kinds, this phenomenon is most intensively studied in economics, under the heading of "hold-up" problems, and is a central to the theory of incomplete contracts and the theory of the firm (Hart, 1995). One approach to the problem in economic contexts can be explicit contracts (e.g., Rogerson, 1992) which are enforceable by the law, backed by the judicial system. Such mechanisms are rarely available in everyday social interactions, of course; and even if they were, the challenge of writing suitably "complete" contracts, and the potential costs of enforcement, would make this type of formal mechanism impractical.

This problem is amplified in dynamic games by what we call the "curse of backward induction," which causes commitments involved in mutually beneficial interactions to unravel, according to conventional rational accounts. As we have noted, if it is the case that once the pen has been lent, the borrower is not really committed to returning it (and has a self-interest in not doing so), then it seems to follow that the pen will not be lent in the first place. Similarly, if two farmers help each other bring in each other's harvests on alternating days, a purely self-interested farmer would stop once their own harvest finished and not help the other on the final day. Knowing this, the other farmer would plan to stop on the day before (because they would anticipate getting no help on the last day); and hence the first farmer would have no reason to help on the day before that, and so on. Backward induction will imply that cooperation will never begin, however mutually beneficial it might be for each farmer. We shall see later that the curse of backward induction looms large in classic experimental games, such as the Centipede game and the finitely repeated Prisoner's Dilemma. These games seem to show that rational agents will be unable to engage in mutually advantageous interactions. Yet experimental studies (Krockow, Colman, & Pulford, 2016; Krockow, Pulford, & Colman, 2018; McKelvey & Palfrey, 1992) have shown that people do engage in mutually beneficial collaboration to a considerable degree in such contexts. Are these successes due to a failure of human rationality? We will argue, instead, that a different theory of rational social interaction, based on virtual bargaining (Chater, Zeitoun, & Melkonyan, in press; Melkonyan, Zeitoun, & Chater, 2018; Misyak, Melkonyan, Zeitoun, & Chater, 2014), provides a reason-based explanation of why even purely self-interested agents need not be subject to the curse of backward induction, and can engage in sequences of mutually beneficial interactions.

A particularly important case where backward induction creates puzzles for standard rational accounts concerns threats of punishment. Suppose farmer *A* does indeed fail to turn up on the final day of harvest to help at *B*'s farm. *B* will surely be angry, and perhaps initiate conflict of some kind. But conflict will typically not merely be aversive for *A*, but also unpleasant for *B*. Indeed, were this not the case, a purely self-interested *B* might engage in conflict, whether *A* helped or not. But if conflict is unpleasant for *B*, then *A* will predict that, once harvest is over, *B* will not punish *A* (because this just makes *B*'s bad situation even worse). Thus, a tacit, or even explicit, threat by *B* appears to be an empty threat, if *B* is a purely self-interested rational agent. So, punishment fails to be a useful mechanism for enforcing behavior after all. Backward induction then leads to the conclusion that *A* will not turn up on the final day, or *B* on the day before, and so on, so that mutually beneficially alternation of help between *A* and *B* is blocked.[2]

There is, moreover, a further puzzling aspect of commitments in most human interactions: the fact that although commitments are sometimes made explicitly—in forms ranging from informal plans and promises, to legally binding contracts—the vast majority of commitments, such as those governing lending a pen, are largely implicit: the agreement is somehow "understood" by both parties, yet never formulated in words. So, for example, a passing customer may hand over some coins to a news vendor and pick up the morning paper without a word being exchanged; a passenger may state an address when getting into a cab, but with no discussion of a tacit commitment to pay the driver on arrival, or that the driver should take the quickest route and not, for example, stop for a lunch break in the middle of the journey. Similarly, people buy rounds of drinks without seeking, or receiving, any explicit commitment from another person to buy the next round; families share food (e.g., a freshly made cake or a box of chocolates) under tacit commitments that one person is not allowed to eat the lot when the others backs are turned (although such commitments are sometimes broken, of course). A social world without reliable commitments would be a hostile one indeed—people would seem continually pitched as potential adversaries, each following their best interests. The degree to which we can expect others reliably to adhere to their typically tacit commitments is closely related to the level of trust between people, which has been viewed as a central determinant of individual and societal well-being (e.g., Fukuyama, 1996; O'Neill, 2002). But, from a rational choice perspective, it is puzzling that trust in commitments—especially tacit commitments—is so ubiquitous.

Why, then, do rational agents make and stick to commitments, and especially, tacit commitments? What are the cognitive mechanisms from which they arise? Existing research in the rational choice tradition offers different answers to these questions. It may be assumed that

---

[2] Reputational factors may, of course, give people an incentive to maintain their agreements—if farmer *A* develops a reputation for violating agreements, then it may reasonably be expected to lead to other people not entering into mutually beneficial agreements with *A* in the future. But in cases where people's interactions are not directly observed by others, this mechanism does not apply for purely self-interested rational agents interested in maximizing their payoffs, because *A* and *B* cannot be relied upon to reliably report *A*'s behavior (as their reports, as well as all their behavior, will be chosen to optimize their interests, rather than to follow a tacit agreement to tell the truth). This is an aspect of the problem of so-called "cheap talk" in economics (Farrell & Rabin, 1996) and animal behavior (Silk, Kaldor, & Boyd, 2000).

individuals have non-standard preferences: e.g., they may be of a certain "cooperative type" (e.g., Kreps, Milgrom, Roberts, & Wilson, 1982) or have altruistic preferences in rewarding cooperative, and punishing uncooperative, behavior by their counterparts (e.g., Fehr & Gächter, 2002). Another strand of rational explanation focuses on how people can utilize external commitment devices (e.g., burning bridges behind them) to modify the costs of implementing commitments (e.g., Benabou & Tirole, 2004), so that it becomes in their best interests to follow the commitment once it is made. While these commitment devices undoubtedly arise in many contexts, this type of analysis does not resolve the fundamental puzzle that people often *do* follow commitments, crucially including entirely tacit commitments, when it would appear to be in their best interests not to do so.

One reaction to these difficulties is to suggest that behavior in making and honoring commitments must arise from something other than a rational attempt to obtain good consequences from an interaction. For example, people might follow their commitments because they feel this is "morally right" (Malle, 2021) or socially "appropriate" (March & Olsen, 2008), or even from habits accrued through a history of reinforcement (e.g., Skinner, 1971). But we suggest that, in some circumstances at least, mutually beneficial interactions between people can be justified in terms of reasons.[3] Indeed, people who engage in such collaborations can still gain enormously (even if a farmer is occasionally cheated on the very last day of harvest). Perhaps, then, we need a shift in our conception of what counts as a rational approach to social interaction—one that avoids the curse of backward induction, and that provides a rational justification for mutually beneficial interactions that involve making, and in large part sticking to, commitments.

In this paper, we outline an explanation for the human ability to create, and follow, tacit commitments in social and economic interactions—one that arises from the individual's reasoning process itself. We argue that people are tacitly committed to some course of action when it is clear that this commitment *would* have been agreed, had a process of actual discussion or negotiation occurred. This requires that each party knows the nature of the agreement—i.e.,

[3] In rational choice theory, it is typically assumed that choosing in one's best interest (allowing that one's own interest may incorporate concern for the well-being of others) is a basic tenet of rationality. Reason-based explanation in philosophy and psychology is much broader than this specific notion of rationality, however—reasons may include procedural justice, prior promises, expectations, norms of politeness, or merely the dictates of tradition (e.g., Sen, 1977, 2009).

who is part of the agreement; what each of them and the others are tacitly committed to; and, moreover, that the agreement is common knowledge between them (in the usage in game-theory and epistemic logic: knowing the agreement, knowing that the others know it, knowing that the others know that they know it, and so on).

Yet if a tacit agreement is never formulated and discussed explicitly, how can this common knowledge of what is agreed arise? Our starting point is that tacit commitments arise from a process of "virtual bargaining"—a mentally simulated process of negotiation about what we *would* agree, were we to bargain explicitly (Chater, Misyak, Melkonyan, & Zeitoun, 2016; Chater et al., in press; Melkonyan et al., 2018; Misyak et al., 2014). Often, we will see, there is a "natural" bargain that would be agreed upon—and this natural agreement can be inferred by all players. Through simulating the bargaining process, all parties can infer, and can infer that others infer, the tacit commitment to which they are collectively "signed up" for.

A crucial question, of course, is what determines the virtual bargain. A simple, but unsatisfactory, approach, would be for both parties to imagine that they could agree on a binding contract, as if it were legally enforceable. This approach would make cooperation easy to explain—for example, the farmers in our example above would simply mentally simulate which written and enforceable contract they would sign up to—and this would presumably be that each would agree to help the other throughout the entire harvest. Having made this mental simulation, they would then behave accordingly. But this approach is inadequate. It would, for example, be equally possible to make agreements in which *A* helps *B* for the first half of the harvest, and then *B* helps *A* for the second half. But such an agreement puts *A* in real jeopardy, because, in reality, there is no legally binding and enforceable agreement, and *B* might not return *A*'s help. Indeed, if people were guided by imagined enforceable agreements, they would continually be in danger of being badly exploited by selfish counterparts.

The theory of virtual bargaining therefore takes a different approach. It assumes that, when considering a possible virtual bargain, each person considers not only the possibility that the bargain is adhered to, but that their counterpart may unilaterally depart from the bargain, to maximize their own selfish interests. Thus, the farmers would consider their own payoff in the scenario in which each helps the other as promised; but also consider the possibility that the other might exploit them to maximize their own gain. Now, the virtues of helping each other on alternating days becomes apparent—for one farmer to exploit the other, this farmer will withdraw her labor only on the last day. While this outcome disadvantages the other farmer, the overall benefit of their collaboration through the preceding days may still make the collaboration worthwhile for both. So, this is an agreement that both parties might sign up to—because even if one attempts to be exploitative, the results for both are still pretty good: the bargain is fairly "exploitation-proof." But if *A* helps *B* for the first half of the harvest, and *B* turns out to be exploitative, then this will be a very bad outcome for *A*. For this reason, *A* would never "sign up" to such an agreement, because it allows for considerable exploitation. The theory of virtual bargaining assumes that people choose which bargain to agree by taking account of both the possibility that the bargain will be followed, but also the possibility that the other may unilaterally depart from the bargain to maximize her own interests. Indeed, the virtual bargaining account assumes that agents are what we will call "cautious" regarding such outcomes—each evaluates the bargain based on whichever case has the worst outcome for themselves.[4]

The theory of virtual bargaining has hitherto been developed primarily for static interactions, in which the parties simultaneously and independently choose a single action. But human interactions, and the tacit commitments that guide them, typically involve sequences of actions unfolding over time—e.g., a pen is requested, lent, not returned, the borrower complains, and so on. One key contribution of this paper is to extend the theory of virtual bargaining to capture the simulated process of forming tacit commitments in this type of "dynamic" setting involving an unfolding, and potentially interacting, sequence of actions from each player.

We shall see that, crucially, the dynamic virtual bargaining analysis may turn seemingly empty threats into credible ones. Relatedly, it may explain cooperative behavior in finitely repeated interactions, but avoid the curse of backward induction, as we saw above with the example of the two farmers. We will see that mutually beneficial cooperation in such finitely

[4] We note that, when contemplating adopting a virtual bargain, it makes sense to wonder whether the other will follow the bargain or unilaterally depart from it to maximize her own interests. But a farmer does not need to worry about where *she herself* will follow the bargain—she can, of course, trust herself to follow whatever hypothetical course of action she thinks best. By contrast, standard game-theoretic reasoning implies that a person cannot trust themselves to follow through with any course of action. Indeed, quite the opposite—it requires that they follow whatever course of action maximizes their own interests, irrespective of any prior commitments, tacit or otherwise. We note, too, that our assumption of "caution" may seem excessively pessimistic, through concentrating on the worst of the two possible scenarios. A more general account that assigns probabilities to each possibility can be imagined, but is left for future work.

repeated interactions can be sustained through both parties forming a tacit agreement and embracing collaborative reasoning while still caring about their own goals.

We extend existing rational choice theory by proposing that tacit commitments can emerge even with purely self-interested preferences, based on the individuals' mode of reasoning. In this paper, to explore the implications of virtual bargaining as a mode of reasoning in dynamic interactions, we will make the simplest possible assumptions otherwise: that our agents are entirely self-interested, have no concerns about reputation, and so on. It is, though, quite possible to introduce these other factors, which aim to increase the behavioral realism of rational choice accounts, alongside the virtual bargaining approach. Our approach is intended to be complementary to, rather than in competition with, accounts that focus on these other factors.

We propose that tacit commitments often arise from a process of mentally simulated, or "virtual," bargaining,[5] which do not emerge from an explicit process of offer and counteroffer, and indeed are not formulated explicitly at all. Explicit processes of bargaining and codification of agreements can be costly in both time and money; but often they appear to be unnecessary. Transacting parties can come to a tacit agreement governing their interaction without communication. They can, moreover, construct tacit agreements in a highly flexible way to deal with specific unexpected situations that they may face—and that might not have been anticipated in an explicit contract. Although many disciplines in the social sciences emphasize the importance of general tacit conventions, customs, and norms of all kinds (e.g., Cialdini, 2007), there has been less attention to the question of how these are applied to create agreements in specific contexts.[6]

---

[5] Here, tacit agreements are closed connected with the notion of "pacts" in psycholinguistics (Brennan & Clark, 1996). The notion is also related to, but not identical with, notions of the psychological contract, especially between employees and employers (Argyris, 1960; Rousseau, 1995; Rousseau & Shperling, 2003)—see Chater, Zeitoun & Melkonyan (in press) for discussion. Note that tacit agreements differ also from "implicit contracts" discussed in economics (e.g., Baker, Gibbons, & Murphy, 2002): tacit agreements can arise spontaneously even in one-shot interactions, whereas implicit contracts typically rely on repeated interactions. The two notions are not mutually exclusive. Moreover, our theory is related to existing literature on "internal commitments" or self-regulation mechanisms such as prudent saving behavior, regular exercising and keeping a healthy diet (e.g., Benabou & Tirole, 2004). Benabou & Tirole (2004) build a theory of internal commitments, which builds on Ainslie (1992, 2001), based on imperfect recall of past feelings and motives. However, rather than examining commitments relating to an individual's self-regulation in non-strategic situations, we focus on tacit (internal) commitments in interactions between individuals.

[6] This distinction between general and specific (tacit) agreements is analogous to the question in the legal literature of how general (written) laws underpin specific (written) contracts governing transactions. For

The human tendency to create and follow tacit agreements has deep psychological foundations. Where people explicitly pre-commit to an action, or are tacitly taken to be committed to it, they appear to have a psychological pull to go through with their commitment (and they object when others do not)—a tendency that humans develop from a very young age. To illustrate, Rakoczy, Warneken, and Tomasello (2008) showed two- and three-year-old children examples of adults engaging in a novel action labeled "daxing." The children strongly objected when a hand-puppet announced that it would dax, but then produced a different action. Crucially, young children would also protest against the violation of tacit rules without prior agreement—e.g., if they observed an agent using an object in a location where they had inferred by observation that this location was not permitted (Rakoczy, Brosche, Warneken, & Tomasello, 2009). Highlighting the importance of tacit agreements even further, some scholars have argued that many of our "moral" emotions focus primarily on commitments to (or violations of) explicit or tacit rules of behavior, rather than solely the outcomes of people's actions (Nichols, 2004). Moreover, the sense of obligation generated by tacit commitments appears fundamental to the rich collaborative behavior observed in human societies (Tomasello, 2020).

To analyze the reasoning process about tacit agreements, we develop a general formal theory of tacit commitment through virtual bargaining (Chater et al., 2016; Chater et al., in press; Melkonyan et al., 2018; Misyak et al., 2014) by developing new equilibrium notions for virtual bargaining in dynamic games. We formally analyze these solution concepts in four games representing broad classes of interactions: the *Gas Station game* (which is an abstract representation of the ubiquitous real-world situations where two individuals sequentially exchange goods for money without being observed by a third party), the *Repeated Bertrand Competition game*, the *Centipede game*, and the *Finitely Repeated Prisoner's Dilemma*. These games approximate many real-life interactions, but they are stylized representations, where talk of prisoners and gas stations is really a "cover story" to explain the structure of the game.[7] In each game, we compare the virtual bargaining solution concepts with each other and with the

---

example, consumer protection laws allow for a range of ways in which a bank can lend money. Within these constraints, the bank and its customer will agree on the type of contract (e.g., a mortgage or a payday loan) and the terms and conditions depending on the nature of their specific transaction. Specific tacit agreements are our main focus here.

[7] Thus, these games are not intended to capture the fine details of actual plea-bargaining challenges faced by prisoners or the specific nature of interactions at gas stations.

subgame perfect equilibria. We find that, under certain conditions, virtual bargaining enables even self-interested individuals to tacitly commit, and stick, to a course of action that deviates from, and may lead to better outcomes than, standard predictions.

To give a taste of our equilibrium notions in dynamic interactions, consider an abstract version of the farming example we gave above: a variant of the well-known Centipede game (Aumann, 1988; Rosenthal, 1981), which provides a classic puzzle for rational choice analysis. In this game, the players move sequentially and, whenever it is a player's turn to move, she can either end the game or continue, with the latter action resulting in a slight loss to themselves and a considerable gain to the other—with them being limited to some known, finite, number of turns. If the players can somehow contrive to continue the game to the end, then each will receive many large gains and suffer much smaller losses—so continuing the game is mutually highly beneficial. But these benefits appear difficult to reconcile with standard models of rationality. The standard game-theoretic analysis assumes that rational players will play a so-called "subgame perfect equilibrium" which, by the logic of backward induction, predicts that the game ends on the very first move, so that the potential mutual benefits of playing the game are lost to both players. This is unsatisfactory not just normatively, in that rational agents appear to forgo a considerable opportunity for mutual benefit; but also in light of a significant body of experimental evidence (see, e.g., Krockow et al., 2016; Krockow et al., 2018; McKelvey & Palfrey, 1992), showing that experimental subjects often choose to continue the game for a relatively long fraction of the strategic interaction and that a non-negligible number of players continue to almost the very end.

We argue that the players contemplate a tacit agreement to continue the game to the very end of the strategic interaction. Each player realizes that her opponent may deviate from this tacit agreement. At the same time, each player realizes that, even if such deviation were to occur, it would take place in the very end, and this course of action would be preferable to alternative tacit agreements. In other words, each player asks the question, "What is the worst thing that can happen to me if I follow the tacit agreement, and, suspecting this, my opponent deviates from the agreement to maximize her own payoff on the assumption that I stick to it (i.e., what game-theorists call "best responding")?"[8] and finds the answer, "The opponent will continue to

[8] A "best-response" to an opponent's strategy is a strategy that has the highest expected payoff, given the opponents' strategies. The concept of best response is central in game theory—indeed, a Nash equilibrium

"almost" the very end by deviating only during her last play," to be rather re-assuring. For both players to embrace this type of reasoning, it must be the case that they both think as virtual bargainers seeking tacit agreements. Thus, in addition to individualistic reasoning, the two players must have their "collaborative hats" on. It seems that the very nature of the Centipede game (as well as many other dynamic games, where there are significant advantages from coordinated actions) will lead many players to embrace such reasoning. As we shall discuss later, however, some individuals may not embrace this type of reasoning, and therefore may require non-standard preferences to achieve cooperative outcomes, or they may not achieve cooperative outcomes at all.

2. DEFINITIONS OF VIRTUAL BARGAINING

We outline the intuition behind the virtual bargaining approach, leaving the formal definitions to the Appendix. We begin by summarizing our existing analysis examining games in which strategies are chosen simultaneously and independently (Chater et al., 2016; Melkonyan et al., 2018; Misyak et al., 2014)—this is known as games in 'normal form' in game theory.

The core intuition is that players consider possible bargains[9] that they might reach: this corresponds to a set of strategies. When considering whether a bargain is acceptable, each player, *i*, considers two possibilities: (i) that all players follow the bargain; and (ii) that while *i* herself follows the bargain, all of the others choose strategies that are individual best-responses to the bargain. The idea is that when agreeing to a bargain, we are concerned with what happens if the bargain is implemented, and with what happens if we are "betrayed" or "exploited". In evaluating the attractiveness of a bargain, we do not have to consider whether we ourselves will follow the terms of the bargain; our concern is only that others may not do so. We assume that each player's payoff does not directly depend on the payoffs of other players but it may depend on the latter indirectly through strategic reasoning and its effect on choices. We also suppose

is defined as a set of strategies such that each strategy is a best response to the others. A strategy in a dynamic extensive form game specifies a player's choice in *every* contingency where it is that player's turn to move. Thus, when determining her best response, a player considers choices in each possible contingency – all possible paths are taken into account, given a fixed strategy of the opponent.

[9] We use bargain, agreement and strategy profile interchangeably.

that players are cautious, in that they evaluate each bargain according to what we shall call the *worst payoff* (i.e., the minimum of their outcomes in cases (i) and (ii)). Note that each bargain has its distinct worst payoff for each player in the game. A player's worst payoff for a bargain depends on the bargain's prescription for the player under consideration and the opponents' behavior (the prescription of the bargain for the opponents and the opponents' individual best responses).

We stipulate that a strategy profile is a *feasible agreement* if no player can improve her worst payoff by unilaterally modifying her own strategy and playing any other strategy (a formal discussion is introduced below). This is analogous to the familiar notion of Nash equilibrium, but applied to the modified game where payoff for each strategy profile is replaced by the worst payoff for that strategy profile. Feasible agreements seem a natural constraint on what people might tacitly agree: if each player attempts to maximize their worst payoff, then they will not choose an agreement where, by a unilateral change, they can find a better agreement in terms of its worst payoff. Note that all Nash equilibria are feasible agreements; but, crucially, there can be feasible agreements that are not Nash equilibria.

Many games have multiple feasible agreements. To select a specific feasible agreement, a process of mentally simulated bargaining is proposed. The chosen feasible agreement(s) is the one that the players would agree upon if they were able to bargain explicitly (if the outcome of this mentally simulated bargaining is unclear to the players, they will not reliably be able to form a tacit agreement). The process by which players "simulate" the bargaining process could, in principle, be modelled by any theory of explicit bargaining (e.g., Kalai, 1977). For concreteness and simplicity, we assume that the choice is made by simulating the results of a particularly simple and well-known model of bargaining known as Nash bargaining (Nash, 1950), which maximizes the product of utility gains from the putative bargain over and above what will be gained if there is *disagreement*: i.e., no bargain is reached (see expression (5) in Appendix). The Nash bargaining solution represents an outcome of a bargaining process where each player gets her disagreement payoff plus a share of the benefits from reaching an agreement. We call the chosen agreement the Virtual Bargaining Equilibrium (VBE).

To illustrate, consider a game that has three feasible agreements (x_1, x_2), (y_1, y_2) and (z_1, z_2), where the variables with index $i = 1,2$ denote the strategies of player i. Suppose that both players get a worst payoff of 2 under the feasible agreement (x_1, x_2). Assume also that player 1's worst payoff under (y_1, y_2) is 7 while player 2's is 9. Under agreement (z_1, z_2), both players' worst payoff is equal to 8. Under this parameterization, the two players' disagreement positions are characterized by the worst payoff of 2, which is the minimum worst payoff in a feasible agreement. The product of utility gains from bargain (x_1, x_2) over the players' disagreement payoffs is equal to $(2-2)\cdot(2-2)=0$. Similarly, the product of utility gains from bargain (y_1, y_2) over the players' disagreement payoffs is equal to $(7-2)\cdot(9-2)=35$. And finally, the product of utility gains from bargain (z_1, z_2) over the players' disagreement payoffs is equal to $(8-2)\cdot(8-2)=36$. The agreement (z_1, z_2) is the VBE since it achieves the largest product of utility gains.

For many generic games, the set of feasible agreements is different from the set of Nash equilibria and the selection from the set of feasible agreements (via Nash bargaining) includes agreements that are better *for all parties*, than any Nash equilibrium (i.e., using Economics jargon, the set of feasible agreements Pareto "dominates" all Nash equilibria). This fact, together with the formal definition provided in Appendix, illustrates that virtual bargaining equilibrium is different from existing equilibrium notions that refine the set of Nash equilibria by invoking Pareto efficiency, coalition-proofness, or some other criterion (see, e.g., Bernheim, Peleg, & Whinston, 1987).

We now turn to examining virtual agreements in dynamic games, where players interact over a sequence of moves. In game theory, such dynamic interactions are often represented as what are called extensive-form games: an explicit representation of the sequencing of players' choices, and of the information of each player at her every decision point about the other players' prior moves (for technical background see, for example, Binmore, 2007). As before, suppose that the game has n players.[10] We will say that a strategy profile is a *virtual bargaining equilibrium of an extensive-form game* Γ if that strategy profile is a VBE of the normal-form game corresponding to Γ.[11] Thus, if a strategy profile is a VBE of an extensive-form game, the players can virtually agree to a course of action under each possible contingency (i.e., in technical game-theoretic

[10] We assume that there are no external factors that also affect the outcome of the game—in the language of game theory, we assume that the game has no moves by Nature.

[11] Thus, we define the VBE of an extensive-form game similarly to how a Nash equilibrium of an extensive-form game is defined.

terms, this means for each information set[12]) and even for contingencies that may be ruled out by the player's earlier choices. In other words, even if a player deviates from the VBE at some point in the game, her opponents may stick to their part of the virtual agreement. Thus, the players have an ability to virtually pre-commit to follow a course of action (at each of her information sets) virtually agreed upon in the very beginning of the strategic interaction. In what follows, we will refer to "VBE of an extensive-form game" as simply "VBE."

Consider now the scenario where the players cannot virtually pre-commit to an agreement governing the entire course of their strategic interaction. Rather, at each point in the game that starts a subgame[13], they can virtually renegotiate the strategy profile that they have previously agreed upon. Formally, we will say that strategy profile σ^V is a *subgame virtual bargaining equilibrium* (SVBE)[14] if for every subgame the strategy σ^V restricted to the subgame constitutes a virtual bargaining equilibrium for the subgame. That is, when one considers any subgame as a game in its own, the prescription of the overall agreement for that subgame constitutes a VBE. Roughly speaking, starting from any point in the game, the players do not have a strict incentive to renegotiate their initial virtual bargain.

Suppose that players do not possess private information, move sequentially, and learn all of the moves made in the game immediately after these moves are made. Consider the very last choice nodes (where a choice leads to the end of the game) of this type of game. A player whose turn it is to move at such a node will choose the action that maximizes her payoff (she will be virtually bargaining with herself and will choose the best course of action). Turning to the penultimate nodes (those that precede the last choice nodes), a player whose turn it is to move at such a node will foresee the subsequent choices that will be made and will virtually renegotiate

[12] A player's information set is a collection of that player's decision nodes such that she cannot distinguish between any of these nodes. To illustrate, consider a game between two players that has 8 stages. Suppose that the two players move sequentially with player 1 choosing in the odd-numbered stages while player 2 choosing in the even-numbered stages. The set of player 1's information sets is the union of *all* of her information sets in *each* odd-numbered stage. For example, the set of player 1's information sets in period 3 is given by the set of all choice nodes in period 3 and each such information set corresponds to a distinct scenario how player 1 played in period 1 and player 2 played in period 2. Finally, the set of player 2's information sets can be defined similarly. For more details see, e.g., Binmore (2007).

[13] A subgame is any part of the game that starts with a singleton information set such that (i) all successors of a node in the subgame belong to the former and (ii) if a node belongs to the subgame then all nodes in the same information set with that node belong to the subgame.

[14] Note that the renegotiation as well as reaching an agreement in each subgame are virtual for SVBE.

(with herself) an action that maximizes her payoff, given the subsequent optimal choices. Continuing this backward induction procedure, we conclude that at each point in the game the player whose turn it is to move will virtually negotiate with herself and choose the action that maximizes own payoff, given the subsequent optimal choices. Hence, SVBE collapses into subgame perfection.

The psychological implications of this observation are intriguing. Virtual bargaining in extensive-form games enables people or organizations to coordinate to follow mutually beneficial agreements (e.g., as we shall discuss further below, engaging in trade in the Gas Station game; cooperating in the Centipede or Finitely Repeated Prisoner's Dilemma games), where such agreements are determined prior to the game. The possibility of commitment to such agreements (i.e., that each player has to allow for the possibility that the other may go through with the agreement, even if this is not the other's best response) enables threats to engage in mutually damaging conflict to be credible deterrents; and blocks the backward induction arguments that would otherwise preclude sequences of cooperative interactions.[15]

One might conjecture that, if the possibility of virtually negotiating tacit agreements helps achieve feasible agreements that are superior to Nash equilibria, then the possibility of continual renegotiation might further expand the players' opportunities for finding mutually beneficial transactions. But we have seen that the reverse is the case: the possibility of renegotiation may fatally undermine the significance of any initial negotiation.

3. VIRTUAL BARGAINING AND TACIT COMMITMENT

In this section, we apply the concept of virtual bargaining to explain how tacit commitments can arise in a range of abstract games, which capture different types of common social and economic interactions, and which pose difficulties in the standard rational choice framework. The scenarios

[15] It is not necessary that the players pre-calculate the entire tree of possible moves at the outset of the game. Instead, the players need merely simulate, at each point in the game, what they would have agreed was reasonable in that situation, prior to beginning the game. For example, they might both agree to punish "defection" by the other prior to the game, even if such punishment were destruction to all players, with the aim of deterring such defection. The possibility of continual renegotiation undermines any such deterrence, as both parties continually start afresh and reason in a forward-looking way, and hence may reject mutually costly punishment.

we sketch are highly simplified, and are intended to help capture the abstract structure of an important type of game, rather than to model specific real-world scenarios.

A. The Gas Station game

Our first example considers a case of potentially mutually beneficial trade, but where there is possible conflict—the type of mutually beneficial interaction that is extremely common, and important, in daily life. The potential trade is sequential, and there is no third-party mechanism for enforcement if one party "defects" (by taking the good without providing payment). Intuitively, the threat of conflict appears to deter such defection. Here, the question is how the threat of conflict can serve as a credible deterrent, where acting on that threat is damaging to both players. This situation arises frequently in transactions that are costly or impossible to enforce.

Consider the transaction between a motorist and a gas station attendant on an isolated road (i.e., there is no law enforcement agency nearby), involving the following three stages.[16] First, the attendant chooses whether to fill the motorist's car with gas. After observing the attendant's choice, the motorist chooses whether to pay the attendant. Finally, the attendant chooses whether to confront the motorist and demand money, ending in mutually damaging conflict. Naively, we might consider that the stronger player (whether motorist or attendant) will prevail (by taking gas with no payment; or taking payment without providing gas). If this were the case, however, a mutually beneficial transaction would not be possible. Yet we have the intuition that a "natural" tacit agreement is possible, whereby the two players exchange gas for money; and this tacit agreement will include the provision that "defection" by either player may lead to the other engaging in mutually damaging conflict.

The mutually damaging nature of any such escalation, however, would seem to rule out that it might occur; i.e., the threat of such conflict may lack credibility and could be ignored. Indeed, the logic of subgame perfection, which is a powerful and appealing concept in many circumstances, suggests that the threat of conflict is not credible in this situation. Nonetheless, tacit agreements that include the threat of conflict do seem to frequently govern our behavior;

[16] The concrete set-up outlined below should not be taken as a realistic model of gas station transactions, but rather as a stylized example of a class of interactions, and one which is both widespread, and puzzling from the standpoint of standard rational choice theory.

moreover, people do sometimes escalate to conflict when the other player "defects," even where this conflict has clear negative consequences for both parties. Intuitively, we may ascribe such acts to emotions such as anger, the desire to "get even," or the urge to punish a perceived sleight or injustice. The psychological basis of such emotions, and the potential for conflict they can lead to, or prevent, is of great empirical and theoretical importance (e.g., Fehr & Gächter, 2002). One proposal concerning how to incorporate such factors has been to add other-regarding preferences, including preferences concerning fairness and "just" punishment, into people's utility functions (e.g., Rabin, 1993).

Here, we aim to provide a fundamentally different explanation based on tacit agreements reached through virtual bargaining. This explanation does not take emotional reactions as given — though it may help explain the origin of such reactions, in the spirit of Frank (1988) and Hirshleifer (1993). We suggest that understanding such tacit agreements is crucial to understanding how individuals can successfully engage in mutually beneficial transactions without external oversight or enforcement (e.g., with no cameras or police to ensure a fair exchange of gas and money).

In the Gas Station game, the natural tacit agreement is that the attendant fills the car with gas and the motorist pays; but if the motorist did not pay, there would be (mutually damaging) conflict. While not a subgame perfect Nash equilibrium, the natural tacit agreement *is* a feasible agreement. Crucially, the motorist cannot improve her *worst case* by not paying: because her worst case is the minimum of what happens if the gas attendant either best-responds (i.e., lets the motorist go unpunished) or follows the agreement (i.e., initiates a mutually damaging conflict). Therefore, focusing on the worst case means that the threat *is* operative: the motorist will pay the attendant to avoid the worst case where mutually damaging conflict is initiated. This mechanism seems important for understanding tacit agreements in unobserved transactions of all kinds: natural tacit agreements will include the provision that reneging on the agreement will be punished, even where such punishment is mutually damaging. Notice that this feasible agreement enables the mutual benefits of trade and is hence superior for both players to the subgame perfect equilibrium in which no transaction occurs.

Formally, consider an extensive-form game between two players; an attendant of a gas station (denoted by A) and a motorist (denoted by M). The attendant moves first and chooses whether to fill the motorist's car with gas. In Figure 1, the attendant's choices are denoted by G (provide gas) and G^c (refuse to provide gas). If the attendant chooses G^c, the game ends and both players get a payoff of 0. If the attendant chooses G, then it is the motorist's turn to make a choice. The motorist has two alternatives; to pay for the gas or to refuse payment. Following the motorist's decision, the attendant chooses whether to confront the motorist. The confrontational choice is denoted by F (for "fight") while the non-confrontational choice is denoted by F^c.

The top payoffs at the terminal nodes in Figure 1 denote the payoffs to player M while the bottom payoffs denote the payoffs to player A. The cost to player A of providing gas is equal to 4. The benefit to player M of receiving gas is equal to 8. The cost of fighting is 10 for both players. The amount of payment by player M to player A is equal to 6. Thus, trade (provision of gas by player A and paying for it by player M) is mutually advantageous.

Table 1 contains the normal form of the game of Figure 1. In this table, the motorist is the row player while the attendant is the column player. The motorist has two pure strategies; "pay" and "not pay." The attendant, Player A, has three information sets. The first is at the top of the game tree in Figure 1, where the attendant chooses whether to provide gas. The other two follow the motorist's choice of payment. The triples at the top of columns in Table 1 denote player A's pure strategies and specify, respectively, his choices at each of these information sets. The first element of a triplet denotes player A's choice whether to provide gas, the second element provides his choice following payment by player M, and the third element provides his choice following non-payment.

The extensive-form game in Figure 1 has a unique subgame perfect equilibrium. In this equilibrium, the attendant does not provide gas and does not fight irrespective of the motorist's choice. The motorist does not pay. Thus, according to the logic of subgame perfection, neither a threat to fight nor a promise to pay are credible.

We now turn to determining the set of virtual bargaining equilibria and the set of subgame virtual bargaining equilibria. Consider the strategy profile where the motorist pays while the attendant provides gas and subsequently does not fight if the motorist pays and fights with probability π if the motorist does not pay: $\big(Pay, (\pi, GF^cF; 1-\pi, GF^cF^c)\big)$. Given player A's strategy, player M will prefer to pay if and if player A fights with a sufficiently large probability: $\pi \geq 0.6$. Conversely, given player M's strategy, player A nets a payoff of 2 (since player M pays) and, hence, she is indifferent between the strategies GF^cF and GF^cF^c. Thus, the strategy profile $\left(Pay, (\pi, GF^cF; 1-\pi, GF^cF^c)\right)$, where $\pi \geq 0.6$, is a Nash equilibrium. Hence, it is also a feasible agreement.

The two players' payoffs for all strategy profiles in the set $\left\{\left(Pay, (\pi, GF^cF; 1-\pi, GF^cF^c)\right): \pi \in [\frac{p}{c_M}, 1]\right\}$ are equal to 2. The corresponding payoff vector (2, 2) Pareto dominates the payoff vectors of all other feasible agreements (i.e., both players are better off under the agreement $\left\{\left(Pay, (\pi, GF^cF; 1-\pi, GF^cF^c)\right): \pi \in [\frac{p}{c_M}, 1]\right\}$ than under other feasible agreements). Hence, the set of VBE of the Gas Station game is given by $\left\{\left(Pay, (\pi, GF^cF; 1-\pi, GF^cF^c)\right): \pi \in [\frac{p}{c_M}, 1]\right\}$. Thus, the transaction is sustained as long as the threat of conflict π is sufficiently high.

The SVBE coincides with the unique subgame perfect equilibrium of the game;[17] according to the logic of subgame perfection as well as the logic of SVBE, any agreement in the set $\left\{\left(Pay, (\pi, GF^cF; 1-\pi, GF^cF^c)\right): \pi \in [\frac{p}{c_M}, 1]\right\}$ entails an empty threat to fight – the threat to fight in the case of non-payment is not credible. Thus, under the VBE, the attendant can tacitly commit to fight in case of non-payment but does not have such commitment ability under the SVBE. This is due to the following key difference between the two concepts. Under VBE, players form a virtual bargain that governs the *entire* course of their interaction. In our example, the attendant and motorist virtually agree on whether (i) the attendant will provide gas, fight following payment, and fight following non-payment and (ii) the motorist will pay. This allows the attendant to tacitly commit to fight in the case of non-payment. For SVBE, players form a virtual bargain at each point in the game that starts a subgame. As a result, even if they virtually agreed on fighting in the case of non-payment, they will virtually renegotiate this commitment, which, in turn, will render it non-credible under SVBE.

This analysis has an important psychological consequence: that a person's tendency to form, and stick to, an entire agreement—including imposing punishments on others at a personal cost—can confer significant advantages, so long as this tendency is known to others. If *B* knows that *A* is using the VBE, then *A*'s implicit threats are credible, and *B* will avoid maltreating *A* (to

[17] In the standard terminology of game theory, this is because all the information sets of the game in Figure 1 are singletons.

A’s advantage). But if *B* suspects that *A* will re-think her actions at each step, in line with SVBE, then *B* will infer that *A*’s threats of punishment are empty, as once *B* has transgressed, *A* only loses further to enforcing punishment. Hence, *B* may maltreat *A* with impunity. Thus, to engage in successful interactions with others, it is important not merely to *be* a virtual bargainer who will fix, and not rethink, future commitments; but it is important to be known to be so. Indeed, it has been argued that one role of emotions, such as anger, is to increase the credibility with which people will stick to past commitments (e.g., implicit or even explicit threats to punish others who maltreat them), even when such punishments will be costly to the punisher. Thus, the outrage of the motorist who receives no gas, or of the gas-station attendant who receives no money, raises the likelihood of a mutually destructive conflict, where a cool reappraisal would lead to punishment being withheld. Strong emotions may overwhelm the possibility of such a cool reappraisal—and maintain the real threat to whichever party maltreats the other (see Frank, 1988 for a related discussion).

Thus, tacit agreements derived through virtual bargaining can support unobserved mutually beneficial transactions under VBE—and indeed such transactions are arguably the foundation of economic and social life: notably, early European explorers would often engage in mutually beneficial exchange with non-Western peoples during the *very first* interaction, before any common means of communication were established (e.g., Hough, 1994). Here the threat of mutually damaging conflict was, of course, very real.

Notice, moreover, that the scope of cases of mutually beneficial transactions where behavior differs from the prescriptions of subgame perfection is quite broad. For example, the situation seems not to be significantly changed if the parties can communicate, at least where “talk is cheap” (i.e., no unforgeable written agreement can later be backed by a judicial system). Moreover, the game is unchanged, from the point of view of a standard game-theoretic analysis, if one or both parties can “report” the behavior of the other to third parties or to a public forum. This is because, in the absence of any third-party verification of what happened, whether a player chooses to make such a report depends on purely forward-looking considerations, and not on what actually happened. Accordingly, they will report “foul” if it is in their interest to do so, and not otherwise; and this is independent of whether the other player actually reneged on the agreement. Thus, any such report will be no more than cheap talk, and would be ignored by receivers; and, by backward induction, would presumably not be sent in the first place. Hence, the possibility of making such reports does not help underpin mutually beneficial transaction. Virtual bargaining can, by contrast, explain how mutually beneficial transaction is possible, where direct retaliation may not be possible, but where one or both sides can report the other's "misbehavior," for example, by posting a negative review. Suppose that, from a forward-looking point of view, reporting misbehavior is somewhat costly. Then, a natural tacit agreement is that players will engage in mutually beneficial trade and will report bad behavior by the other only if it occurs. If the negative impact of such reports is sufficient (e.g., in reputational damage that undermines potentially beneficial future transactions with others), then this tacit agreement is a virtual bargain.[18]

Virtual bargaining also throws light on darker aspects of human behavior. In an interaction between a hostage-taker and the police, the police cannot rely on the observation that harming the hostage is damaging to both the police's interests (of keeping citizens from harm) and the hostage taker (who will not achieve his objectives and will face severe punishment). According to standard Nash-based reasoning, the hostage-taker's threat is empty, and the police can safely advance. Yet even before communication between the two sides has been established, there may be a tacit agreement that neither side will act: that police advance would lead to the hostages being harmed; and conversely that harming the hostages would trigger police advance.

B. The Centipede Game

Our next example, the Centipede game (Aumann, 1988; Rosenthal, 1981), also concerns sequential interactions that are potentially mutually beneficial and are unenforced by a third party. In contrast to the Gas Station game, however, the players have no available threat of retaliation (whether credible or otherwise) that might deter "defection," because the game ends immediately after such defection occurs. In the Centipede game, a mutually beneficial outcome appears to be threatened by the logic of backward induction. In a classic Centipede game, two players move alternately, with a fixed maximum number of moves. Each move either (a) ends the game or (b) continues the game, thereby slightly reducing the mover's payoff, while considerably increasing the payoff to the other player. If both players continue for many moves,

[18] We have here benefitted from discussion with Bob Sugden, although the views expressed here, of course, our own.

they both end up with a substantial positive payoff—to their considerable mutual benefit. A "natural" tacit agreement seems possible whereby both players always choose to continue the game. Moreover, it is intuitively natural to expect that a player will feel annoyed or even angry if the other player drops out at the start (or near the start) of the game—i.e., where the tacit agreement has been violated. By construction, in the Centipede game, backward induction from the last decision node implies that the game will end on the very first move. But we have the strong intuition that this is not what the players would tacitly agree (or, indeed, would explicitly agree if they could engage in pre-play communication). Furthermore, laboratory experiments show that people frequently do play (nearly) to the end of the Centipede game, rather than defecting on the very first round following the unique subgame perfect equilibrium of the game (see, e.g., Krockow et al., 2016; Krockow et al., 2018; McKelvey & Palfrey, 1992). Many existing explanations of the observed behavior in the laboratory experiments rely on an argument that cooperation in the form of continuing the game is sustained by each player's belief that their opponent has a type who prefers to continue the game.[19] Our explanation is fundamentally different: cooperation is sustained by identifying and following a tacit agreement.

In the Centipede game, as we have noted, no threats are possible: the question is how to justify continuing playing the game in the light of a backward induction argument that "defection" will occur on the last round, and hence the second-to-last round, and so on. The intuitively natural tacit agreement that both players continue to the end of the game is a feasible agreement—and indeed it is the best feasible agreement for both players (and hence it will be uniquely selected in the process of virtual bargaining). Suppose, for concreteness, that the game has an odd number of rounds ranging from 1 to some finite odd number. Suppose the "odd" player moves in the odd-numbered rounds while the "even" player moves in the even-numbered rounds. Consider the agreement where the parties cooperate until the end of the game. In this case, the "even" player's best response is to stick to the agreement and gain the benefit of the odd player's final, as well as all of the preceding, moves. Thus, the worst payoff (as defined above) for the odd player is the payoff she receives from the strategy profile where the two players cooperate till the very end. The situation for this agreement is slightly different for the

[19] Similarly, Kreps et al. (1982) demonstrate that cooperation in early rounds of a finitely repeated Prisoner's dilemma game can be supported by the presence of cooperative types who gain utility from acting cooperatively.

even player. The worst scenario for the even player is that the odd player best-responds (rather than follows the agreement): i.e., the odd player stops the cooperation on the very last round. Nonetheless, though, the even player's payoff even in this worst case is still high; and, crucially, higher than the worst case from any other feasible agreement, including the Nash equilibrium, in which the game terminates on the first round.

The extensive form of the game in Figure 2 depicts the version of the Centipede game which was put forth by Aumann (1988). There are two piles of money on the table and two players, A and B, who move sequentially. Initially, one of the piles contains \$1 while the other pile contains \$4. Every time it is a player's turn to make a choice, she can either take the larger of the two piles (denoted by "t" in Figure 2) or pass (denoted by "p"). If a player selects t, the game ends, with her getting the large pile and her opponent getting the small pile. If a player passes, both piles are doubled and the play proceeds to the next player.

The game has a unique subgame perfect equilibrium where both players choose "t" at all of their decision nodes. Player A's payoff is 4 while player B's payoff is 1 in this equilibrium. The set of Nash equilibria includes all strategy profiles where the two players mix between the choices tpp, tpt, ttp and ttt according to arbitrary probability distributions. Thus, the Nash equilibrium behavior leads to the same equilibrium outcome: the game ends on the very first move. We characterize the VBE and SVBE of the game in the following:

Proposition 1. Consider a Centipede game with arbitrary finite number of stages $n \geq 4$.
(a) In the unique VBE of the game, each player chooses "p" whenever it is her turn to move.
(b) In the unique SVBE of the game, each player chooses "t" whenever it is her turn to move.

Thus, under the VBE, the two players tacitly commit to always choose "p". Deviations from this strategy profile are compared through the lens of worst payoffs and not payoffs. Since the worst payoffs of the two players under such commitment exceed the corresponding worst payoffs for individual deviations, the tacit commitment to always choose "p" is credible and the players end up doubling the pile until the very end of the game. More specifically, player B's best response to player A's strategy ppp is to pass the pile twice and to take the money in the last period: ppt. Hence, player A's worst payoff for the strategy profile (ppp, ppp), where both players always pass the pile, is equal to 32 (the minimum of her payoff of 256 for the virtual bargain (ppp, ppp) and her payoff of 32 for the deviation). Consider now a deviation of player A. Suppose player A deviates to strategy ppt, which entails deviating to the virtual bargain (ppt, ppp). Player B's best response to player A's strategy ppt is to pass the pile once and to take the money afterwards: ptt. Hence, player A's worst payoff for the strategy profile (ppt, ppp), where player A passes the pile in the first two rounds while player B always passes, is equal to 8 (the minimum of her payoff of 64 for the virtual bargain (ppt, ppp) and her payoff of 8 for the deviation). Since $8 < 32$, player A will not have an incentive to deviate to the virtual bargain (ppt, ppp). Similar calculations reveal that other deviations are not profitable either.

As in the Gas Station game, the ability to stick to the initial agreement, rather than to renegotiate the bargain at each step in the game, is crucial. The SVBE of the Centipede game coincides with the unique subgame perfect equilibrium—the possibility of continual renegotiation entirely undermines the ability to make a mutually beneficial tacit agreement.

C. The Finitely Repeated Prisoner's Dilemma game

We now turn to the finitely repeated version of what is probably the most widely studied game in both behavioral and theoretical game theory, the Prisoner's Dilemma. In the one-shot version of the Prisoner's Dilemma, players can either defect or cooperate. Irrespective of what the other does, each player will themselves do strictly better by defecting. Hence, the only Nash equilibrium of the game is (defect, defect). But the game is designed so that this pair of strategies is inferior to (cooperate, cooperate) for both parties—i.e., there is a mutually beneficial outcome that both parties could potentially achieve, if only they could make an enforceable agreement to do so. But forward-looking rational agents cannot do this.

In laboratory experiments, both cooperation and defection occur in the Prisoner's Dilemma, in proportions that depend on complex aspects of the payoffs, the setting for the game, other games that have recently been played, and many other factors (Oskamp & Perlman, 1965; Rapoport & Chammah, 1965; Vlaev & Chater, 2006). Can virtual bargaining help explain why cooperation occurs in one-shot Prisoner's Dilemmas? While the informal idea that people may mentally simulate a possible agreement to cooperate (and feel obligated to follow this agreement) may be relevant to explaining cooperation, the formal machinery of virtual bargaining does not give this prediction. This is because the formal notion of virtual bargaining assumes that people choose agreements to maximize the worst case for them—and the worst case for a putative cooperate-cooperate agreement is clearly that the other exploits the cooperator by defecting, so that the cooperator gets the worst possible outcome. Hence the VBE for Prisoner's Dilemma is, like the Nash equilibrium, defect-defect.

But the picture is very different if the Prisoner's Dilemma is not one-off, but is repeated a finite number of times. This set-up shares many features with the Centipede game (i.e., sequential interactions that are potentially mutually beneficial and are unenforced by a third party), but unlike the Centipede game, defection in earlier rounds does not end the game. Although the players would benefit from cooperating throughout the entire game, they have an incentive to defect in the final stage; and according to the logic of backward induction, the players will defect through all prior stages. Here, we have the intuition that a "natural" tacit agreement is for the players to cooperate to the end or nearly the end of the finite number of repetitions of the Prisoner's Dilemma. Moreover, such an intuitive tacit agreement seems to include the provision that, if one player switches to "defect," then the other may also defect thereafter. Of course, if the defecting player "repents" by resuming cooperation, then the other player may, but need not, re-establish cooperation. Laboratory experiments have shown that people engage in such complex patterns of behavior, rather than both defecting on the very first trial and all subsequent stages, as the subgame perfect equilibrium of the game would predict (see, e.g., Andreoni & Miller, 1993; Embrey, Frechette, & Yuksel, 2018).

The intuition for cooperating to the end or nearly the end of the finite number of repetitions is as follows. Backward induction implies that both players should defect in all n rounds. But a tacit agreement to cooperate throughout all rounds can be superior; such an agreement will include the "threat" to defect henceforth as soon as the other defects once. The best response to a player holding to such an agreement will be to cooperate until the very last turn, and then to defect. The worst case for the player holding to the agreement therefore sums up the benefits of $n - 1$ rounds of mutual cooperation followed by low payoff where cooperation is met by defection in the final round. Which of these is chosen as the virtual bargain? The key criterion is whether the worst payoff for this cooperative feasible agreement is superior to the worst payoffs of alternative agreements that involve unilateral defections (this will of course depend on n and the precise payoffs of the game). Where it is, the tacit agreement to cooperate rather than to defect will be the virtual bargain (although, of course, it is still possible that either player will defect in the last round).

Formally, the normal form of a Prisoner's Dilemma game is presented in Table 3. Two players, 1 and 2, simultaneously and independently choose between "cooperate" (denoted by "c") and "defect" (denoted by "d") strategies. In the unique Nash equilibrium of the game, both players "defect".

Suppose now that the game of Table 3 is played finitely many times with both players learning the opponent's choice in the previous round. As we argued above, in the unique subgame perfect equilibrium of this game, both players defect for all possible histories of the game. The set of VBE and the unique SVBE are characterized in the following:

Proposition 2. Consider an arbitrary finitely repeated Prisoner's Dilemma game with $n \geq 2$.
(a) Both players will choose to cooperate along the equilibrium path[20] in any VBE of the game. The players can "enforce" this outcome by using "grim-trigger strategies" (where a player starts by cooperating, cooperates as long as the opponent always cooperated in the past, and defects for the rest of the game after her opponent defects).
(b) In the unique SVBE, the two players will choose to defect for each history of the game.

As we elaborate in Appendix, many strategy profiles can "enforce" a VBE where both players always cooperate along the equilibrium path. One of these is the grim trigger strategy. The intuition behind this finding is as follows. Consider the agreement where the two players use the grim trigger strategies. The best response to the grim trigger strategy is cooperate in the first $n-1$ stages of the game and to defect in the last stage. Hence, the worst payoff from the strategy profile where each player follows the grim trigger strategy is equal to $(n-1)\cdot(-1)+(-10)=-(n+9)$. For illustrative purposes, suppose a player considers deviating from this strategy profile and decides to cooperate in the first stage and to defect in all of the remaining stages of the game. The deviating player's worst payoff from that strategy profile, where she cooperates in the first stage and defects in all of the remaining stages while the opponent plays the grim-trigger strategy, is equal to $(n-1)(-7)+(-10)=-7n-3$. But this worst payoff is strictly smaller

[20] The equilibrium path consists of all information sets that are reached with strictly positive probability when the players follow their equilibrium strategies.

than the worst payoff of $-(n+9)$ from the pair of the grim-trigger strategies. In a similar fashion, one can demonstrate that neither player has a deviation that can improve her worst payoff. Hence, the pair of the grim-trigger strategies is feasible. Note that the subgame perfect equilibrium where both players defect in all stages of the game is also feasible. However, the virtually bargaining player will choose the agreement where they cooperate throughout the game over the subgame perfect equilibrium and other feasible agreements.

In contrast, because the VBE of the static Prisoner's Dilemma game coincides with its unique Nash equilibrium, the SVBE of the repeated game coincides with the unique subgame perfect equilibrium.

The logic of our derivations of the set of VBE in the repeated Prisoner's Dilemma game can also be used to shed new light on repeated interactions where players may find it to their advantage to alternate their moves in a coordinated manner. Recall that a turn-taking bargain minimizes the degree to which either person can be exploited, if the other decides to maximize their own interests, rather than sticking to the agreement (this recalls the farming example, above, of course, where alternation of effort is the preferred strategy).

Consistent with the VBE's prediction, Sheridan, Sharma and Cockerill (2014) argue that turn-taking is a key landmark in a child's development process. Indeed, Melis et al. (2016) demonstrate that 5-year old human children, in contrast to younger human children and chimpanzees, cooperate by taking turns in getting a reward in a repeated collaboration task to obtain that reward. Turn-taking emerges across a variety of environments involving strategic interactions between adults in economic experiments. Leo (2017) finds a significant amount of turn-taking in the repeated "Volunteer's Dilemma" game. Sibly and Tisdell (2018) report a considerable amount of turn-taking in a finitely repeated modified Prisoner's Dilemma game. Moreover, many subjects in their experiment engage in turn-taking until the very last round.

To demonstrate the plausibility of virtual bargaining as an explanation of turn-taking behavior, consider the normal form game in Table 4, which is a modified Prisoner's Dilemma game. As in the standard Prisoner's Dilemma game in Table 3, defection is a strictly dominant strategy in the modified Prisoner's Dilemma game. The subgame perfect equilibrium of the game of Table 4, repeated finitely many times, is for both players to defect for all possible histories of the game. However, the efficient outcome of the interaction requires one of the players to cooperate and the other to defect. For example, if the game of Table 4 is played twice, and in the first round the players choose (c, d) whereas in the second round they choose (d, c), then, assuming no discounting, each player nets a payoff of -10. This is higher than the payoff of -12 from cooperating in both rounds of the game.

Suppose, without any loss of generality, that the game has an even number of stages and consider the turn-taking sequence that begins with (c, d) and then alternates between (c, d) and (d, c) thereafter. It is also assumed that the game of Table 4 has more than four repetitions so that this alternation might occur at least twice. Suppose also that if either player deviates from this sequence, the opponent punishes by defecting for the rest of the game. We will call this strategy profile a turn-taking agreement with punishments, or simply a turn-taking agreement. We have:

Proposition 3: Consider an arbitrary finitely repeated modified Prisoner's Dilemma game.
(a) Both players will choose to play the turn-taking sequence along the equilibrium path in any VBE of the game. The firms can "enforce" this outcome by punishing deviations with defecting for the rest of the game.
(b) In the unique SVBE, the two players will choose to defect for each history of the game.
Proof: See Appendix

Thus, the turn-taking agreement is a VBE. However, the possibility of virtual renegotiation nullifies the possibility of forming such agreement. It is interesting to compare the turn-taking agreement of Proposition 3 to an agreement that also involves alternation but at a less frequent rate. For example, the players could contemplate playing (c, d) twice, then playing (d, c) twice, and so on. For this type of an agreement to be feasible, the game has to be sufficiently long. Moreover, because the worst-case scenario that the opponent will best-respond (instead of sticking to the agreement for these less frequent turn-taking agreements) is less attractive than under the most frequent turn-taking agreement, the worst payoff for the two players under the former agreements will be lower than under the turn-taking agreement where the two players alternate in each stage. Hence, strict alternation, rather than longer cycles of alternation, will be selected as a virtual bargain (although, of course, the players nonetheless face the challenge of deciding the *order* of the alternation—i.e., whether (*c*, *d*), (*d*, *c*), (*d*, *c*), …etc., or the reverse).

As in the Gas Station and Centipede games, the ability to create a bargain concerning behavior throughout the game at the outset, rather than continually renegotiating as the game proceeds, is crucial. The collaborative behavior for mutual benefit (whether cooperating, or turn-taking) is lost if virtual bargains are renegotiated—i.e., in these games, the SVBE is the same as the Nash equilibrium.

D. Collusion: The Repeated Bertrand Competition game

From the examples we have considered so far, it is natural to ask whether or not the notion of a SVBE is actually needed, as it has coincided with the predictions of the conventional subgame perfect equilibrium. But this is not always the case.

An example of competitive interaction illustrates the point. Consider the situation in which two firms set a price for the same good each day (perhaps these are neighboring gas stations). Both firms have an incentive to set the price as high as possible, but each will make more profit by slightly undercutting the other (because they will then obtain more customers, offsetting the loss from the slightly lower price). The standard Nash equilibrium analysis of the one-shot version of this game typically leads to the result that both firms choose relatively low prices, and to the detriment of both; and the sequential version of the game will lead to this worst possible outcome (for the firms) applying relentlessly across all time periods. If the players could collude by explicitly agreeing to set their prices high each day (typically in violation of the law), they would both reap considerable profits. It turns out that virtual bargaining, whether through VBE or SVBE allows them to collude in just this way, without having explicitly to communicate (see Appendix).

Under VBE, the two firms manage to virtually agree to choose a collusive price vector that maximizes the sum of the two firms' profits in every period of the game. This is the outcome that the firms would achieve if they had access to an external mechanism that could perfectly enforce any price choices by the two firms. The firms tacitly commit to "enforce" this outcome by punishing deviations from the collusive price by choosing sufficiently low prices for a sufficiently large number of periods. Although the SVBE does not enable the players to commit to a course of action at the outset of the game, they will still choose a price that is higher than the price under the subgame perfect equilibrium — specifically, they will choose the virtual bargaining equilibrium of the static game in each period and thus achieve a "middle of the road" outcome, which is worse than the VBE of the dynamic game but better than the subgame perfect equilibrium. The ability to partially collude under the SVBE stems from the bite of virtually bargaining in a single shot interaction of the Bertrand game (see Melkonyan, Zeitoun and Chater (2018) for more details). The firms realize that if they raise their prices, the opponent will undercut them. However, they will not undercut by a very large amount, and therefore increasing a price becomes an attractive option for a firm guided by the worst payoff. Thus, the firms collude under both VBE and SVBE, but the degree of collusion is higher under VBE than SVBE.

This example illustrates the dark side of virtual bargaining—tacit collusive behavior may be mutually beneficial for those engaging in the collusion, but potentially very harmful to others (e.g., the customers paying abnormally high prices). This point arises throughout social behavior, not just in market contexts—virtual bargaining provides a way of stabilizing collusive outcomes, involving, for example, engaging in and not reporting corruption or other misdemeanors, or underpinning the codes of behavior in organized crime (e.g., Gambetta, 1996). The conventional rational choice perspective, in contrast, often fails to explain how people can successfully work together in teams over finite, but lengthy, collaborative projects, or maintain long-term social relationships. Suppose a group of people work together on some joint activity of mutual benefit (whether organizing a party, clearing snow, or cooking a meal). The project will succeed in proportion to how much effort is made by the team members. Suppose that outright shirking is punished (perhaps by the team members themselves)—but that, as is true in most contexts, people are unable to detect people making a little less effort than the average. But then each player may be incentivized to work just a little less hard than the average worker—just enough to avoid censure. Then the curse of backwards induction looms. As the project proceeds, people will work, individual workers will make less effort, pulling down the average effort made; and this will allow individual works to adjust downwards their effort levels, pulling the average down further, until effort has collapsed to zero. Indeed, according to a standard rational choice model, the workers should all predict this outcome in advance, and hence stop working from the very beginning. But this “race to the bottom” need not be predicted by virtual bargaining. Instead, each person can reason that the best agreement would be for them all to work hard—if they all do, there is a mutually beneficial outcome for all. If some shirk slightly (it is not in their interest to shirk a lot), then there is still a mutually beneficial outcome. Hence a high level of effort throughout the project may be maintained.

This analysis provides an interesting new perspective on the question of "cheater detection," which some evolutionary psychologists have argued is crucial to the maintenance of social cohesion (Cosmides & Tooby, 1992). It is typically assumed that cheaters must be detected and punished sufficiently that cheating does not 'pay,' to make socially beneficial collaboration possible. But cheating is not always detectable. We have seen that, in contexts where people's incentives are such that, if they maximize their own interests they will only cheat a little, the virtual bargain to maintain high levels of effort on a cooperative activity may be maintained.

4. COMPARISON WITH ALTERNATIVE REASONING-BASED ACCOUNTS

Our analysis above shows how the three equilibrium concepts—the conventional subgame perfect equilibrium (SPE), and our new notions of virtual bargaining equilibrium (VBE), and subgame virtual bargaining equilibrium (SVBE)—make predictions that coincide, or diverge, depending on the type of strategic interaction. We argue that all three concepts are useful to model different phenomena. In this section, we discuss how the three concepts imply different degrees of tacit commitment. Then we contrast the virtual bargaining approach with alternative modes of non-standard reasoning.

Note, first, that formal modelling in psychology in the rational choice tradition typically focuses on finding and choosing between Nash equilibria (for one-shot games) and SPE for dynamic games, although this terminology is not always used. For example, a Bayesian agent may attempt to model the mind of another, by inferring her beliefs from her action; but she may in turn assume that the other is choosing her actions knowing that this will occur, and so on. Typically, equilibria are found by iterating the process of prediction and finding "fixed points" which will correspond to conventional Nash equilibria or SPE (for example, see Bayesian models of teaching or communication). Thus, on the one hand, according to the virtual bargaining account people engage in simulated bargaining to establish tacit commitments — which include mutually beneficial arrangements that may differ from Nash equilibria (or, for dynamic games, SPEs). On the other hand, the above-mentioned psychological accounts assume an iterative process of simulating, and best-responding to, the *actions* of the other agent, hence staying within the domain of Nash equilibria (and SPEs). Thus, the innovations embodied in this important class of psychological models—which help provide rational accounts of mind-reading (Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017), teaching (Shafto, Goodman, & Griffiths, 2014), communication (Frank & Goodman, 2012), and joint planning (Wang et al., 2021)—are orthogonal to the novel contribution of the virtual bargaining approach (for discussion, see Chater et al., in press).

We stress that, in psychological models as well as older economic game-theoretic accounts, the SPE sheds light on many crucial phenomena, such as the conditions under which threats are "empty" (Selten, 1975, 1978) and questions of dynamic consistency (Kydland & Prescott, 1977). Reasoning by SPE implies that at any point in the history of the game, the individual ignores the opponent's past actions when deciding on the present action. Whether my counterpart cooperated or defected in prior stages of the game does not restrict me in any way. This mode of reasoning has the advantage that the individual focuses only on the future consequences of her current decision. But in many strategic interactions, such as the ones discussed above, many people deviate from the SPE.

The SVBE is similar to the SPE in that individuals ignore their counterpart's actions in prior stages of the game and focus exclusively on the future consequences of the current decision. In contrast to the SPE, however, SVBE reasoners adopt virtual bargaining (i.e., a more collaborative mode of reasoning than Nash reasoning) in each stage of the game. In the games in Sections 3.A-3.C, SVBE leads to actions that coincide with those under SPE. But this is not always the case. For example, we noted that in the repeated Bertrand competition game, the players are able to achieve a mutually beneficial collusive outcome under SVBE not allowed by SPE. This finding highlights the limitations of collaborative reasoning if individuals are unable to tacitly commit to a course of action.

The VBE, in contrast, implies that individuals are able to tacitly commit to a certain course of action (in addition to reasoning collaboratively). Crucially, VBE reasoners do not commit to any arbitrary course of action, but to one that represents a "natural" tacit agreement between the players—i.e., the agreement that they would reach if they were able to bargain explicitly. Thus, VBE endogenizes the option to which people may tacitly commit. A threat that is empty under SPE can become credible under VBE, such as the threat of fighting or reporting a defector in the Gas Station game, leading to a mutually beneficial outcome. Moreover, people can experience a psychological cost—a real sense of discomfort or outrage—if their opponent deviates from a "natural" agreement (e.g., Fiske & Rai, 2014). Thus, the VBE provides a psychological foundation for explaining people's tacit commitments, and the psychological costs they experience if their opponent departs from a virtual bargain.

The concept of virtual bargaining offers a unified account that explains our strong intuitions behind the types of games discussed in this paper and many other strategic interactions. It is instructive to compare virtual bargaining with other reasoning-based accounts of strategic behavior that predict deviations from Nash and subgame perfect equilibria, such as quantal response equilibrium (QRE) (e.g., McKelvey & Palfrey, 1995) and the level-k and cognitive hierarchy models (e.g., Camerer, Ho, & Chong, 2004).[21] How do they differ?

First, virtual bargaining explains how players develop and follow tacit agreements based on a mental simulation that enables them to envision the "natural" tacit agreement. This psychological mechanism is different from QRE (that relies on players making errors) and level-k and cognitive hierarchy models (that (i) are defined iteratively starting from a "naïve" level-0 reasoner who best-responds to some distribution of the opponents' strategies[22] and (ii) rely on players believing that they are more sophisticated reasoners (by 1 level) than their counterparts and best-responding based on these beliefs). While all these accounts have a persuasive *raison d'être,* they highlight distinct aspects of human cognition. QRE and level-k/cognitive hierarchy models emphasize deficiencies in reasoning, whereas virtual bargaining develops a positive vision of the agreements that players may want to achieve. Thus, virtual bargaining appears to be particularly suited to explain our intuitions behind "natural" tacit agreements, which are the focus of this paper.

Second, these accounts take different approaches in extending from normal-form to extensive-form games—this can be shown in alternative explanations of players' behavior in the Centipede game. In QRE, McKelvey and Palfrey (1992) model extensive-form games using the agent quantal response equilibrium (AQRE), where each player treats her future self as an independent player with a known probability distribution over actions. In level-k reasoning, Kawagoe and

[21] A related non-Nash equilibrium is the maximin equilibrium (Ismail, 2014). According to this account, when evaluating a strategy profile, a player is guided by the minimum of her payoff from that strategy profile and the minimum payoff taken over the opponents' better responses (rather than the best responses in the present paper). Ismail (2014) also introduces the maximin equilibrium for games with more than two players. The predictions of virtual bargaining and maximin equilibria coincide in some games and diverge in others.

[22] For example, a uniform probability distribution over the opponents' undominated strategies.

Takizawa (2012) assume that players implement behavioral strategies, which correspond to reducing the extensive-form into normal-form games. In this paper, we have presented two possible definitions of the virtual bargaining equilibrium in extensive-form games. Our standard definition, the VBE, aligns with Kawagoe and Takizawa's (2012) approach and assumes that players can tacitly commit to a certain course of actions. Indeed, such commitment is crucial to our account: our alternative definition without commitment, the SVBE, assumes that such commitment is not possible and collapses to the subgame perfect equilibrium in games where players choose sequentially and all information is complete. In contrast to Kawagoe and Takizawa's (2012) approach, Ho and Su (2013) introduce a level-k model for extensive-form games that relies on a game's sequential structure (and does not reduce to normal forms) and allows for belief updating. Both Kawagoe and Takizawa (2012) and Ho and Su (2013) model behavioral traits different from those in the present paper. Most notably, they do not explicitly model tacit commitment.

It is also instructive to compare the virtual bargaining approach with team reasoning (e.g., Bacharach, Gold, & Sugden, 2006). Team reasoning asks how "we" should act in order to benefit the team (whose objective might, for example, be a sum of the utilities of the individual members of the team). Virtual bargaining provides a different mechanism by which people can coordinate in a "team-like" way: the collective behavior is governed not by maximizing a team objective, but by what the parties would agree. Thus, virtual bargaining enables the actors to have individual objectives. Moreover, unlike team reasoning, the virtual bargaining account does not require any psychological identification of players as being, in any sense, "on the same side."

Virtual bargaining can, moreover, lead to coordinated and reasonably efficient behavior in certain cases where Nash or subgame perfection would predict inefficient outcomes; and it can do so without any assumptions about other-regarding preferences (although such preferences can, in the standard way, be built into the utility functions of the players, if required). According to virtual bargaining, players have an inherent tendency to "stick" to agreements to which they are tacitly committed—a tendency strong enough to have at least the potential to deter the violation of the agreement by others, even if punishment is mutually damaging. From the perspective of virtual bargaining, the tendency to punish "defection" by others arises not from other-regarding preferences (e.g., feelings of outrage that require "revenge"), but simply from carrying out the prior agreement. We conjecture that the creation, maintenance, and violation of tacit agreements may be an important source of positive or negative emotions towards other people; such emotions, in turn, may support the tendency to “go through” with the virtual bargain (e.g., feelings of outrage may drive punishment of the other, as specified by the virtual bargain, even where that punishment is mutually destructive, Fiske & Rai, 2014). Notice, too, that the process of virtual bargaining is a highly flexible reasoning mechanism that can lead to a wide variety of complex behaviors (e.g., turn-taking), depending on the nature of the interaction. It is therefore more foundational than specific strategies, such as reciprocation or conditional cooperation.

5. CONCLUSION

Human interactions are frequently guided by tacit agreements; and in many situations such agreements can be generated flexibly and spontaneously, so that explicit communication is unnecessary. According to the virtual bargaining account, tacit agreements can be derived by the parties asking: what would we agree if we were able to communicate? Where it is clear what agreement would be reached, the process of communication and bargaining can be short-circuited.

We have outlined a formal theory of the reasoning that explains tacit commitment and can underpin the creation of such tacit agreements. The virtual bargaining account offers a unified framework to explain how people identify, commit to, and behave according to tacit agreements. Even without infusing imperfect information or assuming non-standard preferences, this theory provides a good fit with some stylized facts of empirical data: people do engage in mutually beneficial exchange when unobserved; collude in price-setting competition; play close to the end in the Centipede game; and cooperate to a considerable degree in the finitely repeated Prisoners’ Dilemma game. The ability to create and maintain such tacit agreements also seems crucial to the creation of the norms, conventions, and institutions that govern human life.

REFERENCES

Ainslie, G. (1992). *Picoeconomics: The Interaction of Successive Motivational States within the Individual*. Cambridge: Cambridge University Press.

Ainslie, G. (2001). *Breakdown of Will*. Cambridge: Cambridge University Press.

Andreoni, J., & Miller, J. H. (1993). Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *Economic Journal, 103*(418), 570-585. doi:10.2307/2234532

Argyris, C. (1960). *Understanding Organizational Behavior*. Homewood, IL: The Dorsey Press.

Aumann, R. (1988). Preliminary notes on integrating irrationality into game theory. Mimeo, International Conference on Economic Theories of Politics, Haifa.

Bacharach, M., Gold, N., & Sugden, R. (2006). *Beyond Individual Choice: Teams and Frames in Game Theory*. Princeton, NJ: Princeton University Press.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour, 1*(4). doi:10.1038/s41562-017-0064

Baker, G., Gibbons, R., & Murphy, K. J. (2002). Relational contracts and the theory of the firm. *Quarterly Journal of Economics, 117*(1), 39–84. Retrieved from <Go to ISI>://000173476500002

Bangerter, A., & Clark, H. H. (2003). Navigating joint projects with dialogue. *Cognitive Science, 27*, 195-225.

Benabou, R., & Tirole, J. (2004). Willpower and personal rules. *Journal of Political Economy, 112*(4), 848-886. doi:10.1086/421167

Bernheim, B. D., Peleg, B., & Whinston, M. D. (1987). Coalition-proof Nash equilibria.1. Concepts. *Journal of Economic Theory, 42*(1), 1-12. doi:10.1016/0022-0531(87)90099-8

Binmore, K. (2007). *Playing for Real: A Text on Game Theory*. Oxford, UK: Oxford University Press.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(6), 1482-1493. doi:10.1037/0278-7393.22.6.1482

Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics, 119*(3), 861-898. doi:10.1162/0033553041502225

Chater, N., Misyak, J. B., Melkonyan, T., & Zeitoun, H. (2016). Virtual Bargaining: Building the foundations for a theory of social interaction. In J. Kiverstein (Ed.), *Routledge Handbook of the Philosophy of the Social Mind* (pp. 418-430). Abingdon, UK: Routledge.

Chater, N., Zeitoun, H., & Melkonyan, T. (in press). The paradox of social interaction: Shared intentionality, we-reasoning and virtual bargaining. *Psychological Review*.

Cialdini, R. B. (2007). Descriptive social norms as underappreciated sources of social control. *Psychometrika, 72*(2), 263-268. doi:10.1007/s11336-006-1560-6

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Clark, H. H. (2006). Social actions, social commitments. In N. J. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality*. Abingdon, UK: Routledge.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp. 163-228). New York, NY: Oxford University Press.

Dunning, D., & Fetchenhauer, D. (2011). Understanding the psychology of trust. In D. Dunning (Ed.), *Social Motivation* (pp. 147–169): Psychology Press.

Embrey, M., Frechette, G. R., & Yuksel, S. (2018). Cooperation in the finitely repeated Prisoner's Dilemma. *Quarterly Journal of Economics, 133*(1), 509-551. doi:10.1093/qje/qjx033

Farrell, J., & Rabin, M. (1996). Cheap talk. *Journal of Economic Perspectives, 10*(3), 103-118. Retrieved from <Go to ISI>://WOS:A1996VD58300007

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature, 415*(6868), 137–140. Retrieved from <Go to ISI>://000173159300032

Fiske, A. P., & Rai, T. S. (2014). *Virtuous Violence: Hurting and Killing to Create, Sustain, End, and Honor Social Relationships*. Cambridge, UK: Cambridge University Press.

Frank, M. C., & Goodman, N. D. (2012). Predicting Pragmatic Reasoning in Language Games. *Science, 336*(6084), 998-998. doi:10.1126/science.1218633

Frank, R. H. (1988). *Passions within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton.

Fukuyama, F. (1996). *Trust: Human Nature and the Reconstitution of Social Order*. New York, NY: Simon and Schuster.

Gambetta, D. (1996). *The Sicilian Mafia: The Business of Private Protection*. Cambridge, MA: Harvard University Press.

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science, 1*(2), 292-304. doi:10.1111/j.1756-8765.2009.01020.x

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics, 18*(2), 141-153. doi:10.1016/0304-4068(89)90018-9

Hart, O. (1995). *Firms, Contracts, and Financial Structure*. Oxford: Oxford University Press.

Hirshleifer, J. (1993). The affections and the passions: Their economic logic. *Rationality and Society, 5*(2), 185-202.

Ho, T. H., & Su, X. M. (2013). A dynamic level-k model in sequential games. *Management Science, 59*(2), 452-469. doi:10.1287/mnsc.1120.1645

Hough, R. (1994). *Captain James Cook: A biography*. London: Hodder & Stoughton.

Ismail, M. (2014). Maximin equilibrium. *GSBE Research Memoranda, No. 037. Maastricht: GSBE*.

Jensen, M. C. (2009). Integrity: Without it nothing works. *Rotman Magazine: The Magazine of the Rotman School of Management*, 16-20.

Kalai, E. (1977). Proportional solutions to bargaining situations: Intertemporal utility comparisons. *Econometrica, 45*(7), 1623-1630.

Kawagoe, T., & Takizawa, H. (2012). Level-k analysis of experimental centipede games. *Journal of Economic Behavior & Organization, 82*(2-3), 548-566. doi:10.1016/j.jebo.2012.03.010

Kreps, D. M., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory, 27*(2), 245-252. doi:10.1016/0022-0531(82)90029-1

Krishna, V., & Serrano, R. (1996). Multilateral bargaining. *Review of Economic Studies, 63*(1), 61-80. doi:10.2307/2298115

Krockow, E. M., Colman, A. M., & Pulford, B. D. (2016). Cooperation in repeated interactions: A systematic review of Centipede game experiments, 1992-2016. *European Review of Social Psychology, 27*(1), 231-282. doi:10.1080/10463283.2016.1249640

Krockow, E. M., Pulford, B. D., & Colman, A. M. (2018). Far but finite horizons promote cooperation in the Centipede game. *Journal of Economic Psychology, 67*, 191-199. doi:10.1016/j.joep.2018.07.002

Kydland, F. E., & Prescott, E. C. (1977). Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy, 85*(3), 473-491. doi:10.1086/260580

Leo, G. (2017). Taking turns. *Games and Economic Behavior, 102*, 525-547. doi:10.1016/j.geb.2017.02.003

Malle, B. F. (2021). Moral judgments. *Annual Review of Psychology, 72*, 293-318. doi:10.1146/annurev-psych-072220-104358

March, J. G., & Olsen, J. P. (2008). The logic of appropriateness. In R. E. Goodin, M. Moran, & M. Rein (Eds.), *The Oxford Handbook of Public Policy*. Oxford: Oxford University Press.

McKelvey, R. D., & Palfrey, T. R. (1992). An experimental study of the centipede game. *Econometrica, 60*(4), 803-836. doi:10.2307/2951567

McKelvey, R. D., & Palfrey, T. R. (1995). Quantal response equilibria for normal-form games. *Games and Economic Behavior, 10*(1), 6-38. doi:10.1006/game.1995.1023

Melis, A. P., Grocke, P., Kalbitz, J., & Tomasello, M. (2016). One for you, one for me: Humans' unique turn-taking skills. *Psychological Science, 27*(7), 987-996. doi:10.1177/0956797616644070

Melkonyan, T., Zeitoun, H., & Chater, N. (2018). Collusion in Bertrand versus Cournot competition: A virtual bargaining approach. *Management Science, 64*(12), 5599-5609.

Michael, J., Sebanz, N., & Knoblich, G. (2016). The sense of commitment: A minimal approach. *Frontiers in Psychology, 6*(1968). doi:10.3389/fpsyg.2015.01968

Misyak, J. B., Melkonyan, T., Zeitoun, H., & Chater, N. (2014). Unwritten rules: Virtual bargaining underpins social interaction, culture, and society. *Trends in Cognitive Sciences, 18*(10), 512-519.

Nash, J. (1950). The bargaining problem. *Econometrica, 18*(2), 155-162.

Nesse, R. (2001). Evolution and the Capacity for Commitment. In. New York: NY: Russell Sage Foundation.

Nichols, S. (2004). *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford: Oxford University Press.

O'Neill, O. (2002). *A Question of Trust: The BBC Reith Lectures*. Cambridge, UK: Cambridge University Press.

Oskamp, S., & Perlman, D. (1965). Factors affecting cooperation in a Prisoner's Dilemma game. *Journal of Conflict Resolution, 9*(3), 359-374.

Pesquita, A., Whitwell, R. L., & Enns, J. T. (2018). Predictive joint-action model: A hierarchical predictive approach to human cooperation. *Psychonomic Bulletin & Review, 25*(5), 1751-1769. doi:10.3758/s13423-017-1393-6

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review, 83*(5), 1281-1302. Retrieved from <Go to ISI>://WOS:A1993ML61400013

Rakoczy, H., Brosche, N., Warneken, F., & Tomasello, M. (2009). Young children's understanding of the context-relativity of normative rules in conventional games. *British Journal of Developmental Psychology, 27*, 445-456. doi:10.1348/026151008x337752

Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology, 44*(3), 875-881. doi:10.1037/0012-1649.44.3.875

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's Dilemma: A study in conflict and cooperation*. Ann Arbor, MI: University of Michigan Press.
Rogerson, W. P. (1992). Contractual solutions to the hold-up problem. *Review of Economic Studies, 59*(4), 777-794. doi:10.2307/2297997
Rosenthal, R. W. (1981). Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory, 25*(1), 92-100. doi:10.1016/0022-0531(81)90018-1
Rousseau, D. M. (1995). *Psychological Contracts in Organizations: Understanding Written and Unwritten Agreements*. London: Sage Publications.
Rousseau, D. M., & Shperling, Z. (2003). Pieces of the action: Ownership and the changing employment relationship. *Academy of Management Review, 28*(4), 553-570. doi:10.5465/amr.2003.10899368
Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences, 10*(2), 70-76. doi:10.1016/j.tics.2005.12.009
Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory, 4*(1), 25-55.
Selten, R. (1978). Chain store paradox. *Theory and Decision, 9*(2), 127-159. doi:10.1007/bf00131770
Sen, A. K. (1977). Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy and Public Affairs, 6*(4), 317-344.
Sen, A. K. (2009). *The Idea of Justice*. Cambridge, MA: Harvard University Press.
Shafto, P., Goodman, N. D., & Griffiths, T. L. (2014). A rational account of pedagogical reasoning: Teaching by, and learning from, examples. *Cognitive Psychology, 71*, 55-89. doi:10.1016/j.cogpsych.2013.12.004
Sheridan, M., Sharma, A., & Cockerill, H. (2014). *From Birth to Five Years: Children's Developmental Progress. 4th edition*. New York: Routledge.
Sibly, H., & Tisdell, J. (2018). Cooperation and turn taking in finitely-repeated prisoners' dilemmas: An experimental analysis. *Journal of Economic Psychology, 64*, 49-56. doi:10.1016/j.joep.2017.11.003
Silk, J. B., Kaldor, E., & Boyd, R. (2000). Cheap talk when interests conflict. *Animal Behaviour, 59*(2), 423-432.
Skinner, B. F. (1971). *Beyond Freedom and Dignity*. New York, NY: Alfred A. Knopf.
Tomasello, M. (2020). The moral psychology of obligation. *Behavioral and Brain Sciences, 43*(e56), 1-58. doi:https://doi.org/10.1017/S0140525X19001742
Vlaev, I., & Chater, N. (2006). Game relativity: How context influences strategic decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(1), 131-149. doi:10.1037/0278-7393.32.1.131
Wang, R., Wu, S., Evans, J., Tenenbaum, J., Parkes, D., & Kleiman-Weiner, M. (2021). Too Many Cooks: Coordinating multi-agent collaboration through inverse planning. In S. Muggleton & N. Chater (Eds.), *Human-like Machine Intelligence* (pp. 152-170). Oxford, UK: Clarendon Press.

Table 1. Normal form of the Gas Station game

Player ***A***

Player ***M***		GFF	GFF^c	GF^cF	GF^cF^c	G^cFF	G^cFF^c	G^cF^cF	$G^cF^cF^c$
	Pay	$\binom{-8}{-8}$	$\binom{-8}{-8}$	$\binom{2}{2}$	$\binom{2}{2}$	$\binom{0}{0}$	$\binom{0}{0}$	$\binom{0}{0}$	$\binom{0}{0}$
	Not pay	$\binom{-2}{-14}$	$\binom{8}{-4}$	$\binom{-2}{-14}$	$\binom{8}{-4}$	$\binom{0}{0}$	$\binom{0}{0}$	$\binom{0}{0}$	$\binom{0}{0}$

Table 2. The worst payoffs for the Centipede game

		Player ***B***							
		ppp	ppt	ptp	ptt	tpp	tpt	ttp	ttt
	ppp	$\begin{pmatrix}32\\64\end{pmatrix}$	$\begin{pmatrix}32\\16\end{pmatrix}$	$\begin{pmatrix}8\\4\end{pmatrix}$	$\begin{pmatrix}8\\4\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$
	ppt	$\begin{pmatrix}8\\16\end{pmatrix}$	$\begin{pmatrix}8\\16\end{pmatrix}$	$\begin{pmatrix}8\\4\end{pmatrix}$	$\begin{pmatrix}8\\4\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$
Player ***A***	ptp	$\begin{pmatrix}2\\4\end{pmatrix}$	$\begin{pmatrix}2\\4\end{pmatrix}$	$\begin{pmatrix}2\\4\end{pmatrix}$	$\begin{pmatrix}2\\4\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$
	ptt	$\begin{pmatrix}2\\4\end{pmatrix}$	$\begin{pmatrix}2\\4\end{pmatrix}$	$\begin{pmatrix}2\\4\end{pmatrix}$	$\begin{pmatrix}2\\4\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$	$\begin{pmatrix}2\\1\end{pmatrix}$
	tpp	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$
	tpt	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$
	ttp	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$
	ttt	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$	$\begin{pmatrix}4\\1\end{pmatrix}$

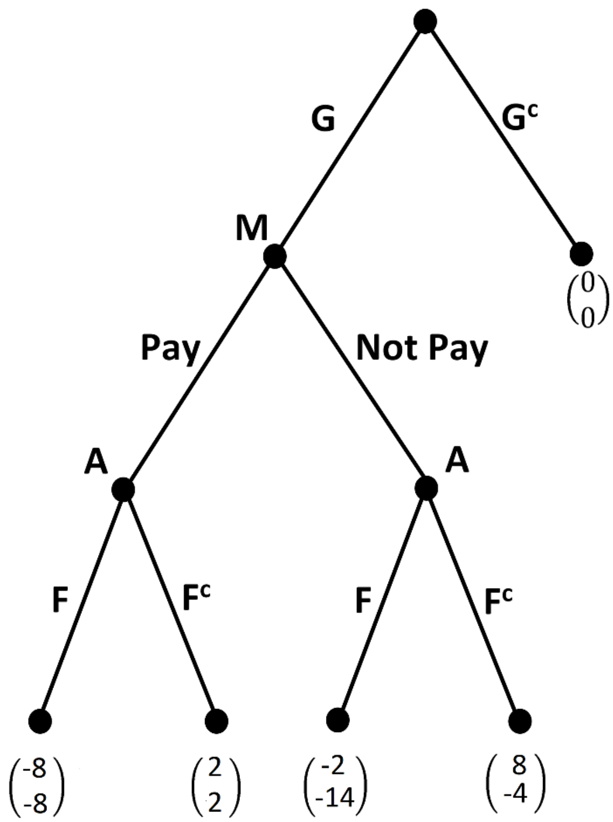
Table 3. A Prisoner's Dilemma game

		Player 2	
		C	*D*
Player 1	*c*	$(-1,-1)$	$(-10,0)$
	d	$(0,-10)$	$(-7,-7)$

Table 4. The modified Prisoner's Dilemma game

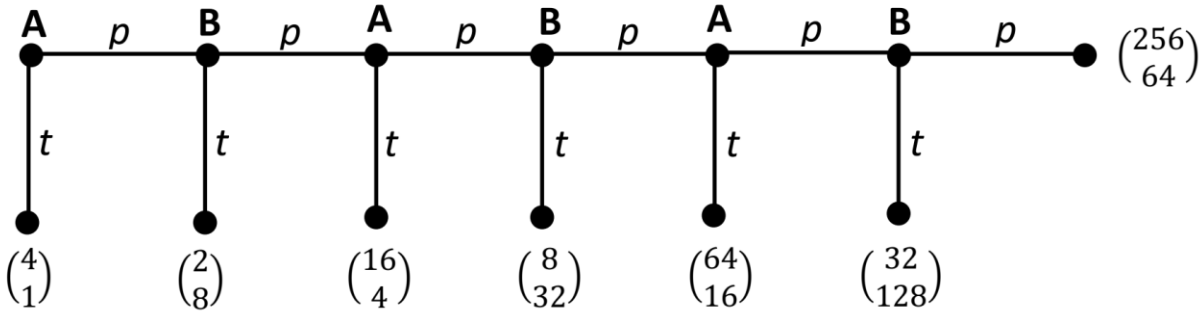
		Player 2	
		C	D
Player 1	c	$(-6,-6)$	$(-10,0)$
	D	$(0,-10)$	$(-7,-7)$

Figure 1. Gas Station game



An extensive-form game between players A and M. Player A moves first and chooses between G (provide gas) and G^c (refuse to provide gas). If player A chooses G^c, the game ends. If player A chooses G, then player M decides whether to pay for the gas or to refuse payment. Following player M's choice, player A chooses between F ("fight") and F^c (non-confrontational choice).

Figure 2. Centipede game



There are two piles of money on the table and two players, A and B, who move sequentially with player A moving first. The game has six stages. Initially, one of the piles contains \$1 while the other pile contains \$4. Every time it is a player's turn to choose, she can either take, "*t*", the larger of the two piles or pass, "*p*". If a player chooses "*t*", the game ends, with her getting the large pile and her opponent getting the small pile. If a player chooses "*p*", both piles are doubled and the play proceeds to the next player (unless the choice is made in the last sixth stage, where the game ends for both choices of player B).

APPENDIX

Formal Definition of VBE

Consider a game between n players who simultaneously and independently choose their strategies. Player i's strategy is denoted by $\sigma_i \in \Sigma_i$, where $i = 1, \dots, n$. Let $u_i(\sigma_i, \sigma_{-i})$ denote player i's ($i = 1, \dots, n$) payoff function, where the set of all opponents of player i is denoted by $-i$, the vector of their strategies is denoted by $\sigma_{-i} = \left(\sigma_{i_1}, \sigma_{i_2}, \dots, \sigma_{i_{n-1}}\right)$, and player i's k-th opponent ($k = 1, \dots, n-1$) is denoted by $i_k \in -i$.

We begin with the first stage of a player's reasoning to choose a strategy: finding the feasible agreements. The *worst payoff* of an agreement[23] $(\sigma_1^A, \dots, \sigma_n^A)$ for player $i = 1, \dots, n$ is defined as

$$(1)\quad w_i\left(\sigma_i^A, \sigma_{-i}^A\right) = \min\left\{u_i\left(\sigma_i^A, \sigma_{-i}^A\right), \sup_{\sigma_{i_{s_l}} \in R_{i_{s_l}}\left(\sigma_{-i_{s_l}}^A\right), l=k+1,\dots,n} u_i\left(\sigma_i^A, \sigma_{i_{s_1}}^A, \dots, \sigma_{i_{s_k}}^A, \sigma_{i_{s_{k+1}}}, \dots, \sigma_{i_{s_{n-1}}}\right) : k = 0, \dots, n-1, s_l \in -i \text{ for all } l = 1, \dots, n-1\right\},$$

where $R_{i_{s_l}}\left(\sigma_{-i_{s_l}}^A\right)$ denotes the set of player i_{s_l}'s best responses to strategy $\sigma_{-i_{s_l}}^A$. Thus, player i allows for the possibility that any subgroup of her opponents $\left(i_{s_{k+1}}, \dots, i_{s_{n-1}}\right)$ may deviate from the agreement σ^A in a "non-cooperative" fashion with each player in the "deviating subgroup" independently playing a best response to strategy σ^A. Note that player i also believes that when players $i_{s_{k+1}}, \dots, i_{s_{n-1}}$ in the deviating subgroup have multiple best responses, they will choose the best responses that yield player i the highest payoff. In other words, player i rules out the possibility that if some of her opponents deviate and play a best response to the agreement σ^A they will do so in a spiteful fashion.[24],[25]

[23] In the light of the model of reasoning envisioned in the present paper, agreement and strategy profile are used interchangeably.

[24] Thus, we assume that the players are "guarded pessimists." Each player is a pessimist because she is guided by the worst payoff of the different contingencies in expression (1) (i.e., the min operator in (1)). However, her pessimism is limited because she views a set of possible deviations by a group of her opponents in an optimistic fashion (i.e., the sup operator in (1)).

[25] We leave the examination of plausibility of different variations of the assumption to future research.

In summary, the worst payoff is equal to the least of a player's payoff under $1 + \sum_{k=1}^{n-1}\binom{n-1}{k} = 2^{n-1}$ possibilities, where $\binom{n-1}{k}$ denotes the number of k-combinations of an $(n-1)$-set. The first of these possibilities corresponds to the scenario where all of player i's opponents go through with their part of the agreement σ^A. There are also $n-1$ qualitatively different scenarios differentiated by the number of player i's opponents who choose to deviate and play a best response to the agreement. Each such scenario $k \in \{1, \dots, n-1\}$ has $\binom{n-1}{k}$ cases that correspond to different identities of the players in the "deviating subgroup."

It follows immediately from the definition in (1) that each player's worst payoff is bounded from above by her payoff. The two are equal to each other if either all of the opponents do not have beneficial deviations or all non-spiteful deviations by the opponents do not negatively affect the player's payoff.

In a game with two players, the worst payoff of an agreement (σ_1^A, σ_2^A) for player $i = 1,2$ is equal to

(2) $$w_i(\sigma_i^A, \sigma_{-i}^A) = \min\left\{u_i(\sigma_i^A, \sigma_{-i}^A), \sup_{\sigma_{-i} \in R_{-i}(\sigma_i^A)} u_i(\sigma_i^A, \sigma_{-i})\right\}.$$

In this case, the worst payoff is equal to the least of two possibilities: (i) player i's opponent goes through with her part of the agreement (σ_1^A, σ_2^A) and (ii) best-responds to σ_i^A. The use of the min operator can be "rationalized" by introducing ambiguity on behalf of the players about their opponents' strategies and assuming that the players have maximin expected utility preferences (Gilboa & Schmeidler, 1989).

Each player is guided by the worst payoff and discards all agreements for which she can improve her worst payoff via a unilateral deviation. The remaining agreements are called "feasible." Formally, an agreement σ^F is *feasible* if, for all $i \in \{1, \dots, n\}$,

(3) $$w_i(\sigma_i^F, \sigma_{-i}^F) \geq w_i(\tilde{\sigma}_i, \sigma_{-i}^F) \text{ for all } \tilde{\sigma}_i \in \Sigma_i.$$

Let $R_i^F(\sigma_{-i}) \equiv argmax_{\sigma_i \in \Sigma_i} w_i(\sigma_i, \sigma_{-i})$ denote player i's best-response correspondence for the worst payoff function. With this notation, agreement σ^F is feasible if and only if

(4) $$\sigma_i^F \in R_i^F(\sigma_{-i}) \text{ for all } i \in \{1, \dots, n\}.$$

We let F denote the set of feasible agreements. The definition in (4) reveals that an agreement is feasible if and only if it is a Nash equilibrium of the game where the payoff of each strategy profile has been replaced by the worst payoff of that profile.

In the second stage of the player's reasoning to select a strategy, each player simulates a bargaining process that, given her and the opponents' status quo positions, chooses one of the feasible agreements. The *status quo* position for each player in the bargaining process is given by her worst payoff from a feasible agreement that is worst for her. This assumption is consistent with our modeling of the players as decision-makers who view outcomes through the lens of a guarded pessimist. Formally, player i's fallback position is defined as $w_i^m = \inf_{\sigma^F \in F} w_i(\sigma^F)$.[26] In what follows, we will call w_i^m the minimum feasible worst payoff of player i.

The players choose from the set of feasible agreements according to the Nash bargaining solution where the players' status quo positions are their minimum feasible worst payoffs.[27,28] Formally, a *virtual bargaining equilibrium* (VBE) $\sigma^V = (\sigma_1^V, \ldots, \sigma_n^V)$ satisfies:

(5) $\sigma^V \in \underset{\sigma^F \in F}{\operatorname{argsup}} \prod_{i=1}^n (w_i(\sigma_i^F, \sigma_{-i}^F) - w_i^m)$.

Note that the optimization problem in (5) may have multiple solutions. There are multiple virtual bargaining equilibria in this case.

Proof of Proposition 1. (a) For illustrative purposes, consider the case where each player has an opportunity to move at most three times. The result and the proof hold for arbitrary fixed number of stages of the game. Table 2 contains the two players' worst payoffs for different combinations of pure strategies. It follows immediately from this that the set of feasible agreements consists of the strategy profile (ppp, ppp) as well as the set of Nash equilibria comprised of all possible probability distributions over the strategies tpp, tpt, ttp and ttt for each of the two players. Note that both players' payoffs are larger for the strategy profile (ppp, ppp) than for any Nash equilibrium of the game. Hence, from (5), (ppp, ppp) is the unique VBE of the game.

[26] For space considerations, we do not consider alternative assumptions for the fallback position. We note, however, that the predictions of our model of reasoning are robust for many sensible alternatives.

[27] Krishna and Serrano (1996) characterize a bargaining procedure akin to Rubinstein's alternating offer game for the general case of n players and demonstrate that when the players are sufficiently patient the equilibrium agreement in their setup approximates the n-player Nash bargaining solution.

[28] In what follows, we assume that the players have equal bargaining powers.

(b) Since all of the information sets of the game are singletons, the SVBE coincides with the unique subgame perfect equilibrium of the game. QED

Proof of Proposition 2. (a) We will say that a player plays the $(n-k)$-stage (for $k=0,..,n-1$) grim trigger strategy if (i) in the first $(n-k)$ stages, the player starts by cooperating, cooperates as long as the opponent always cooperated in the past, and defects for the rest of the game after her opponent defects and (ii) the player always defects in the last k stages of the game, irrespective of own and the other player's choices. When $k=0$, we will simply say that a player uses the grim trigger strategy.

Consider the case where a player, say Player 1, plays the grim trigger strategy. A best response[29] for Player 2 to Player 1's strategy is to cooperate in the first $(n-1)$ stages of the game and to defect in the last stage. Hence, the worst payoff to Player 1 from the strategy profile where each player follows the grim trigger strategy is equal to $(n-1)\cdot(-1)+(-10)=-(n+9)$. To verify that the strategy profile where each player uses the grim trigger strategy is a feasible agreement, consider a deviation by Player 1. A deviation by Player 1 will have her defecting instead of cooperating in at least one of the stages of the game. Let $k=1,\ldots,n$ denote the first stage that Player 1 defects. Given that the agreement stipulates that Player 2 choose the grim trigger strategy, Player 1's payoff from such deviation will be equal to $(k-1)\cdot(-1)+(n-k)(-7)$. When $k>2$, Player 2's best response is to defect in all stages from $(k-1)$ through n and to cooperate in stages 1 through $(k-2)$. When $k\leq 2$, Player 2's best response is to defect in all stages of the game. Player 1's payoff when Player 2 plays this best response is equal to

$$(9)\qquad \begin{cases}(k-2)\cdot(-1)+(-10)+(n-k+1)(-7), & if\ k>2\\ (n-1)(-7)+(-10), & if\ k=2.\\ n\cdot(-7), & if\ k=1\end{cases}$$

Each of the expressions is less than Player 1's worst payoff of $-(n+9)$ under the agreement where both players choose the grim trigger strategy. Thus, deviation in any stage of the game

[29] Due to a certain degree of freedom to specify actions off the "equilibrium path," there might be several best responses to a player's strategy. The behavior on the "equilibrium path" is, however, identical for all of these best responses.

yields a strictly lower worst payoff, which implies that the pair of the grim trigger strategies is a feasible agreement.

Consider now the unique subgame perfect equilibrium of the game where both players defect in all stages of the game irrespective of the history of the game. By the definition of feasible agreement, this strategy profile is also feasible. The worst payoff of both players for this strategy profile is equal to $-7n$.

Note that these two strategy profiles, the pair of the grim trigger strategies and the unique subgame perfect equilibrium, do not exhaust the set of feasible agreements. Consider, for example, the following variant of the grim trigger strategy. Suppose that, similarly to the grim trigger strategy, a player starts by cooperating and cooperates in all n stages of the game. In contrast to the grim trigger strategy, the new strategy doesn't prescribe defection for the rest of the game following a defection by the opponent. Rather, in the case of an early first deviation[30] by the opponent the player defects only for a number of periods that is sufficient to deter the deviation. For example, following the first deviation by the opponent, the player may cooperate once and defect in the remaining stages of the game. The worst payoff to either player when the players virtually agree that each will play this new strategy is equal to $-(n+9)$, which is the same worst payoff for the agreement consisting of the grim trigger strategies.

In addition, there are a number of equilibria that qualitatively differ from the above strategy profiles both on and off the equilibrium path. For example, an agreement where both players play the $(n-k)$-stage grim trigger strategy is feasible. It follows immediately from our derivations above that the worst payoffs to the two players in all feasible agreements, including the subgame perfect equilibrium, cannot strictly exceed the worst payoff for the agreement where both players follow the grim trigger strategy. Hence, by the definition of the VBE in (5) and because the players have equal bargaining powers, the latter strategy profile is a VBE. Moreover, any VBE of the game will entail cooperation by both players in all n stages of the game.[31]

(b) To determine the SVBE, first consider the last stage of the game. In the SVBE for any history of the game in this stage, each player will defect. Solving the game backwards, similarly to the

[30] By early deviations, we mean deviations from the path prescribed by the pair of the grim trigger strategies which occurs in a relatively early stage of the game.

[31] There are also asymmetric virtual bargaining equilibria where the players use different strategy profiles as well as equilibria where the players mix off the equilibrium path.

repeated Bertrand competition game, we find that the SVBE coincides with the unique subgame perfect equilibrium of the game where both players defect for all possible histories of the game. QED

Proof of Proposition 3. (a) Suppose that each player follows her part of the turn-taking agreement with punishments. Any best response for Player 2 to Player 1's strategy will entail sticking to the turn-taking sequence in the first $(n-1)$ stages of the game and defecting in the last stage. Hence, the worst payoff to Player 1 from the turn-taking agreement is equal to $\left(\frac{n}{2}-1\right)\cdot(-10)+(-10)+(-7)=-(5n+7)$.

Consider now the worst payoff to Player 2 from the turn-taking agreement with punishments. Given Player 2's strategy, Player 1 will not have any incentive to deviate from the agreement. Player 1 has no incentive to change her choice in the last stage. She also has no incentive to change her choice in the penultimate stage because she would get a continuation payoff of -14 from a deviation to "*d*" in the stage as opposed to a continuation payoff of -10 from sticking to turn-taking. Advancing this logic toward the top of the game tree, we find that Player 1 does not have any incentives to deviate in any of the stages of the game. Hence, Player 2's worst payoff from the turn-taking agreement is equal to $-5n$.

To verify that the pair of the turn taking strategies with punishments is a feasible agreement, consider a deviation by Player 1. Let $k=1,\dots,n$ denote the first stage when Player 1 defects. A deviation by Player 1 will have her defecting instead of cooperating if k is odd and cooperating instead of defecting if k is even. Given that the agreement stipulates that Player 2 defects for the remainder of the game after any deviation by player 1, Player 1's payoff from such deviation will be equal to $\left(\frac{k-1}{2}\right)\cdot(-10)+(n-k+1)\cdot(-7)=-7n+2(k-1)$ if k is odd and equal to $\left(\frac{k}{2}-1\right)\cdot(-10)-10-6+(n-k)\cdot(-7)=-7n+2k-6$ if k is even. Hence, if Player 1 deviates in any stage preceding the penultimate stage $(k=1,\dots,n-2)$ or in the last stage $(k=n)$, her payoff, and hence her worst payoff, will be smaller under the deviation than her worst payoff under the agreement where both players choose the turn taking with punishments strategy.

Finally, consider a deviation where Player 1 cooperates in stages 1 through $(n-2)$ but defects in stages $(n-1)$ and n. Player 2's best response to this strategy of Player 1 is to defect starting in stage $(n-3)$. Hence, Player 1's worst payoff for the agreement where she cooperates in stages 1 through $(n-2)$ but defects in stages $(n-1)$ and n while Player 2 plays the turn taking with punishments strategy is equal to $(n-4)\cdot(-5)+(-10)+(-7)+2\cdot(-7)=-5n-11$. This worst payoff is also lower than the worst payoff $-(5n+7)$ from the pair of the turn taking strategies with punishments. Thus, deviation in any stage of the game yields a strictly lower worst payoff for Player 1. A similar argument can be used to establish that, starting from the agreement comprised of the pair of the turn taking strategies with punishments, Player 2 cannot improve her worst payoff by deviating to an alternative agreement. Thus, the turn taking agreement is feasible.

Using derivations similar to those for the repeated Prisoner's Dilemma game, one can demonstrate that the worst payoffs for the two players in all feasible agreements cannot strictly exceed the worst payoffs for the turn taking agreement. Take, for example, the agreement consisting of the grim trigger strategies for both players. The worst payoff of both players under this agreement is equal to $(n-1)\cdot(-6)+(-10)=-(6n+4)$, which is lower than the worst payoff $-(5n+7)$ for the turn-taking agreement as long as the game has more than three stages. Hence, by the definition of the VBE in (5), the turn-taking agreement is a VBE because the players have equal bargaining powers. Moreover, any VBE of the game will involve alternation by both players in all n stages of the game.

(b) Similar to the corresponding part of Proposition 2. QED.

Collusion and SVBE. Consider a Bertrand game repeated $n \geq 3$ times as, for example, described in Melkonyan, Zeitoun and Chater (2018). There, we examine a one-shot Bertrand game and characterize its virtual bargaining equilibrium. To focus on the dynamic considerations, we adopt the same linear demand function setup as in their paper. In each stage of the game, two firms, $i=1,2,$ simultaneously and independently choose their prices $p_i \in \mathbb{R}_+$. The marginal costs of the firms are equal to zero while their direct demand functions are $q_1 = a - bp_1 + cp_2$ and $q_2 = a - bp_2 + cp_1$, where $b \geq c > 0$.[32]

[32] The results in this section hold for any specification of the demand and cost functions.

First, consider the one-shot interaction. Firm i's profit function is given by $\pi_i(p_i, p_{-i}) = (a - bp_i + cp_{-i})p_i$ while her best response to price p_{-i} is given by $R_i(p_{-i}) = \frac{a+cp_{-i}}{2b}$. The unique Nash equilibrium price is given by $p^N \equiv p_1^N = p_2^N = \frac{a}{2b-c}$. The price combination that maximizes the total of the two firms' profits, called the collusive price vector, is given by $p^* \equiv p_1^* = p_2^* = \frac{a}{2b-2c}$. From Melkonyan, Zeitoun and Chater (2018), the unique virtual bargaining equilibrium is given by $p^V \equiv p_1^V = p_2^V = \frac{a(2b+c)}{2(2b^2-c^2)}$. It follows from these expressions that the price under virtual bargaining, as well as the joint profits, fall strictly between their respective values under Nash and the collusive outcome (see Proposition 2 in Melkonyan, Zeitoun and Chater, 2018).

Consider now the finitely repeated version of the game. In the unique subgame perfect equilibrium of the game, the two firms will pick the Nash equilibrium price p^N for all possible histories of the game. We are left with characterizing the set of VBEs and SVBEs of the dynamic game.[33] It can be shown that the two firms will choose the collusive price vector p^* along the equilibrium path[34] in any VBE of the repeated Bertrand competition game. The firms can "enforce" this outcome by punishing deviations from p^* via sufficiently low prices for a sufficiently large number of periods. In the unique SVBE, the two firms choose the VBE of the static game p^V for each history of the game. Thus, the SVBE diverges from the conventional SPE—even when players re-think their behavior at each step, there are circumstances in which they can arrive at mutually beneficial outcomes not attainable through the standard SPE.

Thus, the firms "enforce" the collusive price vector by punishing the opponent. An example is the grim trigger strategy where the firm starts by choosing the collusive price p^* and continues with that price as long as both players always chose p^* in the past, and chooses Bertrand price p^N for the rest of the game after a deviation from p^*.

It also follows that the prices and joint profits under SVBE fall strictly between the prices and joint profits under subgame perfect and virtual bargaining equilibria. The repeated nature of the

[33] To minimize our notation, we abstract away from discounting here and in the other games.

[34] By an equilibrium path, we mean the collection of all information sets that are reached with strictly positive probability given the players' strategies. A deviation from the equilibrium path involves a different choice at an information set that is reached with a strictly positive probability. A deviation off the equilibrium path involves a different choice at an information set that is reached with zero probability.

game and the commitment capacity characterizing VBE of extensive form games open up possibilities for the two firms to achieve outcomes that are preferred by both firms to the corresponding SVBE. Moreover, in the repeated Bertrand game, the ability to commit, entailed in VBE, enables the two firms to achieve the outcome that is first-best from their perspective.