TESTING MODELS OF STRATEGIC UNCERTAINTY: EQUILIBRIUM SELECTION IN REPEATED GAMES

EMANUEL VESPA, TAYLOR WEIDMAN, AND ALISTAIR WILSON

ABSTRACT. In repeated-game applications where both the collusive and non-collusive outcomes can be supported as equilibria, researchers must resolve underlying selection questions if theory will be used to understand counterfactual policies. One guide to selection, based on clear theoretical underpinnings, has shown promise in predicting when collusive outcomes will emerge in controlled repeated-game experiments. In this paper we both expand upon and experimentally test this model of selection, and its underlying mechanism: strategic uncertainty. Adding an additional source of strategic uncertainty (the number of players) to the more-standard payoff sources, we stress test the model. Our results affirm the model as a tool for predicting when tacit collusion is likely/unlikely to be successful. Extending the analysis, we corroborate the mechanism of the model. When we remove strategic uncertainty through an explicit coordination device, the model no longer predicts the selected equilibrium.

1. INTRODUCTION

Assumptions about which of the possible equilibria best captures the participants' economic behavior take on a central role in applications with repeated interaction. For example, in models of oligopoly where firms interact repeatedly both collusive and non-collusive equilibria can arise. Researchers must often find evidence in the data to identify which equilibrium best captures the status quo. But equilibrium-selection problems persist when trying to predict counterfactual behavior, where appeals to the data are not possible. Consider a market where the researcher can establish that the current behavior of firms matches the non-cooperative equilibrium, but where their goal is to predict the effect from industry consolidation on consumer welfare. Embedded within the question is a concern on whether the selected equilibrium remains the same or will move under the counterfactual. Were a change isolated solely to the number of firms in the market, a directional comparative-static on selection could be posited: With fewer firms in the market, ceteris paribus, the equilibrium is likely to be pushed toward greater collusion.

But counterfactual analyses are often more involved than a simple comparative static. In our oligopoly example, while a lower number of firms might push behavior in the collusive direction, consolidation tends to be accompanied by other concurrent changes. For example, bigger firms can generate efficiency gains, increasing the marginal returns to price competition and product innovation. These accompanying shifts to the profit

1

functions could plausibly pull more towards the non-collusive equilibria.[1] As such, consolidation has the potential for conflicting effects on the likelihood of collusion. More broadly, whenever counterfactual policies involve richer changes to more than one variable, predicting the resulting effect on the selected equilibrium can become substantially more challenging.

In this paper we evaluate a model of equilibrium selection constructed over the theoretical primitives; a model capable of integrating counterfactual shifts to many primitives at the same time. In doing this, we build on experimental studies of the infinitely repeated prisoner's dilemma (RPD), where strategic uncertainty has been offered as an explanatory model for rationalizing several findings across this literature. Specifically, Dal Bó and Fréchette (2011) outline the organizational power of a selection notion from Harsanyi and Selten (1988) when applied to data from their experimental treatments.[2] The central idea here is to formally model the strategic uncertainty surrounding the collusive, history-dependent, equilibrium. In so doing, the analyst can reduce the equilibrium-selection problem to an understanding of the risk/reward trade-off over collusion attempts. The developed strategic-uncertainty measure is therefore a pure function of the primitives of the game. Thus, the model can help understand the trade-offs between two or more competing shifts of the primitives—for example, a change to the payoffs that increases the rewards to successful collusion combined with a reduction to the discount rate that makes such an attempt riskier. The model mediates the effects of any multivariate change into a single dimension, the effect on strategic uncertainty, which can then be used to predict the eventual likelihood of collusion. However, while the literature has shown the explanatory power of the strategic-uncertainty measure, previous studies have not been designed to test it, to examine the mechanism, and its potential limitations.

To stress test this equilibrium-selection model we expand the set of primitives under consideration. In so doing we introduce a new source of strategic uncertainty to the repeated environment, one that plays an important role in applications, as motivated by our industry consolidation example above. Similar to the original two-player RPD experiments, our treatments will manipulate payoff primitives (primarily a stage-game payoff-parameter x) that affects the riskiness of collusion. But we also manipulate the number of players within the environment (N), where collusion in the model becomes riskier in larger groups. Hence, we can study joint changes in both x and N—changes that push the equilibrium selection in two different directions, where the model's ability to predict the joint effect is the object of interest. As such, we will test the main hypothesis that equilibrium-selection across two distinct sources can be understood through the strategic-uncertainty model. We therefore construct treatment manipulations that modify x and N relative to a baseline in a way that the strategic-uncertainty measure is held exactly constant. Finding constant cooperation across the double-barreled change would

---

[1]The industry-consolidation literature provides ample evidence that mergers are regularly accompanied by other changes in addition to a lower number of market participants. For instance, accompanying effects include lower costs (Farrell and Shapiro, 1990), higher capacities (Perry and Porter, 1985), new product offerings (Deneckere and Davidson, 1985), and asymmetry among firms (Compte, Jenny, and Rey, 2002).
[2]A subsequent meta-study across the literature (Dal Bó and Fréchette, 2018) confirmed the pattern.

be consistent with the modifications to x and N being perfect substitutes, per their modeled effect on strategic uncertainty.[3] Complementing the construction and test of the core strategic-uncertainty model, our design also embeds an alternative hypothesis (which we characterize the behavioral underpinnings of) in which the group-size variable N does not affect strategic uncertainty, with a clearly distinct set of predictions from our main model.

Our main results allow us to clearly separate and validate the equilibrium-selection model based on strategic uncertainty. Across the six possible treatment comparative-statics in our design, four are predicted to have clear directional effects on selection (with the other two being perfect substitutes). In every single one of these four comparisons, data from our experiments move in the predicted direction, with effects sizes that are both statistically and economically significant. This is true for both initial and ongoing behavior for the repeated games we examine. The more-challenging comparisons involve treatments with compound shifts that move selection in opposed directions. Two of our treatment comparisons are constructed so that changes across the payoff and group-size are perfectly substituting as captured by the strategic-uncertainty model. Relative to successful coordination on the collusive outcome, the metric that likely matters most for applications, our results show strong support for the selection model, and our designed shifts to the payoffs and number of players are direct substitutes. Meanwhile, we document discrepancies between the measure's predictions and behavior in the first period of interaction. As such, our results indicate that the extended selection model is more accurate in predicting ongoing collusion than in predicting initial intentions—where changes in N provide the key variation to identify this. Overall, our results validate the theoretical model, providing a tool to understand which effects might dominate in counterfactuals where policy shifts over multiple primitives. In particular, the results lend credence to the idea that a reduction of the environment to a single dimension capturing strategic uncertainty can be predictive of equilibrium selection.

Uncertainty regarding the other players' actions is put front-and-center in our main treatments as the driver of equilibrium selection. Hence, if strategic uncertainty is the causal mechanism driving selection, then removing or reducing the doubts about others' play should make the selection model's predictions moot. We pursue this idea in an additional set of treatments where we allow for pre-play communication in a setting where we would otherwise expect little collusion. Participants are given the opportunity to exchange free-form messages prior to the repeated game, a feature that can be used to reduce uncertainty over others' intentions (see Kartal and Müller, 2018). For the same experimental parameterization where we observe ongoing cooperation rates below one percent without communication, the effect from adding it is to shift behavior to the other extreme: with initial (ongoing) cooperation rates of 95 (80) percent.[4] Given these results,

---

[3]Our treatments also include manipulations where the strategic-uncertainty effects of one of the two variable changes theoretically dominates the competing effect, with a predicted effect on selection. Moreover, the design also embeds comparative-static comparisons where a single variable is manipulated.

[4]As a check that the provided communication channel is not driving results separate from the equilibrium coordination effects, we implement a second treatment with communication but where the collusive outcome is not a robust equilibrium outcome. Here we find that communication does not lead to successful

we conclude that strategic uncertainty, the mechanism underlying the model, is the causal channel. In addition to mechanism validation, these results clearly outline the limitation of the model to understanding *tacit* collusion, in the sense that it fails to provide useful guidance when collusion is more explicit.

In terms of the paper's organization, we next outline the related literature. In section 2 we outline our generalization of the selection model from the two-player repeated PD game, outline the predictions from that literature, and provide design details. In section 4 we analyze the results from our core treatments and demonstrate the model's predictive properties generalize to N. Section 5 then examines some extensions: one showing that our results hold for both between- and within-subject identification; and the other pinning down the strategic uncertainty mechanism via the opportunity for explicit collusion. Finally, in section 6 we conclude.

1.1. **Literature.** Our paper builds on the recent consolidation of the experimental RPD literature presented in Dal Bó and Fréchette (2018). While one of our baseline treatments replicates a standard finding in the literature,[5] our core contribution is to generalize the equilibrium selection model outlined in the meta-study, adding an additional source of strategic uncertainty: the number of players, N.[6] Where the literature has developed this model to be explanatory, our approach is both to expand the selection model to a new setting, but also to take the model as the core experimental object to be tested.

Our generalization of the strategic uncertainty model in the RPD literature is carried out in two alternative ways. The first extension (and most-standard, given its use of independent beliefs) formalizes a distinct source of strategic uncertainty from the payoff-based source in the meta-study. An alternative extension (based on fully correlated beliefs) reflects a null effect, that the newly introduced source has no effect. As such, our generalization offers a potentially profitable design approach for future research examining other channels for strategic uncertainty effects—asymmetries in the action space or payoffs, the effects of sequentiality, etc.[7]

Our environment also allows us to better distinguish between the empirical measures linked to the selection model. That is, using literature-level data assembled by Dal Bó and Fréchette (2018), we show that the two-player RPD strategic uncertainty model predicts both initial and ongoing cooperation well. However, neither is identified particularly well from the other. With more than two players though, this is no longer the case, where we

---

collusion. In other words, communication only has an effect when there are clear theoretical incentives for collusion.

[5]As pointed out in Berry, Coffman, Hanley, Gihleb, and Wilson (2017) many experimental replications become harder to find where the papers do not explicitly point out the replication components.

[6]The extension of the notion of equilibrium selection described in Harsanyi and Selten (1988) to the RPD was first proposed by Blonski and Spagnolo (2015) and first shown to organize data by Dal Bó and Fréchette (2011). See also Fudenberg, Rand, and Dreber (2010) for an examination of the effects with imperfect monitoring, and Kartal and Müller (2018) for a test of a selection theory based on individual heterogeneity in preferences over dynamic strategies.

[7]See Ghidoni and Suetens (forthcoming) and Kartal and Müller (2018) for experimental examinations the effect sequentiality has in RPD settings through a reduction in strategic uncertainty.

use this to demonstrate that the developed strategic uncertainty model is better suited for predicting successful collusion rather than initial intentions.[8]

This paper is part of a broader literature seeking to understand and document regularities in equilibrium selection. In particular, regularities that are amenable to theoretical modeling. To this end, the strategic-uncertainty measure that we examine is particularly promising, as the equilibrium objects required for calculation are computationally simple: the stationary non-collusive equilibrium and the history-dependent collusive equilibrium. In environments beyond the RPD where the equilibrium outcomes are held constant, the model can be similarly extended per our illustration with a move to N players. However, in more complex environments where the equilibrium set changes more substantively, the constraint to two focal equilibria in the strategic uncertainty model may lose validity, and/or raise questions as to which two strategies are focal. Examples of more-complex settings includes dynamic games, where the stage environment can change across the supergame and the space of strategies becomes substantially larger. Vespa and Wilson (2020) focus on a horse-race examination of which two equilibria are focal (from a wider set of possible alternatives) for rationalizing behavior in dynamic games. The paper identifies a similar strategic uncertainty measure constructed around the most-efficient Markov perfect equilibrium and the best *symmetric* collusive equilibrium. A strategic-uncertainty model based on these strategies predicts behavior, where these strategies dovetail with the repeated-game strategies in the simpler environment studied here.[9]

An experimental literature on behavior in oligopolies documents that collusion clearly responds to the number of players. Both Cournot (Huck, Normann, and Oechssler, 2004; Horstmann, Krämer, and Schnurr, 2018) and Bertrand settings (Dufwenberg and Gneezy, 2000) indicate that as the number of players increases collusion becomes less likely, often as soon as N exceeds two.[10] We contribute to this literature on two margins. First, papers here focus on out-of-equilibrium behavior in settings with a finite-time horizon, and a subsequently unique theoretical prediction. In contrast, we examine how changes to N affect outcomes in an infinite-horizon environment with both collusive and non-collusive equilibria. Second, and crucially, our focus is not just on the qualitative directional effects from N, but on validating a model of how it affects strategic uncertainty. The model allows for an understanding of the extent of substitutability between primitives under consideration; and where validated can predict the directional effects of more-nuanced, multi-dimensional counterfactuals. Clearly, any equilibrium-selection notion suggested for such a task requires a great-deal of scrutiny. However, our findings do suggest substantial optimism that this path may be fruitful.

---

[8]Ongoing cooperation is a measure that is likely to be more relevant for empirical applications where collusion may be a worry. For instance, from Harrington, Gonzalez, and Kujal (2016), page 256: "(...) collusion is more than high prices, it is a mutual understanding among firms to coordinate their behavior. (...) Firms may periodically raise price in order to attempt to coordinate a move to a collusive equilibrium, but never succeed in doing so; high average prices are then the product of failed attempts to collude.

[9]Work on equilibrium selection in dynamic games builds on recent contributions in this area. For example: Battaglini, Nunnari, and Palfrey (2012, 2016); Agranov, Frechette, Palfrey, and Vespa (2016); Kloosterman (2019); Vespa and Wilson (2019); Rosokha and Wei (2020); Salz and Vespa (2020); Vespa (2020).

[10]See also references in Potters and Suetens (2013) for similar findings.

Our work is also related to an experimental literature on mergers that manipulates the number of players. As surveyed by Goette and Schmutzler (2009), some experiments deal with a "pseudo-mergers," where a subset of the original firms remain in the market (Huck, Konrad, Müller, and Normann, 2007, for example). Others experiments implement "real mergers," where mergers introduce other changes in the market beyond N (Davis, 2002). A contribution of our approach is that the strategic-uncertainty measure can predict counterfactual behavior in both settings. Another discussion in this literature is whether merger effects are evaluated within the same group of participants (within-subject designs) or across different groups (between-subject design). In this paper, while our main treatments make use of between-subject identification, we also conduct within-subject sessions at the same parameterization, demonstrating that though there can be meaningful short-run differences, with enough experience the results do not depend on the approach.[11]

The effects of communication devices as a bolster for collusion are well-established in the experimental literature. As surveyed in Cason (2008) and Harrington, Gonzalez, and Kujal (2013), more-structured, limited forms of communication usually result in small, temporary collusive gains, where free-form communication generates large, long-lasting effects.[12] For these reasons, in one of our extensions we examine unrestricted chat messages as a strong coordination device. The collusive results here clearly indicate that the domain for our strategic-uncertainty measure based on tacit-collusion does not include environments where explicit collusion is likely. However, we do show that there are clear limits on the effects of explicit collusion, and that these limits are directly predicted by theory. Using a change to the payoff primitives (here the discount rate) we make collusion a knife-edge, non-robust equilibrium, and show that the effects of communication are entirely dissipated.

While much of the literature on repeated games studies the standard two-player PD game, there is a large literature studying a canonical N-player social dilemma: the voluntary contribution public-goods game (see Vesterlund, 2016, for a survey). Though much of this literature focuses on finite implementations, one notable exception is Lugovskyy, Puzzello, Sorensen, Walker, and Williams (2017). Similar to our paper, the authors use experimental variation over both N and the payoff primitives (in their case the return to group contribution). However, this is done with a different end goal. They design off-setting variation in the payoffs and N to identify the isolated effect of the stage-game's MPCR. We instead do so to isolate strategic uncertainty, to test a predictive theory of selection.

---

[11] Differences in behavior can be stickier if changes are small or introduced gradually. Weber (2006) shows that increasing the number of players gradually in a coordination game leads to different results relative to a situation when play starts with a large group. In repeated games gradual introduction of changes to the payoff primitives have also been found to have effects in repeated games, see Kartal, Müller, and Tremewan (2017). These results suggest that the selection-notions we are examining are relevant for 'large' counterfactual changes, and where future research can help clarify how to integrate 'large' into a predictive model of selection.

[12] For further details on the effect of communication in repeated games with an unknown time horizon, see Fonseca and Normann (2012), Cooper and Kühn (2014), Harrington et al. (2016), Wilson and Vespa (2020) and Fréchette, Lizzeri, and Vespa (2020).

Beyond social dilemmas, our paper is also connected to the literature on coordination games (see Devetag and Ortmann, 2007, for a survey). The strategic-uncertainty measure examined in our paper works because the infinitely repeated prisoner's dilemma game has a stag-hunt normal-form representation (Blonski and Spagnolo, 2015), adapting the risk-dominance notion for one-shot coordination games in Harsanyi and Selten (1988).[13] Risk dominance (and the cardinal implementation through the measure of strategic uncertainty) has been shown to have substantial predictive content in simple coordination games with trade-offs over payoff-dominance and risk-dominance (see Battalio, Samuelson, and Van Huyck, 2001; Brandts and Cooper, 2006, and references therein). Strategic uncertainty has therefore demonstrated its usefulness as a theoretical selection device both in directly presented one-shot games, and in repeated games. Our contribution to this literature is to design an experiment to explicitly test it and show that the predictive effects extend further still, to multi-player infinite-horizon settings.

2. Generalizing the Basin of Attraction

Developing empirical criteria for equilibrium selection in games where collusion is possible requires two measures: one theoretical, one empirical. On the theory side we need a prediction, a model that maps the primitives of the game into a scalar where upward/downward movements are clearly interpretable as increasing/decreasing the likelihood of collusion. On the empirical side we need a precise target against which the theoretical notion can be contrasted and validated. This empirical measure should examine a behavior that differs starkly under the collusive and non-collusive equilibria.

We begin this section by summarizing the progress made towards validating a theoretical prediction in the two-player infinitely repeated prisoner's dilemma (RPD) literature. The theoretical notion here is the size of the basin of attraction for the strategy *always defect*. The focal outcome measure is the cooperation rate of individual players in the initial rounds. In what follows we discuss the issues as we extend the exercise to a new source of strategic uncertainty, the number of players N—both for the theoretical and empirical measures.

2.1. **Two-players.** Consider an RPD with discount rate $\delta \in (0,1)$. In each period $t = 1, 2, \dots$ players $i \in \{1, 2\}$ simultaneously select actions $a_i \in \mathcal{A} := \{(C)\text{ooperate}, (D)\text{efect}\}$. The period-payoff for player i is a function of both players' choices, $\pi_i(a_i, a_j)$, where all symmetric PD games can be expressed in a compact form by normalizing all payoffs relative to the joint-defection payoff $\pi_0 = \pi(D, D)$, and rescaling with the relative gain from joint cooperation: $\Delta\pi := \pi_i(C, C) - \pi_0$.[14] Defining scale and normalization in this way, the PD stage-game can be expressed with two parameters g and s for the different-action payoffs $\pi_i(D, C) = \pi_0 + (1+g)\Delta\pi$ and $\pi_i(C, D) = \pi_0 - s \cdot \Delta\pi$. The parameters $g > 0$ and $s > 0$ thereby capture the relative temptation- and sucker-payoff parameters, respectively.

---

[13]The difference in our setting is that neither total payoffs nor strategic choices are directly provided to participants, as these are extensive-form objects. Instead they are provided with the stage-game payoffs/actions, where strategies (such as the grim trigger or tit-for-tat) and gross payoffs must be endogenously formulated.

[14]More exactly, game payoffs π can be transformed via $\tilde{\pi} = (\pi_i - \pi_0)/\Delta\pi$ to measure all payoffs relative to joint defection in units of the optimization premium.

The RPD payoffs can be used as primitive inputs into a risk/reward model of collusion based upon strategic uncertainty. Strategic uncertainty here is distilled into a decision between two focal extensive-form RPD strategies:

    (i) *Always defect*, $\alpha_{\text{All-D}}$, which plays the stage-game Nash in all rounds (the unique MPE of the game).[15]

    (ii) The *Grim Trigger*, α_{Grim}, a strategy that begins by cooperating but switches to always-defect after observing any defections in past play (the best-case collusive SPE).[16]

As functions of the observable history h_t, these two strategies are given by:

$$\alpha_{\text{Grim}}(h_t) = \begin{cases} C & \text{if } t = 1 \text{ or } h_t = ((C,C),(C,C),\ldots,(C,C)), \\ D & \text{otherwise;} \end{cases}$$

$$\alpha_{\text{All-D}}(h_t) = D.$$

Strategic uncertainty in the two player RPD is measured through the *size of the basin of attraction for always defect*. The model considers the expected payoff to player i when uncertainty on the other player j is represented by a believed strategy mixture $p \cdot \alpha_{\text{Grim}} \oplus (1-p) \cdot \alpha_{\text{All-D}}$. The basin for always defect is defined as the set of beliefs p for which player i yields a higher expected payment from $\alpha_{\text{All-D}}$ than α_{Grim}. The always-defect belief basin is therefore the interval $[0, p^\star(g,s,\delta)]$, where the critical-point/interval-width is given by:[17]

$$(1) \qquad\qquad\qquad p^\star(g,s,\delta) \equiv \frac{(1-\delta) \cdot s}{\delta - (1-\delta) \cdot (g-s)}.$$

The basin-size $p^\star(g,s,\delta)$ has a clear interpretation for strategic uncertainty: for any belief $p > p^\star(g,s,\delta)$ on the other player using the collusive strategy α_{Grim}, the player does strictly better choosing α_{Grim}; for any belief $p < p^\star(g,s,\delta)$, they do strictly better by selecting the non-collusive strategy $\alpha_{\text{All-D}}$. As such, the smaller $p^\star$ is, the lower the strategic uncertainty surrounding collusion. Moreover, the cardinal basin-size measure directly implies the ordinal risk-dominance relationship between the two strategies. If $p^\star(g,s,\delta) < 1/2$ then the collusive strategy α_{Grim} risk-dominates $\alpha_{\text{All-D}}$, and vice versa. Henceforth, by 'basin of $p^\star$' we mean the maximal belief on the other cooperating for which the strategy $\alpha_{\text{All-D}}$ is optimal.

Equation (1) represents an easy-to-derive theoretical relationship between the payoff primitives of the game (here g, s and δ) and a critical strategic belief on the other player's likelihood of collusion. The hypothesized relationship is monotone, where the higher is $p^\star$, the

---

[15]A Markov strategy is history independent, removing any conditioning on past play. A Markov-perfect equilibrium (MPE) is a subgame-perfect equilibrium (SPE) in which agents use Markov strategies. In an RPD, if choices are forced to be history independent then there is a unique equilibrium: playing the stage-game Nash equilibrium in all periods.

[16]The strategy here is 'best case' in three senses: (i) it can support the best-case outcome; (ii) it uses the harshest possible punishment, and so can support collusion at smaller values of δ than any other strategy; and (iii) any realized miscoordination is minimal, being resolved in a single round.

[17]In the case that the strategy $(\alpha_{\text{Grim}}, \alpha_{\text{Grim}})$ is not an SPE of the repeated game, the basin size for always defects is defined as $p^\star(x_T, x_S, \delta) = 1$.
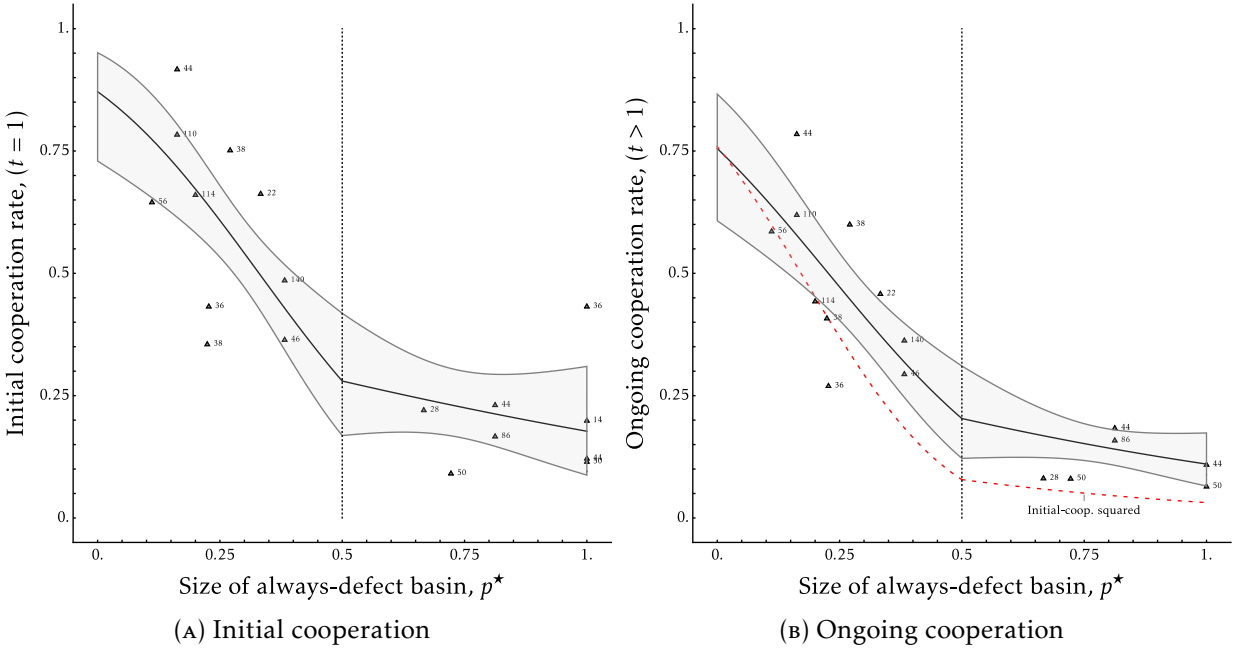
Figure 1. Meta-study relationship: Strategic Uncertainty and RPD Cooperation

Note: Figures show estimated effects and 95-percent confidence intervals for initial/ongoing cooperation in RPD meta-study (Dal Bó and Fréchette, 2018). Red dashed line in Panel (B) indicates implied theoretical relationship for ongoing cooperation based on the estimated relationship illustrated in Panel (A). Each point indicates a separate treatment, with the number of participants indicated.

lower is the likelihood of cooperation, thereby allowing for unambiguous directional predictions for any counterfactual change in the primitives. The posited mechanism within this model is also clear-cut: strategic uncertainty introduces a risk/reward trade-off for collusion attempts, which can be solved using standard economic analysis. This has two benefits: First, we can test the underlying strategic-uncertainty mechanism through other channels outside of the model, where we will do exactly that in an extension examining coordination devices. Second, the necessary assumptions for extending this model to analyze alternative strategic-uncertainty sources are straightforward.

We now turn to the empirical criteria used to validate this theoretical measure. As a summary of the RPD literature we focus on the recent meta-study on the two player RPD (Dal Bó and Fréchette, 2018). Two of the main results here are to show that the scalar basin-size measure of strategic uncertainty is highly predictive of behavior, though with a non-linear relationship. We illustrate this relationship in Figure 1 for two empirical outcome measures.

In both panels in Figure 1 the theoretical measure of strategic uncertainty (the basin-size $p^\star$) is presented on the horizontal axis. The empirical outcome measures are presented on the vertical axes. In Panel (A) we present the results for *initial cooperation* in the supergame, a measure tracking collusive *intentions* at the individual level as there is no interaction with matched partners at this point. In contrast, in Panel (B) we present

results for *ongoing cooperation*, choices in rounds two and beyond, a measure of the extent to which collusion attempts are successful. While there is no mechanical relationship between initial and ongoing cooperation, the theoretical model used in the basin construction posits an indirect linkage between the two. If collusion functions through conditional-cooperation with grim-trigger punishments, then the expected ongoing cooperation rate is the probability that the players jointly cooperate in the first round: the initial cooperation rate all squared. Thus, the posited relationship between initial and ongoing cooperation in the two-player RPD is not separately identifiable through any of the underlying payoff primitives.

Using the meta-study data from 996 participants across 18 experimental treatments, we estimate the relationship between late-session cooperation (supergames 16–20, the data we focus on in our experiments) and the strategic-uncertainty basin-size measure. Per the meta-study, the fitted relationship is estimated using a piecewise–linear probit model.[18] The solid line in each figure indicates the predicted cooperation rate at each $p^\star$ from the probit estimates, where the shaded region indicates the 95 percent confidence interval for the prediction (clustering by treatment). For both initial and ongoing cooperation the same pattern is found: a consistently low cooperation level when always-defect is risk dominant ($p^\star > \frac{1}{2}$); and a significantly decreasing relationship with $p^\star$ when collusion is risk dominant ($p^\star < \frac{1}{2}$). As such, Figure 1(A) illustrates Results 3 and 4 from the meta-study.

Figure 1(B) indicates a similar qualitative relationship between strategic uncertainty and ongoing cooperation. Going beyond just the form though, layered on top of the estimated ongoing-cooperation relationship we also indicate the implied theoretical relationship with the initial-cooperation results. Using the square of the estimated *initial* cooperation rate given in Panel (A) we therefore calculate the predicted *joint* cooperation rate in round one. Under grim-trigger coordination, the initial joint-cooperation rate is exactly the ongoing cooperation rate, and so we illustrate this implied relationship in Panel (B) as the red dashed line. There are certainly some differences between the estimated relationship using actual ongoing cooperation decisions (black solid line) and the implied rate from initial cooperation (red dashed line). However, as the figure illustrates, these differences are mostly in levels for the region where always-defect is risk dominant.[19] Slopes and levels for $p^\star$ in the $[0, \frac{1}{2}]$ region where the strategic uncertainty input has a clearer empirical response are not significantly different. The implication from this is that with only two players we cannot separately identify the extent to which the strategic-uncertainty measure is predicting initial intentions and/or successful coordination. However, as we

---

[18]Individual-level cooperation decisions are the left-hand side variable, where the basin-size enters the right-hand side in a piecewise-linear fashion around the risk-dominance dividing point. The econometric specification is motivated by Dal Bó and Fréchette (2018, Table 4); however to enforce level-continuity in the estimated relationship we remove a degree of freedom from their specification that allowed for a discontinuity at $p^\star = 1/2$.

[19]In particular, we find significantly more ongoing cooperation when the basin measure is above $1/2$ than we might expect from the initial rates and the use of a grim-trigger. This is consistent with another finding in the literature, the frequent use of more-forgiving/lenient strategies than grim. Tit-for-tat, for instance, can generate substantial *ongoing* cooperation when matched with its variant that starts out defecting.

will show, adding more players provides a channel to identify between the initial and ongoing measures.[20]

2.2. **Extending to $N > 2$.** We now extend the strategic-uncertainty model above to an N-player environment. The core extension is intuitive: instead of considering one other player choosing a mixture $p \cdot \alpha_{\text{Grim}} \oplus (1-p) \cdot \alpha_{\text{All-D}}$, we consider $N-1$ other players choosing this same strategic mixture.

To outline the extension, and set-up our design, we consider a family of symmetric social dilemmas that nest the standard RPD when $N = 2$. However, to hold constant a 2×2 stage-game representation for all N, our family of dilemmas make use of an aggregate (and deterministic) signal of the other agents' actions. Each player $i = 1,\dots,N$ continues to make a binary action choice $a_i \in \mathcal{A} \equiv \{C, D\}$, but their payoffs do not vary with (and they do not receive feedback on) the separate actions of the other $N-1$ players. Instead, payoffs are determined by the own-action a_i and a deterministic binary signal $\sigma(a_{-i}) \in \{S(\text{uccess}), F(\text{ailure})\}$ of the others' actions, a_{-i}. In particular, the generic player i's stage-game payoff and signal function are given, respectively, by:

$$\pi_i\big(a_i, \sigma\big) = \begin{cases} \pi_0 + \Delta\pi & \text{if } a_i = C, \sigma = S, \\ \pi_0 + \Delta\pi \cdot (1 + x) & \text{if } a_i = D, \sigma = S, \\ \pi_0 - \Delta\pi \cdot x & \text{if } a_i = C, \sigma = F, \\ \pi_0 & \text{if } a_i = D, \sigma = F; \end{cases}$$

$$\sigma\big(a_{-i}\big) = \begin{cases} S & \text{if } a_j = C \text{ for all } j \neq i, \\ F & \text{otherwise.} \end{cases}$$

These choices lead to a symmetric game, where payoffs can be summarized with a 2×2 table over: (i) the own action C or D; and (ii) the signal outcome, an S signal if the other $N-1$ players jointly cooperate, or an F signal if at least one matched player defects. The exact payoffs (with the same implicit scale $\Delta\pi$ and normalization π_0 as before) implement a PD-like environment where payoff-based strategic uncertainty is collapsed to a single parameter x.[21]

The success/failure signal here is a deterministic function of the $N-1$ other player's actions that dovetails with the standard RPD game when $N = 2$. The choice for the signal function $\sigma(\cdot)$ here maximizes the coordinative pressure, duplicating a Bertrand-like tension: collusion is successful only when the other $N-1$ players all cooperate.[22]

---

[20]For any setting where collusion requires N agents to initially cooperate to produce ongoing cooperation then the relationship is simply initial cooperation rate to the N-th power. Separate identification between the two measures then comes about through comparison across treatments with different values of N.

[21]In the meta-study notation this is implemented with $s = g = x$. This single-parameter formulation is equivalent to the Fudenberg et al. (2010) benefit/cost formulation, where their benefit/cost ratio b/c parameter is given by $(1+x)/x$ here.

[22]While potentially interesting, one disadvantage of a Cournot-like parameterization where the positive externality from others' cooperation is strictly increasing in the number of cooperators is that the cardinality of the signal/history spaces will then scale with N. However, note that the precise basin-size formula is *identical* under the Cournot formulation if own cooperation has an additive cost $x \cdot \Delta\pi$ and the punishment path is triggered on *any* level of defection. An alternative choice for a signal $\sigma(\cdot)$ that does not scale

11

Ignoring the game's scale and normalization (held constant in our experiments with $\Delta\pi = \$9$ and $\pi_0 = \$11$) the repeated games we examine are summarized by three primitives: (i) The relative cost of cooperating, x. (ii) The number of players, N. (iii) The continuation probability, δ. Our experiments will fix $\delta = \frac{3}{4}$ in all but one diagnostic treatment in the extensions section. This leaves us with two key experimental parameters: the relative cost x (actual cost $X = x\Delta\pi$) and the number of participants N.

In building a model of strategic uncertainty for arbitrary N, we will use a symmetric belief over the others' choices. That is, we assume each player chooses a mixture $p \cdot \alpha_{\text{Grim}} \oplus (1-p) \cdot \alpha_{\text{All-D}}$ over the two strategies.[23] Our family of social dilemmas require cooperation from all N players for everyone to get a success signal. Because of this, the strategic uncertainty reduces to the probability that the other $N-1$ players *jointly* coordinate on the collusive strategy,

$$Q(N) = \Pr\{N-1 \text{ others all choose } \alpha_{\text{Grim}}\}.$$

In every other case a failure will be received by at least $N-1$ players, and the punishment path will be triggered.

In an identical manner to the two-player case, the critical belief $Q^\star(N)$ is given by the point of indifference between the amount given up with certainty from a single round of cooperation, $x \cdot \Delta\pi$, and the continuation gain from collusion, $\frac{\delta}{1-\delta} \cdot \Delta\pi$, obtained with probability $Q(N)$. Indifference is therefore given by:

$$Q^\star(N) = \frac{(1-\delta)}{\delta} x,$$

where the RHS is identical to the two-player construction in (1) for $x = g = s$.

To complete the extension we therefore need to relate the joint-cooperation of the other $N-1$ players to the probability p that each other player individually attempts to collude. However, even though we have specified the marginal belief distribution and assumed symmetry, we must still resolve the relationship between the joint and marginal distributions: the extent to which beliefs are correlated. In particular our design and extension will focus on two extreme models as we extend the model to N-players. The 'standard' extension, where beliefs are fully independent; and an alternative/null-effect model where beliefs are perfectly correlated.[24] Assuming perfect correlation for the other $N-1$ agents,

---

with N would be that successes are achieved whenever some minimal number of players cooperate (≥ 2 say). Notice though, that such a choice introduces a second coordination problem over who cooperates and who defects. Efficiency requires the minimal number of cooperators, with the rest free riding. Changing the signal function in this way would therefore introduce a distributional asymmetry (see Wilson and Vespa, 2020, for an illustration that these additional dimensions lead to non-trivial problems). Given our focus on understanding the effects from adding a new source of strategic uncertainty, we therefore chose an environment that did not introduce additional trade-offs/confounds.

[23]For the N-player dilemma we define the grim-trigger strategy with imperfect-signals as:

$$\alpha_{\text{Grim}}(h_t) = \begin{cases} C & \text{if } t = 1 \text{ or } h_t = ((C,S),(C,S),\ldots,(C,S)) \\ D & \text{otherwise.} \end{cases}$$

[24]See Cason, Sharma, and Vadovič (2020) for a clear example of correlated beliefs arising where independence would be the standard prediction.

joint and individual probabilities are identical, $Q(N) = p$, and so the extended critical belief (and correlated model outcome) is given by:

$$p^{\star}_{\text{Corr.}}(x) = \frac{1-\delta}{\delta} \cdot x. \tag{2}$$

In constrast, when the beliefs are independent we have $Q(N) = p^{N-1}$, and the critical belief (and independent model extension) is:

$$p^{\star}_{\text{Ind.}}(x, N) = \left(\frac{1-\delta}{\delta} \cdot x\right)^{\frac{1}{N-1}} \equiv \left(p^{\star}_{\text{Corr.}}(x)\right)^{\frac{1}{N-1}}. \tag{3}$$

Obviously, when $N = 2$ the basin measures in (2) and (3) are identical, dovetailing with the standard construction. However, for $N > 2$ the two measures of strategic uncertainty are distinct, where the standard model extension under independence has strategic uncertainty vary both through the payoff parameter x and the group-size N.

3. Experimental Design

The basin of attraction for always defect under independent beliefs serves as our measure of strategic uncertainty. Ceteris paribus, the greater the uncertainty on successful strategic coordination, the more likely the subject is to take refuge in the safer strategy—in the case of an RPD game, the stage-game Nash outcome of defecting.

As outlined above, as we extend the environments to $N > 2$, this opens up a question as to how strategic uncertainty is affected by N. If others' behavior is thought to be highly correlated, adding players but holding constant the payoffs will do little to affect the selected behavior in our games. If this is so, the null-effect $p^{\star}_{\text{Corr.}}$ measure of strategic-uncertainty will predict behavior. In contrast, in a more-standard extension where beliefs on the other players are independent, then $p^{\star}_{\text{Ind.}}$ will model the strategic uncertainty. Under this extension, we will be able to use shifts in $p^{\star}_{\text{Ind.}}$ to understand changes in the selected behavior. While one implication of this is that the addition of more players will ceteris paribus reduce collusion, a deeper implication of the model is to help us understand substitution effects across the two sources of strategic uncertainty, x and N.

Our experimental design attempts to untangle the effects of strategic uncertainty. The aim of the design is to embed comparative-static tests on the effect of N (and so rule out the $p^{\star}_{\text{Corr.}}$ null-effect model), but also to examine the possible substitution effects by constructing perfect-substitutes treatments with the $p^{\star}_{\text{Ind.}}$ model. We achieve this through a series of experimental implementations of the N-player 2-action–2-signal repeated game outlined above. In particular, the first part of our design will use this family of games to distinguish and separate between the two extremes of independence and perfect correlation, leveraging the theoretical relationships derived above.

While we cannot directly manipulate strategic uncertainty—as the basin-size measures are indirect, theoretical relationships derived from the primitives—equations (2) and (3) allow us to implicitly manipulate each measure through shifts in x and/or N. Increases in x increase the strategic uncertainty in both models: as higher costs of cooperation require a greater belief that the other(s) are cooperating. In contrast, increases to the number of

TABLE 1. Experimental Design

Panel A.	Stage-game payoffs	$X = \$9$		$X = \$1$	
		$\sigma(a_{-i}) = S$	$\sigma(a_{-i}) = F$	$\sigma(a_{-i}) = S$	$\sigma(a_{-i}) = F$
Coop., $\pi_i(C,\sigma)$		$20	$2	$20	$10
Defect, $\pi_i(D,\sigma)$		$29	$11	$21	$11

Panel B.	All-D Basin Size	$X = \$9$ $(x = 1)$		$X = \$1$ $(x = 1/9)$	
		$N = 2$	$N = 4$	$N = 4$	$N = 10$
Cor. basin, $p^{\star}_{\text{Cor.}}(x)$		$p^{\star}_0$ [0.33]	$p^{\star}_0$ [0.33]	$p^{\star}_0 - \Delta p^{\star}_{\text{Cor.}}$ [0.04]	$p^{\star}_0 - \Delta p^{\star}_{\text{Cor.}}$ [0.04]
Ind. basin, $p^{\star}_{\text{Ind.}}(x,N)$		$p^{\star}_0$ [0.33]	$p^{\star}_0 + \Delta p^{\star}_{\text{Ind.}}$ [0.69]	$p^{\star}_0$ [0.33]	$p^{\star}_0 + \Delta p^{\star}_{\text{Ind.}}$ [0.69]
Sessions		3	3	3	2
Subjects		60	60	72	60

Panel C.	Meta-study predictions	$p^{\star}_0$	Marginal effect from:	
		[0.33]	Basin increase to [0.69]	Basin decrease to [0.04]
Initial coop. $(t = 1)$		0.50	-0.26	$+0.35$
Ongoing coop. $(t > 1)$		0.37	-0.21	$+0.50$

Note: Meta-study prediction indicates the cooperation-rate estimation from the 2-player basin recovered from the treatment-clustered probits illustrated in Figure 1.

players N only increases strategic uncertainty for the independent basin-size measure, interacting with x in a non-linear manner.

Using equations (2) and (3), the two notions can be varied in isolation from one another. As such, it is possible to construct a 2×2 design that orthogonally varies each strategic-uncertainty measure. Next, we outline our design, summarized in Table 1.

Panel (A) of Table 1 illustrates our first treatment dimension, which manipulates the payoff cost of cooperating $X = x \cdot \Delta\pi$, where $\Delta\pi = \$9$. The two values of X—a high temptation of $9 (illustrated on the left, $x = 1$), and a low temptation of $1 (on the right, $x = 1/9$)—lead to two payoff environments over own-actions and the signals. The given stage-game payoffs capture what a participant would see on their screens.[25]

Our design also manipulates the number of players N, captured in the column headings of Panel (B) in Table 1. In total, we create four treatment environments, each defined by an (N, X)-pair, where we will refer to treatments with the directly varied parameter pair. In Panel (B), the rightmost four columns are headed by the chosen game variable, where the two rows of the table indicate how these choices affect the two basin-size measures of strategic uncertainty.

---

[25]See Figure A.3 in the Online Appendix for representative screenshots.

To manipulate each basin-size measure separately, our design takes $\binom{N=2}{X=\$9}$ as its starting point. The values for both the independent and correlated basin-size measures are the same: $p_0^\star = 0.33$. Holding the relative cooperation cost fixed and increasing the number of players to $N = 4$ does not affect the correlated measure in (2). However, a shift to $N = 4$ increases the independent basin measure to $p_0^\star + \Delta p_{\text{Ind.}}^\star = 0.69$.

Now consider the manipulation of X. Comparing $\binom{N=4}{X=\$1}$ to $\binom{N=2}{X=\$9}$, we hold constant the independent basin-size measure at $p_0^\star = 0.33$. The shifts in both X and N have perfectly substituting effects in equation (3). However, the same change in both variables under the correlated-basin measure has a substantial effect, where the change in N does not provide an offsetting effect. As such, the correlated-basin measure is lowered to $p_0^\star - \Delta p_{\text{Corr.}}^\star = 0.04$. Finally, in the $\binom{N=10}{X=\$1}$ treatment we complete the 2×2 design over the two basin-size measures. Comparing $\binom{N=4}{X=\$1}$ to $\binom{N=10}{X=\$1}$ holds constant the correlated basin, which does not depend on N, at $p_0^\star - \Delta p_{\text{Corr.}}^\star = 0.04$. But the increased number of players does increase strategic uncertainty in the independent basin. In particular, our parameterization matches the independent basin sizes for $\binom{N=10}{X=\$1}$ and $\binom{N=4}{X=\$9}$ at $p_0^\star + \Delta p_{\text{Ind.}}^\star = 0.69$.

Through variation in the primitives x and N, our design thereby generates four correlated/independent basin measure pairs with a 2×2 structure:[26]

$$\left(p_{\text{Corr.}}^\star, p_{\text{Ind.}}^\star\right) \in \left\{p_0^\star, p_0^\star - \Delta p_{\text{Corr.}}^\star\right\} \times \left\{p_0^\star, p_0^\star + \Delta p_{\text{Ind.}}^\star\right\} := \left\{0.33, 0.04\right\} \times \left\{0.33, 0.69\right\}.$$

The design therefore achieves the goal of orthogonal variation over the two basin measures. However, the parameterization was also chosen so that the shifts in each dimension would be expected to have quantitatively large effects. Through the Dal Bó and Fréchette (2018) meta-study we can generate level predictions for the behavioral effects of each change. Our design generates three basin-size measures: $p_0^\star = 0.33$, $p_0^\star - \Delta p_{\text{Corr.}}^\star = 0.04$, and $p_0^\star + \Delta p_{\text{Ind.}}^\star = 0.69$. Using the probit models illustrated in Figure 1 we leverage the meta-study to make *quantitative* predictions for the cooperation rates under each basin-size measure. These predictions are indicated in Panel (C) of Table 1.

Shifts in the basin-of-attraction under either model generate large changes in the predicted outcomes. The first column in Panel (C) indicates the baseline initial (ongoing) cooperation rate expected at $p^\star = 0.33$ using the meta-data estimated relationship is 50 percent (37 percent).[27] The next column pair then indicates the expected treatment effect for a shift in the strategic uncertainty level from $p^\star = 0.33$ to either of the two alternative

---

[26]We note that our choices of $\Delta \pi = \$9$ and $\delta = \sfrac{3}{4}$ were necessary choices here, as we aimed for exact-integer solutions for N and X for simplicity of the presentation. Our design over the basin measures is more-exactly given by:

$$\left(p_{\text{Corr.}}^\star, p_{\text{Ind.}}^\star\right) \in \left\{3^{-1}, 3^{-3}\right\} \times \left\{3^{-1}, 3^{-1/3}\right\}.$$

[27]All predictions based on late-session meta-study data (supergames 16–20), the same point in the sessions used to assess our hypotheses.

values. Panel (C) indicates that if either of the two strategic uncertainty measures captures and generalize behavior, then we would expect a large treatment effect on one of the two design dimensions.

We formalize the two competing hypotheses as:

Correlated-Basin/Null-effect Hypothesis. *Cooperation increases as we decrease X, but there is no effect as we vary the number of players N.*

Independent-Basin Hypothesis. *Cooperation decreases as we increase X and/or N. Moreover, the substitution effects between X and N indicate no effect on cooperation if we decrease X and increase N to hold constant the $p^\star_{Ind.}$ measure of strategic uncertainty.*

That is, consider the prediction if the standard independence-based extension of strategic uncertainty is the correct model. In treatments $\binom{N=2}{X=\$9}$ and $\binom{N=4}{X=\$1}$ the independent basin-size is 0.33, increasing to 0.69 in treatments $\binom{N=4}{X=\$9}$ and $\binom{N=10}{X=\$1}$. If the strategic uncertainty relationship estimated from the two-player RPD meta-data is perfectly extrapolated then we should expect: (i) A reduction of 26 (21) percentage points in initial (ongoing) cooperation across the treatment pair, caused by the increase in strategic uncertainty. (ii) A null effect on cooperation within each treatment pair, reflecting the designed perfect substitution across X and N.[28]

Notice that our hypotheses are silent with respect to which of the two outcome measures, initial and/or ongoing cooperation, we are supposed to match. The two measures have different interpretations—where initial cooperation captures intentions, and ongoing cooperation captures successful coordination. In the case of the two-player RPD, Figure 1 shows that the basin size tracks both cooperation measures relatively well, and that the effects are hard to disentangle. Through N, we are able to generate additional variation in the theoretical relationship between initial and ongoing cooperation variables that separates the two relationships on another dimension. An advantage of our design is that it will allow us to better identify if either of the two measures is better predicted by either basin-size measure.

Experimental Specifics. We use a between-subject design over the four distinct environments outlined in Table 1. Participants for each treatment were recruited from the undergraduate population at the University of Pittsburgh, participating in only one session. We recruited a total of 520 participants, with 252 for the first four treatments and 268 for the extensions that we outline in Section 5. Three sessions were conducted for each environment, with an aim to recruit at least 20 participants per session—the one

---

[28]Alternatively, under a null-effect from N, given by the correlated basin measure, the basin-size is reduced from 0.33 to 0.04 as we move between the $\binom{N=2}{X=\$9}$ and $\binom{N=4}{X=\$9}$ treatment pair and the $\binom{N=4}{X=\$1}$ and $\binom{N=10}{X=\$1}$ pair. The RPD prediction from the meta-study is then for an increase in the initial (ongoing) cooperation rate of 35 (50) percentage points (and again, a null effect within each pair).

exception to this was the $\binom{N=10}{X=\$1}$ treatment, where we ran two sessions of 30.[29] Sessions lasted between 55 and 90 minutes with participants receiving an average payment of $18.61.

Each session comprised 20 supergames, where supergames used a random termination chance of $1-\delta = \frac{1}{4}$ after each completed round. However, supergame-lengths are matched by session across treatment. Participants were randomly and anonymously matched together across the 20 supergames in a stranger design.[30] The 20 supergames were divided into two parts of ten supergames.[31] For final payment, one supergame from each part was randomly selected, where only the actions/signals from the last round in the selected supergame counted for payment.[32]

4. Results

In the first section we begin by describing the aggregate treatment-level cooperation rates. We postpone inferential tests of our two basin-extension hypotheses to the second section. The main finding here is that while neither extension contains all the relevant information for predicting initial cooperation, we do find more definitive results for ongoing cooperation. Overall, we conclude that the basin-size measure based on a standard independence assumption contains all relevant information for predicting successful coordination within the experimental supergames.

4.1. **Main Treatment Differences.** The top two lines of Panel (A) in Table 2 report cooperation rates broken out across the four treatments, where we separately report initial (the first round) and ongoing cooperation (all subsequent rounds). The averages in the table are for the last five supergames—late-session behavior, after subjects have amassed experience in the environment—though including all rounds generates similar results (see Table A.1 in the Online Appendix). The results point to large cooperation shifts across both the cost to cooperating X and the group size N.

The initial cooperation rate in our $\binom{N=2}{X=\$9}$ treatment is 50.3 percent, essentially identical to the 50 percent cooperation rate predicted by the PD meta-study. However, holding constant the cooperation cost at $X = \$9$ and doubling the group size to four virtually eliminates cooperative behavior, with just 3.5 percent initial cooperation in $\binom{N=4}{X=\$9}$. In

---

[29]In more detail, our design called for sessions to have at least 20 participants, but allowed us to recruit an additional group of size N depending on realized show ups. For $\binom{N=10}{X=\$1}$ we instead opted to recruit 30 for each session so that we had at least three groups in each supergame.

[30]All subjects received written and verbal instructions on the task and payoffs, where instructions are provided for readers in the Online Appendix.

[31]Subjects received full instructions for the first part and were told they would be given instructions on part two after completing supergame ten. For the four between-subject treatments outlined in section 3, part two was then identical to part one. Later in the paper we will outline a further set of treatments with a within-subject change across the parts. The design choice for two identical parts here allows for direct comparisons in first-half play.

[32]This method is developed in Sherstyuk, Tarui, and Saijo (2013) to induce risk neutrality over supergame lengths. Another benefit from this design choice is that there are no wealth effects within a supergame, and where history only matters as an instrument for others' future play.

Table 2. Cooperation rates and basin-effect decomposition

Panel A.	Action and signal rates	X = \$9			X = \$1	
		$N = 2$	$N = 4$		$N = 4$	$N = 10$
	Initial coop.	0.503 (0.058)	0.035 (0.017)		0.792 (0.042)	0.357 (0.055)
	Ongoing coop.	0.450 (0.055)	0.006 (0.003)		0.409 (0.050)	0.184 (0.048)
	Initial success	0.503	0.000		0.578	0.000
	Ongoing success	0.450	0.000		0.293	0.000

Panel B.	Cooperation decomposition	$p_0^\star$	Marginal effect from:	
			Ind. basin increase to	Corr. basin decrease to
		[0.33]	$p_0^\star + \Delta p_{\text{Ind.}}^\star = [0.69]$	$p_0^\star - \Delta p_{\text{Corr.}}^\star = [0.04]$
	Initial coop.	0.464 (0.058)	−0.395 (0.048)	+0.357 (0.053)
	Ongoing coop.	0.366 (0.051)	−0.293 (0.051)	+0.115 (0.061)

Note: Results are calculated using data from the last-five supergames. Cooperation rates present raw proportions (with subject-clustered standard errors). The cooperation decomposition runs two subject-clustered probits on the cooperation decision (initial and ongoing) where variables are dummies for a low correlated basin treatment ($X = \$1$, both N values) and a high-independent–basin treatment ($X = \$9/N = 4$ and $X = \$1/N = 10$). Coefficients shown are the predicted level at just the constant (the $p_0^\star$ column) and the predicted cooperation changes from each estimated dummy.

low-temptation settings ($X = \$1$), groups of $N = 4$ show highly cooperative behavior (79.2 percent for initial decisions), while groups of $N = 10$ generate moderate cooperation rates (35.7 percent). The second row of Panel (A) indicates the ongoing ($t > 1$) cooperation rate. Here the data indicates a decline in cooperation over the initial behavior in all treatments, though the quantitative effects are largest in the $\$X = 1$ treatments.[33]

The third and fourth rows of Table 2 present the fraction of participant rounds where a success signal was observed.[34] Focusing on success signals, similar patters emerge to ongoing cooperation, though with starker quantitative effects. While a success is the modal signal in the $\binom{N=2}{X=\$9}$ and $\binom{N=4}{X=\$1}$ treatments, in the $\binom{N=4}{X=\$9}$ and $\binom{N=10}{X=\$1}$ treatments we observe *no* successes at all.[35]

---

[33]In the Online Appendix, Table A.2 further breaks out ongoing cooperation by the observed history in the previous round. The results indicate clear evidence that individual cooperation is highly conditional on successful coordination. However, strategies are significantly more-forgiving after failed cooperation at $X = \$1$ than $X = \$9$.

[34]A success requires that the other $N − 1$ participants jointly cooperate. Success is a direct function of group-level cooperation, where the *expected* success rate with an independent cooperation rate q is q^{N-1}. In two-player games, the success rate is identical to the cooperation rate. For the initial round the expected success rates (in the Table 2 column order) are: 0.503, 4.2×10^{-5}, 0.497 and 9.5×10^{-5}.

[35]As success is a direct aggregate of individual level cooperation we do not report standard errors (where we also cannot calculate standard errors when there is no variation). However, the starkness of the effect with no successes when the independent basin-size is high make clear the underlying economic effects.

Using just the raw averages in the top panel of Table 2, the evidence clearly falsifies the correlated-basin/null hypothesis on the effect of N, for both initial and ongoing cooperation. The experimental results indicate large shifts in behavior as we move N as a comparative-static, fixing the value of X. Fixing $N = 4$ we do find the comparative-static effect predicted by the correlated-basin measure as we move X, though this directional effect is also predicted by the independent basin.

The evidence suggests that the independent-basin hypothesis fares slightly better. Both initial and ongoing cooperation clearly respond in the predicted direction as we shift either X or N in isolation. However, for initial cooperation, we do not find perfect substitution as we move both x and N, as predicted by the independent-basin hypothesis. The hypothesis predicts no change in cooperation rates comparing either $\binom{N=2}{X=\$9}$ to $\binom{N=4}{X=\$1}$ or $\binom{N=4}{X=\$9}$ to $\binom{N=10}{X=\$1}$. When this hypothesis is scrutinized using initial cooperation rates, we find large differences in either comparison.[36] To see this finding from a different perspective, let us pool treatments at each value of X. That is, we compare the average initial cooperation rate pooling $\binom{N=2}{X=\$9}$ and $\binom{N=4}{X=\$9}$ to the average pooling $\binom{N=4}{X=\$1}$ and $\binom{N=10}{X=\$1}$. Initial cooperation in the pooled treatments for $X = \$9$ and $X = \$1$ are 28.6 and 59.4 percent, respectively, and the difference is more than 30 percentage points. In other words, initial cooperation rates are inconsistent with the independent-basin hypothesis.

Results for ongoing cooperation though are substantially better for the independent basin, where $\binom{N=2}{X=\$9}$ and $\binom{N=4}{X=\$1}$ are not substantially different. We do note a difference between $\binom{N=4}{X=\$9}$ and $\binom{N=10}{X=\$1}$, this is primarily driven by the very stark finding of near-zero cooperation in $\binom{N=4}{X=\$9}$. We explore this further below. In the aggregate though, pooling the ongoing cooperation rates across X, we find similar rates at 22.8 percent for $X = \$9$ and 29.8 percent for $X = \$1$.

4.2. **Evaluation of Independent- and Correlated-Basin Hypotheses.** Panel (B) in Table 2 provides a direct statistical evaluation of our two competing hypotheses. The table reports results of a probit model that assesses subjects' cooperation decisions using dummy variables for the 2×2 design in Table 1 Panel (B). The dummy covariates are an indicator for the $\Delta p^{\star}_{\text{Cor.}}$ decrease in the correlated basin-size (as we decrease X), and an indicator for the $\Delta p^{\star}_{\text{Ind.}}$ increase in the independent basin-size (as we increase N within each X).

Each row in Panel (B) provides the results from a separate estimation, one over initial cooperation, one over ongoing. The first column reports the estimated cooperation rate when both dummy variables are zero: essentially the cooperation rate for a game with a basin size of $p^{\star}_0 = 0.33$. The final column-pair then report the estimated marginal effect on the cooperation rate for each basin shock, holding the other constant. If either of the two basin hypotheses fully explains behavior, we would expect a significant estimate for the corresponding dummy and an insignificant effect on the other.

---

[36]The differences are economically large: 29 percentage points in the first comparison and 35 percentage points in the second.
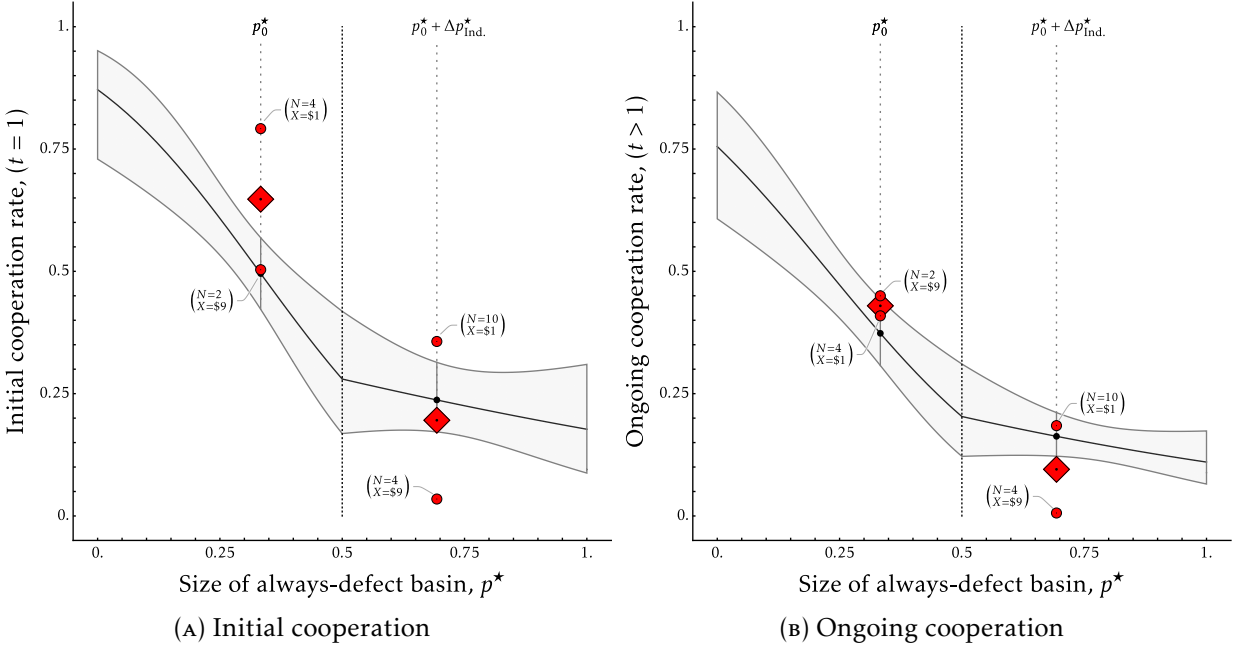
Figure 2. Cooperation and the Independent Basin-Size Model

Note: Figures show pooled data by the independent basin (diamonds) while the separate treatments are illustrated as the surrounding circles. See Figure A.1 in the appendix for an analogous figure for the correlated basin.

The estimation procedure here is designed to directly parallel the probit-model used to generate predictions from the meta-study. The estimated cooperation rates at $p_0^\star = 0.33$ is in fact quantitatively very close to the meta-study prediction in Panel (C) of Table 1. Where the meta-study predicts an initial (ongoing) cooperation of 49.5 (37.3) percent, our data at $p_0^\star = 0.33$ indicates similar (and statistically inseparable) rates of 46.4 (36.6) percent. To illustrate this, in Figure 2 we provide the fitted relationships from the meta study but where we overlay results from our four treatments (the smaller circles) using the independent basin-size on the horizontal axis. In addition we also indicate the results pooled over each value for the independent basin measure (the larger diamonds). While there is substantial divergence for initial cooperation in Panel (A), the quantitatively similar results for ongoing cooperation in Panel (B) are clear.[37]

Our tests of the two competing hypotheses are focused on the second and third columns of Panel (B) in Table 2. If the independent (correlated) basin measure captures all relevant facets of behavior, we would expect a significantly negative (insignificant) estimate for the independent-basin increase (correlated-basin-decrease), and an insignificant (significantly negative) effect on the correlated (independent) basin-increase.

For initial cooperation, we find that changes to both basin-measures generate significant effects ($p < 0.001$). The magnitudes of each estimated marginal effect are very similar,

---

[37]In Figure A.1 in the Online Appendix we present an analogous figure organized under the correlated basin-size model, illustrating much poorer organization of the data, both relatively across the treatment comparisons, and quantitatively.

though each moving in opposite directions, as predicted. Since neither effect dominates, we conclude that both X and N contain information for predicting initial cooperation that is not fully absorbed by either basin measure.

However, consistent with our descriptive presentation of the raw treatment rates in Table 2 Panel (A), the probit estimates for ongoing cooperation in Panel (B) do break towards the independent-basin construction. The coefficient on the increase in the independent basin is negative and significant ($p < 0.001$), while the estimate on the decrease for the correlated basin is much smaller in magnitude and is not significant at the 5 percent level ($p = 0.061$). Beyond just the qualitative directional effects, the quantitative change in ongoing cooperation under the independent-basin measure is close to the predicted effects from the relevant basin shifts that we would expect from the RPD meta-study data. That is, the predicted effect from the meta study for ongoing cooperation when the size of the basin increases from $p_0^\star = 0.33$ to 0.69 is a 21 percentage point drop (see Table 1, Panel C), where our estimates indicate a 29 percentage point drop.[38]

As alluded to above, the differences on ongoing cooperation rates between our data and the out-of-sample prediction from the meta-study is driven by the stark (essentially boundary) behavior in the $\binom{N=4}{X=\$9}$ treatment. As illustrated in Figure 1(B) a two-player repeated game with a a basin of size $p^\star = 0.69$ has a predicted ongoing cooperation rate of 16.3 percent, where we should be able to reject 11 percent cooperation at 95 percent confidence. While the $\binom{N=10}{X=\$1}$ treatment is close to the predicted rate (as are the two other treatments at $p^\star = 0.33$), Figure 1(B) clearly indicates the $\binom{N=4}{X=\$9}$ treatment being significantly below the prediction. However, when the non-collusive strategy is risk dominant (when the independent basin size is greater than $\frac{1}{2}$), the argument made in the literature is that we should not expect substantial cooperation in this region. Where the model predicts low cooperation at this basin-size, the evidence from Treatment $\binom{N=4}{X=\$9}$ pushes towards this same conclusion, just in a starker way. Aside from the more-extreme coordination effects at $\binom{N=4}{X=\$9}$, for the other three treatments the theoretically standard extension of the strategic uncertainty measure comes very close to *quantitatively* predicting the ongoing cooperation level using the out-of-sample relationship estimated from the two-player RPD meta-data.

We summarize the main findings:

Result 1 (Independent-Basin Measure)**.** *The independent-basin measure qualitatively organizes the results for ongoing cooperation, and in all but one treatment matches the quantitative level predictions. However, it does not contain all relevant information for predicting initial intentions to collude.*

Result 2 (Correlated-Basin Measure)**.** *Our data is not consistent with the predictions from the correlated-basin hypothesis, neither for initial, nor ongoing cooperation. In particular,*

---

[38]In contrast, for a decrease to 0.04 we should expect an ongoing cooperation-rate increase of 50 percentage points, where we instead we see 11.5 percent.

where the correlated basin predicts that behavior should ceteris paribus be unaffected by N we instead find large cooperation decreases with increases to N.

5. Extensions

Our analysis so far has abstracted away other features of the coordination problem to focus on the pure effects from the primitives of the strategic game. In this section we consider two extensions—with relevance both inside and outside the laboratory—that allow us to study possible limitations of the strategic-uncertainty model to predict changes in equilibrium selection.

First, we consider the extent to which beliefs on others' collusive behavior may be distorted by prior experience. While a policy change can change market primitives and the strategic-uncertainty measure, the underlying variable in this model is beliefs on others' collusion. It seems plausible that beliefs may be driven by experience before any change in the primitives, and so the model may fare poorly at predicting changes within a population. For example, if a player has engaged extensively with the *same* population of market participants under a status quo that led to non-collusive behavior, their beliefs on others acting in this way may be sticky, and therefore unresponsive to shifts in the primitives that have large effects in the strategic uncertainty model. Our treatments in the previous section used a between-subjects design. Identification relied on comparisons of late-session behavior across different populations, each with experience under a fixed environment. In a modified treatment pair we examine the effects of varying the number of players N within the same population, so that participants have experience at two distinct values. In this extension we show that outcomes do not exhibit long-run stickiness, where the between-subject results detailed previously are not substantially dissimilar when we examine them within-subject.

In a second set of extensions, we examine the strategic uncertainty mechanism underlying the basin-size selection device. In particular we examine the extent to which our results are affected by the possibility of explicit coordination, holding constant X and N. Here we seek to mirror an empirical finding that when collusion in industries is detected, it is often accompanied by evidence of explicit collusion—despite the illegality of such meetings.[39] In one treatment we show that once explicit collusion is allowed for, neither the independent nor the correlated basin measures does a good job of predicting collusive behavior levels. Once parties can explicitly collude we find very high levels of sustained cooperation. This suggests that indeed uncertainty over the other strategic choices is a main driver of behavior in our main treatments. However, the extremeness of the effect once communication is allowed for raises a question over the extent to which explicit collusion might lead to high cooperation rates even when collusive outcomes are not equilibrium. To examine this, we show that there a clear limits on what explicit collusion can achieve is defined by the theoretical existence of collusive equilibria.

5.1. **Between vs. Within Identification.** The motivating idea for our first extension is that in many settings of interest the policy-relevant comparative static is being varied

---

[39]See Marshall and Marx (2012) for a more comprehensive treatment.

within a population. However, if agents have strong beliefs about others due to prior experiences, it may be that our theoretical construction lacks bite at predicting outcomes as the policy shifts. If selected equilibria are very sticky within a population, then more-standard assumptions maintaining the equilibria across the counterfactual may have greater validity. For example, within the experiment if a participant's experience with others is that they play the stage-game defection action every period, then this belief can persist despite a policy shift that makes collusion easier.

Ideally, we would introduce a primitive change within a supergame—for example, a move from $N = 4$ to $N = 2$, where the matched player after the modification is one of the matched participants from before. In exploring potential designs for this, we were not satisfied that they would produce clear results. First, it is well-documented that repeated-game environments require several supergames of experience for participants to internalize the environment (Dal Bó, 2005). While implementing a surprise change in N as a mid-supergame manipulation would mirror an outside-the-laboratory consolidation, this would provide a single supergame observation, and require substantial explanation across the surprise. An alternative design choice could implement a change in N with some probability within each supergame. However, any observed effects would then be confounded with the expectations over the primitive change (and greater complication in the instructions) and would no longer comparable to our between-subject treatments.

Given the potential confounds with other designs we instead opted—certainly as a first approach— for a design with a surprise one-time shift in the number of players N, but where this occurred in a fixed session-level population. Holding constant the cooperation cost at $X = \$9$, we initially set a value of N (either two or four) for the first ten supergames. We then change the value of N for the last ten supergames (to either four or two, respectively).

This led to two further experimental treatments, one with $\binom{N=2}{X=\$9}$ in the first half, and $\binom{N=4}{X=\$9}$ in the second; and the converse treatment from $\binom{N=4}{X=\$9}$ in the first, to $\binom{N=2}{X=\$9}$ in the second half. In both treatments, the change in N comes as a surprise: subjects know there is a second part, but do not receive instructions on it until supergame ten concludes.[40] In terms of the standard strategic-uncertainty model this creates a shift across the session from a low basin-size of 0.33 when $N = 2$, and a high basin-size of 0.69 when $N = 4$. In particular, this is a change in N where our between-subject design indicates a substantial treatment effect. Given that we hold constant $X = \$9$, for simplicity we label the treatments as $2 \rightarrow 4$ and $4 \rightarrow 2$, for the first-half to second-half shifts.

In Figure 3(A) we present the average initial cooperation-rates across the session's twenty supergames. The between-subject treatments with $N = 2$ and $N = 4$ are indicated by the two gray dashed lines (separately labeled), while the within-subject treatment results are represented by two black lines: a solid line for the $2 \rightarrow 4$ treatment, and a dash-dotted-line for the $4 \rightarrow 2$ treatment. The figure illustrates the substantial between-subject effect, with more cooperation in $\binom{X=\$9}{N=2}$ over $\binom{X=\$9}{N=4}$ for all twenty supergames. Our within and

---

[40]Our between-subject treatments also divided the session into two parts, except that once subjects reached the second half of the session they were told that part two was identical to part one.

between sessions are identical for the first ten supergames, where the figure indicates replicated effects. Pooling the between and within treatments with $N = 2$ in supergames 6–10 the initial cooperation rate is 47.4 percent. In contrast, the pooled cooperation rate for $N = 4$ is just 13.9 percent.[41]

As we move into supergames 11–20, the number of players matched in each supergame changes for our within treatments. Figure 3(A) indicates the immediate shift in behavior as the primitive changes as the two vertical dotted lines. For the $2 \rightarrow 4$ treatment (the black solid line) initial cooperation levels remain fairly high after the shift from $N = 2$ to $N = 4$. In fact, cooperation in the first four-player interaction (supergame 11) actually exhibits an increase to 59.7 percent from the 53.0 percent from the last two-player interaction (supergame 10). However, while there is no immediate cooperation drop-off, and thus stickiness, as subjects gain experience at $N = 4$ the cooperation rate falls rapidly, reaching 16.7 percent by supergame 20. In contrast, moving in the other direction from $N = 4$ to $N = 2$ (the black dash-dot line) we find an immediate jump with the primitive shift. Where initial-round cooperation in supergame 10 with four-players is 18.3 percent, the reduction to $N = 2$ pushes this rate up to 60.0 percent for supergame 11. The immediate jump in behavior is then sustained across the remaining supergames, with 58.3 percent cooperation in supergame 20.

Inspecting the session time-series illustrated in Figure 3(A) it is clear that the there is little evidence for the hypothesis that equilibrium selection is sticky in the long-run under a within-population shift in N. Despite exposure to the alternative environment in the first half, longer-run behavior in the second-half is not substantially dissimilar from the between-subject levels. This is indicated by the close proximity of the two black/gray line pairs at supergame 20, and relative distance from the other pair.

In Online Appendix B we provide a more detailed like-with-like comparison of the between-subject and within-subject results. These more-detailed findings indicate no difference with the between-subject results as we move from $2 \rightarrow 4$. However, in opposition to the hypothesis that the selected equilibrium is sticky, we actually find a significant *increases* in the response to N over the between analysis as we move from $4 \rightarrow 2$. We summarize the result:

Result 3 (Between vs. Within). *Switching the identification to within does not substantially change our qualitative results, finding no evidence that the selected equilibrium is sticky in the long run as we shift a primitive within the population. If anything, our within-subject identification shows a larger shift than the between-subject results as we decrease N.*

5.2. **Explicit Correlation.** The focus so far, and for the RPD literature more generally, is an examination of *implicit* coordination. Our results thus far suggest that a model of strategic uncertainty extended via independence is relatively successful at organizing the data. This is particularly so for ongoing cooperation. In contrast, a measure based on

---

[41]Testing the initial cooperation rate differences in supergames 6–10 over N (so across the between and within sessions with identical treatment at this point) we find $p = 0.150$ for $N = 2$ and $p = 0.981$ for $N = 4$ from t-tests for a level difference, and $p = 0.353$ for a joint test.
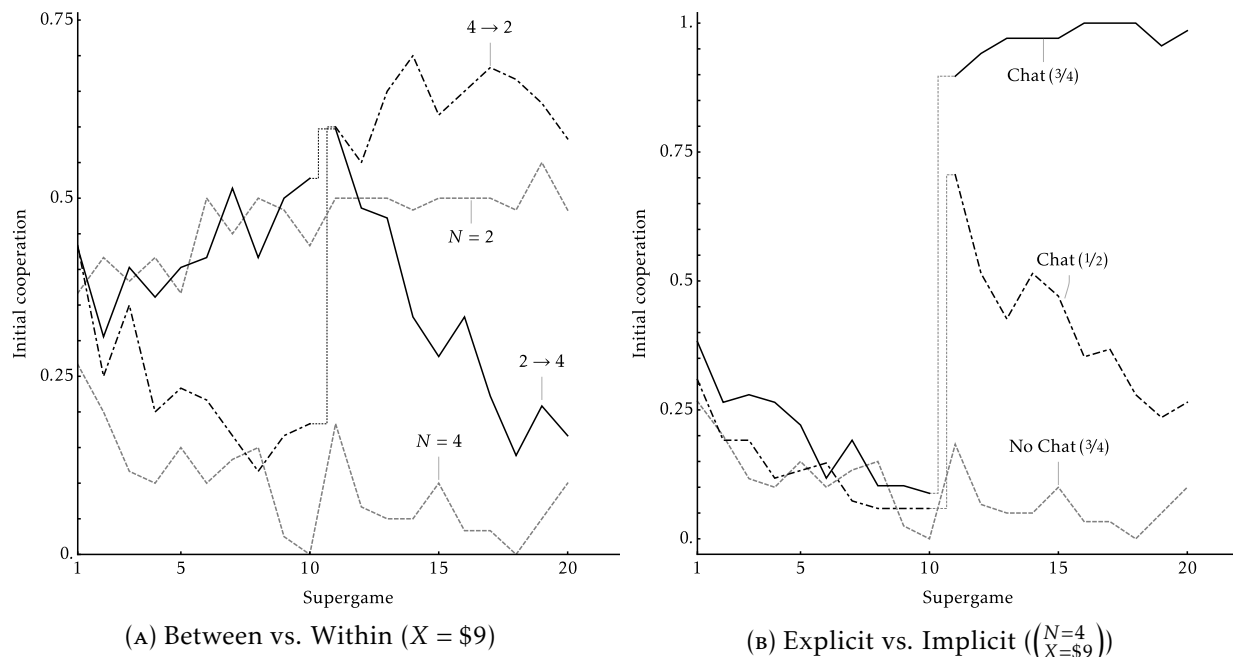
FIGURE 3. Initial cooperation rates in extensions (by supergame)

perfectly correlated play fails to predict the substantial effect of N for both initial and ongoing behavior.

These results are therefore a validation of a standard extension of the basin of attraction notion. However, correlation across participants' beliefs becomes much more plausible when participants have access to a channel through which to explicitly coordinate. In an extension to our previous experiments we now examine the extent to which multiple parties successfully collude when given a coordination device.

To this end, we designed a further treatment pair that modifies our environment with the starkest coordination on the non-collusive outcome, the $\binom{N=4}{X=\$9}$ treatment. In the first chat treatment, the first 10 supergames were identical to the $\binom{N=4}{X=\$9}$ treatment, but we introduce pre-supergame chat between the four matched players in supergames 11–20. A second chat treatment is identical to the first in terms of the timing for when we introduce the chat coordination device; the difference is that we reduce the continuation probability to $\delta' = 1/2$ for the entire session. The designed effect of this change is that while the stage-game payoffs and number of players is held constant, at this value of δ the grim trigger is now only a knife-edge SPE, where the critical belief $p^\star(\delta') = 1$ under both (2) and (3). The $\delta = \frac{1}{2}$ treatment therefore serves as a test for whether explicit coordination can implement outcomes that are not supportable as a robust equilibrium (that is, with arbitrarily small trembles in behavior).

In an online supplement we provide chat logs for the interested reader; however, we here focus on the of effects of chat on cooperation levels. In Figure 3(B) we plot the average

25

Table 3. Cooperation: Implicit vs. Explicit

	Implicit	Explicit	
	NoChat(¾)	Chat(¾)	Chat(½)
Initial coop.	0.035 (0.017)	0.988 (0.007)	0.300 (0.037)
Ongoing coop.	0.006 (0.003)	0.806 (0.030)	0.044 (0.018)
Initial success	0.000	0.971	0.094
Ongoing success	0.000	0.756	0.002

Note: All treatments have $x = \$9, N = 4$, where NoChat(¾) is from the core 2×2 between-subject design outline in section 4. Data taken from last five supergames, participant-clustered standard errors in parentheses.

initial cooperation rates across sessions (with the $\binom{N=4}{x=\$9}$ treatment provided as a baseline, labeled here as NoChat(¾)). Late-session cooperation and success rates (assessed in supergames 16–20 with subject-clustered standard errors) are provided in Table 3.

Our first chat treatment delivers unequivocal results: providing pre-play communication at $\delta = ¾$ takes the near-zero cooperation rate in the absence of chat to almost complete cooperation. While ongoing cooperation drops slightly from the very high initial-cooperation levels, the large majority of supergames exhibit coordination by all four participants on the efficient/collusive outcome. Such high levels of cooperation with communication are inconsistent with the predictions of either model of equilibrium selection for these primitives. This suggests that once explicit coordination devices are allowed for and strategic uncertainty is reduced, the independent model–that captured behavior when collusion is tacit–is no longer helpful.

However, at the $\delta = ½$ boundary for dynamic strategies to be incentive compatible, even with pre-play communication participants find it hard to coordinate. While initial cooperation levels are substantially higher than the treatment without chat at 30.0 percent, ongoing cooperation falls to just 4.4 percent, where successful joint-cooperation across an entire group in an ongoing round becomes exceedingly rare at 0.2 percent. The findings therefore indicate that for explicit communication to play a role, clear incentives for collusion do need to be present.

Given the very stark findings here, we simply summarize the extension result:

Result 4 (Implicit vs. Explicit). *Explicit coordination leads to very high cooperation levels with multiple players, in a setting where implicit cooperation achieves near-zero cooperation. However, in the limiting case, where cooperation is a knife-edge SPE outcome, even pre-play chat fails to support cooperation in our experiments.*

6. Conclusion

Our paper examines equilibrium selection in repeated games, and the extent to which it can be predicted with a model of strategic uncertainty. We leverage a model of equilibrium selection that rationalizes behavior in data from repeated prisoner's dilemma (RPD) experiments and design an experiment to stress test this specific theoretical model. The predictive model works by mediating the effects from multiple-primitives into a single dimension, strategic uncertainty. As such, even for rich counterfactual policies with many changes to the setting, the model can still generate a directional prediction. The experimental design builds on this motivation for the model, introducing a novel source of strategic uncertainty that has not been studied in the RPD setting (the number of players) while also manipulating a payoff parameter. We can thus change both sources of strategic uncertainty simultaneously and study the extent to which the evidence is consistent with the predictions of the selection model. Our main finding is that the model of equilibrium selection can indeed be used as a device for understanding successful coordination on the collusive outcome. In particular, the model performs well in trading-off the competing effects from the two distinct sources of strategic uncertainty.

After illustrating the theoretical power of the model for implicit coordination, we turn our attention to two application-motivated extensions that probe its limitations. In the first, we are motivated by the extent to which prior history and experience could make the equilibria "sticky," even where the model suggests a change. To do this, we study the extent to which our findings still hold when treatment-variable manipulations takes place within the same population, as opposed to between populations in our main study. The results indicate that eve when the participants experience a treatment-variable shift, the model still continues to predict the longer run outcomes. While we do find some hysteresis in the short-run response to a within-population policy change—initial stickiness in behavior in one direction and a large immediate response to the change in the other—behavior after some experience under the new parameters is not distinct from the between-population treatments. As such, while a plausible limitation of the model was the effect of prior experience, these results suggest that the model prediction fares better than the more-standard assumption of maintaining a selected equilibrium across a policy change.

Finally, in a second extension, we examine the potential effects from explicit collusion. Shifting the experimental environment by providing an explicit coordination device (pre-play chat) we examine a necessary limitation on the model if the underlying strategic uncertainty mechanism is correct. Our main finding here is that once we allow for explicit coordination then the selection-model prediction is very far from observed behavior. In fact, where the model suggests very low levels of cooperation, actual behavior is at the other extreme once coordination devices are present. The evidence from the second extension therefore indicates both that strategic uncertainty plays a clear role, but also that the selection model based upon it is entirely inappropriate for predicting behavior when explicit collusion is suspected.

Though our results are encouraging, for the theoretical selection model we examine to be useful in applications further tests are needed. First, as we point out when describing

our experimental design, there are many ways in which expanding the multiple players might impact the game.Because we do not want to introduce further coordination issues, we focus on a setting where collusion requires joint cooperation form all of the agents. But in other interesting environments cooperation among a subset of the agents may be enough for reaching an efficient outcome. Such environments introduce further distributional coordination problems, over who cooperates and who free-rides. Further research is needed to evaluate how and if models of equilibrium selection might be extended to such problems. Outside from the role of multiple players, there are also outstanding questions with regard to how asymmetries for the players—in terms of actions, payoffs, and the sequentiality of moves—should be modeled.

From another perspective, our results also outline that there is room to evaluate the role for equilibrium selection model in environments with explicit collusion. Our findings suggest one clear limit on where explicit collusion can play a role: the existence of a collusive equilibrium. However, beyond this point, not much is known about the conditions for which explicit communication between the parties is enough for collusion. Another direction for future research is to try to understand if there are ways to manipulate the primitives to understand when explicit collusion will and will not be successful.

References

Agranov, Marina, Guillaume Frechette, Thomas Palfrey, and Emanuel Vespa (2016), "Static and dynamic underinvestment: An experimental investigation." *Journal of Public Economics*, 143, 125–141.

Battaglini, Marco, Salvatore Nunnari, and Thomas R Palfrey (2012), "Legislative bargaining and the dynamics of public investment." *American Political Science Review*, 106, 407–429.

Battaglini, Marco, Salvatore Nunnari, and Thomas R Palfrey (2016), "The dynamic free rider problem: A laboratory study." *American Economic Journal: Microeconomics*, 8, 268–308.

Battalio, Raymond, Larry Samuelson, and John Van Huyck (2001), "Optimization incentives and coordination failure in laboratory stag hunt games." *Econometrica*, 69, 749–764.

Berry, James, Lucas C Coffman, Douglas Hanley, Rania Gihleb, and Alistair J Wilson (2017), "Assessing the rate of replication in economics." *American Economic Review*, 107, 27–31.

Blonski, Matthias and Giancarlo Spagnolo (2015), "Prisoners' other dilemma." *International Journal of Game Theory*, 44, 61–81.

Brandts, Jordi and David J Cooper (2006), "A change would do you good.... an experimental study on how to overcome coordination failure in organizations." *American Economic Review*, 96, 669–693.

Cason, Timothy N. (2008), "Price signaling and 'cheap talk' in laboratory posted offer markets." *Handbook of Experimental Economics Results*, 1, 164–169.

Cason, Timothy N., Tridib Sharma, and Radovan Vadovič (2020), "Correlated beliefs: Predicting outcomes in 2×2 games." *Games & Economic Behavior*, 122, 256–276.

Compte, Olivier, Frederic Jenny, and Patrick Rey (2002), "Capacity constraints, mergers and collusion." *European Economic Review*, 46, 1–29.

Cooper, David J. and Kai-Uwe Kühn (2014), "Communication, renegotiation, and the scope for collusion." *American Economic Journal: Microeconomics*, 6, 247–78.

Dal Bó, Pedro (2005), "Cooperation under the shadow of the future: experimental evidence from infinitely repeated games." *American Economic Review*, 95, 1591–1604.

Dal Bó, Pedro and Guillaume R Fréchette (2011), "The evolution of cooperation in infinitely repeated games: Experimental evidence." *American Economic Review*, 101, 411–429.

Dal Bó, Pedro and Guillaume R Fréchette (2018), "On the determinants of cooperation in infinitely repeated games: A survey." *Journal of Economic Literature*, 56, 60–114.

Davis, Douglas D (2002), "Strategic interactions, market information and predicting the effects of mergers in differentiated product markets." *International Journal of Industrial Organization*, 20, 1277–1312.

Deneckere, Raymond and Carl Davidson (1985), "Incentives to form coalitions with bertrand competition." *RAND Journal of Economics*, 473–486.

Devetag, Giovanna and Andreas Ortmann (2007), "When and why? a critical survey on coordination failure in the laboratory." *Experimental economics*, 10, 331–344.

Dufwenberg, Martin and Uri Gneezy (2000), "Price competition and market concentration: an experimental study." *International Journal of Industrial Organization*, 18, 7–22.

Farrell, Joseph and Carl Shapiro (1990), "Horizontal mergers: an equilibrium analysis." *American Economic Review*, 107–126.

Fonseca, Miguel A and Hans-Theo Normann (2012), "Explicit vs. tacit collusion—the impact of communication in oligopoly experiments." *European Economic Review*, 56, 1759–1772.

Fréchette, Guillaume, Alessandro Lizzeri, and Emanuel Vespa (2020), "Collusion in multimarkets: An experiment." NYU Working Paper.

Fudenberg, D., D.G. Rand, and A. Dreber (2010), "Slow to anger and fast to forgive: Cooperation in an uncertain world." *American Economic Review*.

Ghidoni, Riccardo and Sigrid Suetens (forthcoming), "The effect of sequentiality on cooperation in repeated games." *American Economic Journal: Microeconomics*.

Goette, Lorenz and Armin Schmutzler (2009), "Merger policy: What can we learn from competition policy." *Experiments and Competition Policy; Hinloopen, J., Normann, HT, Eds*, 185–216.

Harrington, Joseph E, Roberto Hernan Gonzalez, and Praveen Kujal (2013), "The relative efficacy of price announcements and express communication for collusion: Experimental findings." *Working paper. University of Pennsylvania, The Wharton School*.

Harrington, Joseph E, Roberto Hernan Gonzalez, and Praveen Kujal (2016), "The relative efficacy of price announcements and express communication for collusion: Experimental findings." *Journal of Economic Behavior & Organization*, 128, 251–264.

Harsanyi, John C and Reinhard Selten (1988), *A general theory of equilibrium selection in games*. MIT Press, Cambridge, MA.

Horstmann, Niklas, Jan Krämer, and Daniel Schnurr (2018), "Number effects and tacit collusion in experimental oligopolies." *Journal of Industrial Economics*, 66, 650–700.

Huck, Steffen, Kai A. Konrad, Wieland Müller, and Hans-Theo Normann (2007), "The merger paradox and why aspiration levels let it fail in the laboratory." *Economic Journal*, 117, 1073–1095.

Huck, Steffen, Hans-Theo Normann, and Jörg Oechssler (2004), "Two are few and four are many: number effects in experimental oligopolies." *Journal of Economic Behavior & Organization*, 53, 435–446.

Kartal, Melis and Wieland Müller (2018), "A new approach to the analysis of cooperation under the shadow of the future: Theory and experimental evidence." University of Vienna working paper.

Kartal, Melis, Wieland Müller, and James Tremewan (2017), "Building trust: The costs and benefits of gradualism." University of Vienna working paper.

Kloosterman, Andrew (2019), "Cooperation in stochastic games: a prisoner's dilemma experiment." *Experimental Economics*, 1–21.

Lugovskyy, Volodymyr, Daniela Puzzello, Andrea Sorensen, James Walker, and Arlington Williams (2017), "An experimental study of finitely and infinitely repeated linear public goods games." *Games & Economic Behavior*, 102, 286–302.

Marshall, Robert C and Leslie M Marx (2012), *The economics of collusion: Cartels and bidding rings*. MIT Press.

Perry, Martin K and Robert H Porter (1985), "Oligopoly and the incentive for horizontal merger." *American Economic Review*, 75, 219–227.

Potters, Jan and Sigrid Suetens (2013), "Oligopoly experiments in the current millennium." *Journal of Economic Surveys*, 27, 439–460.

Rosokha, Yaroslav and Chen Wei (2020), "Cooperation in queueing systems." *Working Paper*.

Salz, Tobias and Emanuel Vespa (2020), "Estimating dynamic games of oligopolistic competition: An experimental investigation." *RAND Journal of Economics*, 51, 447–469.

Sherstyuk, Katerina, Nori Tarui, and Tatsuyoshi Saijo (2013), "Payment schemes in infinite-horizon experimental games." *Experimental Economics*, 16, 125–153.

Vespa, Emanuel (2020), "An experimental investigation of cooperation in the dynamic common pool game." *International Economic Review*, 61, 417–440.

Vespa, Emanuel and Alistair J Wilson (2019), "Experimenting with the transition rule in dynamic games." *Quantitative Economics*, 10, 1825–1849.

Vespa, Emanuel and Alistair J Wilson (2020), "Experimenting with equilibrium selection in dynamic games." *Working Paper*.

Vesterlund, Lise (2016), "Using experimental methods to understand why and how we give to charity." *Handbook of Experimental Economics*, 2, 91–151.

Weber, Roberto A (2006), "Managing growth to achieve efficient coordination in large groups." *American Economic Review*, 96, 114–126.

Wilson, Alistair J. and Emanuel Vespa (2020), "Information transmission under the shadow of the future: An experiment." *American Economic Journal: Microeconomics*, 12.