Robust Cooperation with First-Order Information*

Daniel Clark, Drew Fudenberg, and Alexander Wolitzky

January 10, 2019

Abstract

We study the repeated prisoner’s dilemma in a large population with random matching and overlapping generations. Our goal is to study the extent to which cooperation can be supported by equilibria where players use simple strategies and have very limited information about their opponents. To this end, we assume players have access to only first-order information about their current partners, meaning that a player’s record tracks information about her past actions only, and not her partners’ past actions (or her partners’ partners’ actions). Cooperation in strict equilibrium is impossible if payoffs are submodular, or if players can erase their records. If payoffs are sufficiently supermodular, then cooperation can be sustained by a tolerant version of grim trigger strategies, where a player retains good standing until her record reflects a certain number of defections. Players close to the threshold may cooperate even with defectors; this “unconditional cooperation” is crucial for sustaining maximal cooperation. If players can forge records of fake past interactions and successfully coordinate with their current partners, this tolerant version of grim trigger is the unique family of strategies that sustains cooperation in equilibrium.

*We thank V Bhaskar, Glenn Ellison, Yuval Heller, Takuo Sugaya, Satoru Takahashi, and Caroline Thomas for helpful comments and conversations, and NSF grants SES 1643517 and 1555071 and Sloan Foundation grant 2017-9633 for financial support.
1 Introduction

In many settings of economic interest, individuals interact with different partners over time, and bad behaviour against one partner causes a negative response by other members of society. This paper studies such “community enforcement” in the repeated prisoner’s dilemma. Our two goals are to better understand what sorts of information suffice for good community outcomes and to provide a foundation for the analysis of relatively simple strategies that we think might be descriptively plausible. To this end, we depart from past models of community enforcement (which we review in Section 2) by dropping the assumption of a common start date and calendar time, and by considering quite restrictive forms of information and record-keeping. We then study when cooperation is possible under these restrictions, and how the maximal level of cooperation varies with the parameters of the underlying game.

To place our work in context, recall that in the standard repeated game model, a fixed set of players interacts repeatedly with a commonly known start date and a common notion of calendar time. When each player’s signals are sufficient to statistically identify the vector of her opponents’ actions, equilibria that support cooperation usually exist when players are patient, but the most efficient equilibria are typically “complicated” if there is any noise in the monitoring structure. This model seems natural for studying some long-term relationships with well-defined start dates among a relatively small number of relatively sophisticated players, such as business partnerships or collusive agreements among firms. However, laboratory studies of repeated games suggest that many subjects use fairly simple strategies.\(^1\) Moreover, repeated games have also been used to model cooperation in large populations, and for these applications the assumptions of a fixed population, a common start date, and common calendar time seem less appropriate, and there is all the more reason to doubt whether players will use complicated strategies.

Thus, instead of analyzing interactions in fixed pairs, we consider a class of repeated strategies.

\(^1\)See e.g. Fudenberg, Rand, and Dreber (2012) and the survey by Dal Bó and Fréchette (2018).
games with random matching, no commonly known start date or calendar time, and restrictive forms of information about past play. In our model, there is continuum of players with geometrically distributed lifespans. Each player has a series of pairwise interactions with different partners. We assume players have only “first-order” information about their partners, meaning that their information depends only on the actions the partner has taken, and not on the actions or histories of the partner’s past partners. In particular, each player carries a “record” which depends only on her own past actions (perhaps stochastically), and when two players meet they observe each other’s record and nothing else. We study steady states of this population game, where the share of the population with each record is constant over time, and each player’s strategy depends only on her own record and the record of her current opponent. A preliminary result is that a steady state always exists.

We then assume the stage game is the standard prisoner’s dilemma:

\[
\begin{array}{ccc}
C & D \\
C & 1,1 & -l,1+g \\
D & 1+g,-l & 0,0 \\
\end{array}
\]

Figure 1: The Prisoner’s Dilemma

with \( g, l > 0 \) and \( l + 1 > g \), so \((C,C)\) maximizes the sum of payoffs. Here \( g \) measures the gain to defection (that is, playing \( D \)) when one’s opponent cooperates, for example the cost savings from providing a low quality product or service, or the profit gained by undercutting competitors in a cartel. Moreover, because \( l \) measures the gain from playing \( D \) against \( D \), the comparison of \( g \) and \( l \) is a measure of the complementarity in the interaction. As we will see, the possibility and maximal extent of equilibrium cooperation in our setting depend on the degree of complementarity as well as the temptation to deviate. Intuitively, this is because first-order information is not sufficient to distinguish between opportunistic deviations to \( D \) and equilibrium plays of \( D \) that punish opponents with bad records, and so players must sometimes be willing to
worsen their record by playing $D$ against $D$ when their continuation payoff would be higher if they played $C$ and incurred a short run loss.

Throughout the paper we restrict attention to strict equilibria; this captures a simple form of robustness and, in particular, rules out “belief-free” equilibria and related constructions. The steady state where everyone always plays $D$ regardless of the records is always a strict equilibrium. As in the related random matching models of Takahashi (2010) and Heller and Mohlin (2018), we find that when the prisoner’s dilemma stage game is “submodular,” that is when $g \geq l$, the only strict equilibrium is *Always Defect*, regardless of the community’s record-keeping system. We thus focus on the supermodular case where $g < l$. In our setting, though, this is not sufficient to permit cooperative equilibria; we will see that we additionally require sufficiently high complementarity.

The space of possible record-keeping systems that one could consider is very large. To sharpen our focus, we note that in some economic environments players may be able to manipulate their records, most notably by evading record-keeping or by forging fake positive records. We show that, if players can evade record-keeping, the only strict equilibrium is again *Always Defect*. We thus consider record-keeping systems where evasion is impossible, and instead ask for robustness to forging records of fake interactions, a property we call *forgery-proofness*.\(^2\) This consideration, along with an interest in simplicity, leads us to study the very simple technology that records only the number of times a player has played $D$, but not when those times were or how many times the player has played $C$ (as rewarding players for having more $C$’s in their record would encourage them to fake $C$’s).

For this information structure, we study how much cooperation can be supported in\(^2\)

---

\(^2\)Fake reviews are an important problem for many real-world reputation systems. For example, fake reviews on Yelp are sufficiently common that Yelp uses an algorithm to (imperfectly) filter many of them out (Luca and Zervas, 2016). For an empirical analysis of review manipulation on Expedia and TripAdvisor, see Mayzlin, Dover, and Chevalier (2014). In the context of online book reviews, Chevalier and Mayzlin (2006) find that the relatively rare one-star reviews carry a lot of weight with consumers. They argue that this is because “the author can post a large number of meaningless five-star reviews cheaply, (but) cannot prevent others from posting one-star reviews,” (pp. 349-350).
the limit as expected lifespans grow to infinity (i.e. the continuation probability $\gamma \to 1$) and “noise” (i.e., the $\varepsilon$ probability that a play of $C$ is mis-recorded as $D$) shrinks to 0, and say that there is a limit efficient equilibrium when this share converges to 1. Our first observation here is that limit efficient strict equilibria always exist if the prisoner's dilemma is mild, meaning that $g < 1$. The strategies we use in this construction are “forgiving,” in the sense that no matter how many $D$’s there are in a player’s record, the appropriate sequence of play will lead to them being forgiven and treated as a cooperator. Moreover, as in the strategy “perfect tit-for-tat” (Fudenberg and Maskin, 1990), the way a player earns forgiveness is by acquiring additional $D$’s in her record. Thus, a player with more $D$’s in her record is sometimes treated better than a player with fewer $D$’s. For this reason, the equilibrium is not forgery-proof.

However, forgery-proofness on its own does not rule out all the equilibria that we find implausible. In particular, it allows there to be multiple disjoint sets of records, where players with records in each set cooperate with each others but not with the reciprocators in other sets: For example, players with 2 $D$’s might cooperate only with other players with 2 $D$’s, and players with 4 $D$’s might only cooperate with other players whose record is 4. Thus we also require that strategies are also coordination-proof, which means that, if the prisoner’s dilemma faced by a pair of players becomes a coordination game when it is augmented by their equilibrium continuation payoffs (as a function of their records), they play the Pareto-dominant equilibrium. Combined, these restriction imply that each player uses a strategy of a form we call GrimKL. This strategy partitions players into three groups: players with $K - 1$ or fewer $D$’s in their records cooperate with any player with $K + L - 1$ or fewer $D$’s; players with between $K$ and $K + L - 1$ $D$’s are unconditional cooperators (who cooperate with everyone); and players with $K + L$ or more $D$’s are unconditional defectors. We call the special case where $L = 0$—so there are no unconditional cooperators—GrimK.\footnote{GrimK was introduced by Fudenberg, Rand, and Dreber (2012) who noted that while it is never an equilibrium in a two player game with perfect monitoring, it can be an equilibrium when actions are observed with noise, and that some experimental subjects seem to use such strategies}
Having identified GrimK and GrimKL as the only plausible strategies in our environment, we then study when these strategies can support cooperation, and in particular ask when they are limit efficient (for appropriately chosen \(K\) and \(L\)). We find that GrimK cannot support any cooperation when the static incentive to deviate from cooperation is “large” in the sense that \(g \geq 1\), regardless of the value of the continuation probability \(\gamma\). When instead \(g < 1\) and \(l > g/(1 - g)\), GrimK can support some cooperation in the limit where \(\gamma \to 1\) and then \(\varepsilon \to 0\), and the degree of limit inefficiency shrinks as \(l\) grows. To see the intuition, note that, since a player gains at least \(g\) by deviating from \(C\) to \(D\), a player’s continuation payoff must be reduced by \(g\) whenever her record acquires a \(D\). With GrimK strategies, this reduction must be achieved by “switching” \(g\) future plays of the stage game from \((C, C)\) to \((D, D)\). Since each such switch also adds a \(D\) to the partner’s record, this leads to additional switches. When \(g < 1\), these switches dampen over time, while when \(g \geq 1\) they “snowball.” Thus, positive steady-state cooperation is possible when \(g < 1\), but not when \(g \geq 1\).\(^4\)

Our last main result is that the more flexible GrimKL strategies cannot support cooperation when \(l < g \left(1 + g\right)\) (regardless of \(\gamma\)) but are limit efficient when \(l > g \left(1 + g\right)\).\(^5\) (The condition \(g \geq 1\) plays no role here.) Thus, unlike GrimK strategies, GrimKL strategies are sometimes fully limit efficient, and they can be limit efficient even for parameters where GrimK strategies cannot support any cooperation. Specifically, GrimKL does not require a “small” static gain from defection, and requires less complementarity than GrimK does. The reason GrimKL strategies can be effective even when \(g \geq 1\) is that they incorporate “extra effort” in addition to “punishment”: With GrimK strategies a player’s continuation payoff can fall only by switching play from \((C, C)\) to \((D, D)\), while with GrimKL strategies a player’s continuation payoff can also fall by switching play from \((D, D)\) to \((C, D)\). This additional loss makes it possible to avoid the “snowballing” of \(D\)’s that precludes cooperation with GrimK strategies.

\(^4\)Moreover, as we show in section 5.1, this result extends to any strategy that does not involve unconditional cooperators.

\(^5\)Technically, the lower bound on \(l\) required for our limit efficiency result may be slightly greater than \(g \left(1 + g\right)\) due to an integer problem. See Section 6.3.
when $g \geq 1$.

We find it particularly interesting that GrimKL strategies can support cooperation even in severe prisoner’s dilemmas ($g \geq 1$) by using extra effort from players with marginal records to cut off the snowballing defections that would arise under GrimK. This is somewhat related to notions of “repentance” or “restitution” that have been shown to support sustainable cooperation in some well-known case studies (Milgrom, North, and Weingast, 1990; Ellickson, 1994). However, under repentance a player with marginal standing can improve her standing by cooperating, while under GrimKL the best she can do is prevent further deterioration of her status. This difference comes from the fact that we want equilibria to be forgery-proof: if players can forge records that say they cooperated, repentance strategies break down, while GrimKL is robust.

All our main results involve records that (noisily) track the number of times a player has played $D$. If records instead track the number of times a player has played $C$, we show cooperation is impossible. More powerful record-keeping systems—for example, systems that track the number of both $C$’s and $D$’s, or that track the timing of actions—could sometimes support more cooperation. However, exploiting these systems may require more complicated strategies, and we show that (very simple) GrimKL strategies can already attain full limit efficiency when only $D$’s are tracked, as long as the game is “sufficiently supermodular.” Nonetheless, investigating different information structures is one promising direction for future research.

2 Literature Review

2.1 Random Matching with Limited Information

Rosenthal (1979) and Rosenthal and Landau (1979) introduced the study of repeated games with random matching. Rosenthal (1979) considered the special case of first-order information where players know only the action that their current opponent played in the previous period. He showed that Markovian equilibria exist, and that in
the prisoner’s dilemma cooperation can be supported by pure strategy equilibria only for a particular knife-edge value of the discount factor.\textsuperscript{6}

Kandori (1992) and Ellison (1994) showed that cooperation in the prisoner’s dilemma can be enforced by “contagion equilibria” in finite populations even when the players have “zero-order” information—that is, no information at all—about each other’s past, but the required discount factor converges to 1 as the population becomes infinitely large. Kandori constructed simple contagion equilibria that exist only under a fairly strong restriction on the payoff functions—the loss parameter $l$ needs to be sufficiently large, and in particular must diverge to $\infty$ as $\delta \to 1$. The issue is that if a player is very patient then even when he sees another player defect, and so knows that contagion has begun, he may still choose to cooperate to slow the spread of the contagion.\textsuperscript{7} Ellison extended Kandori’s results both to arbitrary payoff parameters and to approximately efficient equilibria in settings with a small amount of noise by using either public randomizing devices or “threading,” which both serve to lower the players’ effective discount factor, and so make it incentive-compatible for players to carry out the punishments prescribed by the equilibrium. We find these sorts of strategies unintuitive, and some of the details of our model are designed to rule them out.\textsuperscript{8}

Three previous papers have studied cooperation in continuum-player repeated games with anonymous random matching and first-order information: Takahashi (2010), Heller and Mohlin (2018), and Bhaskar and Thomas (2018). Takahashi (2010) shows how cooperation can be supported when players know the entire record of each part-

\textsuperscript{6}Rosenthal and Landau (1979) study two particular sorts of record-keeping technologies in the setting of an asymmetric battle of the sexes game. Their second, simpler, model has an index that goes up or down by 1 depending on which of the two actions the player uses, so it is related to the sorts of records we analyze in this paper; their first model of “comparative records” allows records to encode more than first-order information.

\textsuperscript{7}Kandori also shows that cooperation can always be supported if higher-order information is available. Okuno-Fujiwara and Postlewaite (1995) derived a similar result in a continuum-population model. See also the literature on “standing” in evolutionary biology, starting with Sugden (1986) and more recently surveyed by Sigmund (2012).

\textsuperscript{8}Deb, Sugaya, and Wolitzky (2018) prove the folk theorem for finite-population repeated games with anonymous random matching. Their construction relies heavily on a finite population, common calendar time, and non-strict incentives.
ner’s past play—all of the “first-order” information—but no higher-order information, such as the action that today’s partner’s partner took yesterday. Takahashi’s general construction requires the use of somewhat complex and unintuitive “belief-free” mixed strategies. These strategies track the expected payoff of each player given her history of play and the equilibrium distribution of states and actions, and reduce a player’s continuation payoff each time she plays D by just enough that she is always indifferent between C and D at every history. Because of this indifference, each player is willing to play C with exactly the probability corresponding to the current partner’s record.

Takahashi also shows that first-order information allows cooperation to be supported in strict equilibria when the players are patient and the game is strictly supermodular. To do this, he follows Ellison and uses threads to calibrate the effective discount factor to be within the interval where players want to cooperate against a player who cooperates and defect against the player who defects. Conversely, Takahashi shows that no strategies support cooperation as a strict equilibrium when the payoffs are strictly submodular. The intuition is simple: with only first-order information, a player’s continuation payoff depends on her record and what she does today but not on her current partner’s action. Thus, in order for a player to strictly prefer to cooperate with a partner who will cooperate while defecting against a player who defects, payoffs must be strictly supermodular, and moreover the effective discount factor must be low enough that the difference in payoffs caused by the opponent’s current action can offset the difference in future continuation payoffs.

Heller and Mohlin (2018)’s study of the prisoner’s dilemma with anonymous random matching assumes that a small fraction of players are commitment types—which rules out belief-free equilibria—and that players live forever and are infinitely patient, but see only a finite sample of their partners’ past actions. Players are restricted to use stationary strategies that condition only on the sampled actions of their partners and not on their own histories. Here, as in Takahashi, there are no cooperative equilibria when payoffs are submodular, and for much the same reason: players will play C or D depending on how likely it is their opponent plays C, so when their observation of
their partner’s play consists only of D’s they are most likely to play C. But this means there is no incentive to play C, and cooperation breaks down. Conversely, Heller and Mohlin show that when payoffs are supermodular, the presences of commitment types allows the construction of an efficient and relatively simple mixed-strategy equilibrium without threads or public randomization, and that this equilibrium is essentially the only one that supports cooperation in their model.\footnote{Heller and Mohlin also consider alternative information structures where players observe, for example a finite sample of their partners’ past action \textit{profiles}. The conditions for cooperation in this setting depend on payoff parameters via “snowballing” considerations similar to those in some of our results. However, the many differences between our models make the results difficult to compare directly. In another model with commitment types, Dilmé (2016) constructs a belief-free cooperative equilibrium for the case where $g = l$.}

Bhaskar and Thomas (2018) study a sequential-move “lending game” with one-sided moral hazard, where borrowers are constrained to default with a fixed i.i.d. probability. They show that cooperation can be supported in a strict (or at least purifiable mixed) equilibrium if lenders are told only whether or not a borrower has defaulted in any of the last $K$ periods for some sufficiently large $K$. The distinction between submodular and supermodular games does not arise here due to the sequential nature of the game.

Nowak and Sigmund (1998) and many subsequent papers study the enforcement of cooperation using “image scoring,” which means that each player has first-order information about her partner, but conditions her action only on her partner’s record and not her own record. These strategies are never a strict equilibrium, and are typically unstable in environments with noise (Panchanathan and Boyd, 2003). One interpretation of our model is that it shows that image scoring-type strategies can be strict equilibria, provided the game is supermodular and players condition on their own record as well as their partner’s.

There is also a literature on repeated games with overlapping generations of non-anonymous players, e.g. Cremer (1986), Kandori (1992), Salant (1991), and Smith (1992). The lack of anonymity makes these papers less directly relevant.
2.2 Simple Strategies in Fixed-Pair Interactions

Rubinstein (1986) and Abreu and Rubinstein (1988) introduced the study of repeated games played by automata. They assumed perfect monitoring, and as they acknowledge, their results are not robust to even a slight amount of noise. Compte and Postlewaite (2015) study fixed pairs playing the gift-exchange version of the prisoner's dilemma (which is on the boundary between the supermodular and submodular cases described above) with imperfect binary private signals of the partner's action. To model simplicity, they assume that strategies can be represented by automata with only two states, and they determine how the extent of cooperation is limited by the accuracy of the monitoring technology. Joe et al. (2012) computationally study whether “k-period mutual punishment strategies”, which are conceptually similar to perfect tit-for-tat, can support cooperation in a two-player repeated prisoner’s dilemma with private monitoring, and show that a simple version of this strategy can sustain fairly high payoffs for a range of parameters. Möbius (2001) and Olszewski and Safronov (2018a,b) consider simple “chips strategies” in repeated games, where a player gives her partner a token whenever he does her a favor, and a player who runs out tokens stops receiving favors until she reciprocates. These strategies are sometimes approximately efficient when players are patient.

Most evolutionary models of repeated games restrict to simple strategies to make the analysis tractable. This is the case for example in Axelrod and Hamilton (1981), who used evolutionary stability to argue that people will use the strategy “Tit-for-Tat” in the repeated prisoner’s dilemma.\(^{10}\)

Finally, the modern literature on repeated games in the laboratory provides evidence that simple strategies are used in play of fixed-partner repeated games (see e.g. Dal Bó and Fréchette (2018)). This helps motivate our interest in simple strategies.

\(^{10}\)Axelrod (1984) showed that Tit-for-Tat and other simple strategies such as Tit-for-Two-Tats were selected in tournaments where participants submitted fully complete strategies to play the repeated game, but the requirement to submit strategies, as opposed to playing the extensive form version of the game period-by-period, may have led to simpler strategies than would otherwise have been used.


3 Random Matching with First-Order Information

Although most of the paper studies the prisoner’s dilemma, we first present a model of
first-order information in general stage games that may be of use in future work. We
consider a discrete time model with a constant unit mass of players, each of whom has
a geometrically-distributed lifespan with continuation probability $\gamma \in (0, 1)$, with exits
balanced by a steady flow of new entrants. To motivate our exclusion of strategies that
condition on calendar time, we assume the time horizon is doubly infinite.

Fix a finite symmetric game with action space $A$ and payoff function $u : A \times A \rightarrow \mathbb{R}$. When players match they observe each other’s record, which for now is an arbitrary
integer $n \in \mathbb{Z}$. New players all enter with the same initial record $n_0$, which we set equal
to 0 without loss of generality. One key restriction that we maintain throughout the
paper is that these records track only first-order information—that is, data about how
the player played—and do not depend on the play or records of the players she has
been matched with.\footnote{This contrasts with the “status levels” studied by Okuno-Fujiwara and Postlewaite (1995) and the “standing” models of Sugden (1986) and Kandori (1991).} The state of the system is then the share of players with each
possible record; we denote this by $\mu \in \Delta(\mathbb{Z}) := M$.

To operationalize anonymous random matching in a continuum population, we
specify that, when the current state of the system is $\mu$, the distribution of matches is
given by $\mu \times \mu$, so that for each $(n_1, n_2) \in \mathbb{Z}^2$ there are $\mu(n_1)\mu(n_2)$ matches between
players of record $n_1$ and players of record $n_2$.

**Definition 1.** A record-keeping system is a function $r : \mathbb{Z} \times A \rightarrow \Delta(\mathbb{Z})$ that specifies
a probability distribution over a player’s record tomorrow given the player’s current
record and her realized action in their current match. A record-keeping system $r$ has
bounded-support updates if there exists $B \in \mathbb{Z}$ such that support $(r(z, a))$ has at most
$B$ elements for all $(z, a) \in \mathbb{Z} \times A$.

Here the stochastic term can represent errors in record-keeping, but it could also
correspond to imperfect implementation of the intended action.\footnote{To restrict records to first-order information, we do not allow the update of a player’s record}
We assume that when players meet each sees the record of her current opponent. In principle, each player can condition her play on the entire sequence of outcomes and past opponent records that she has seen. However, since we work in a model with a continuum of players, only the player’s current record and that of her current partner matter for the player’s current payoff, and only the player’s own record will matter in the future. For this reason, all strict equilibria are record-dependent, meaning that they condition only on the player’s current record and the record of her current partner. We write a record-dependent pure strategy as a function \( s : \mathbb{Z} \times \mathbb{Z} \to A \), with the convention that the first coordinate is the player’s own record and the second coordinate is that of the partner. Since we will restrict attention to strict equilibria, we consider only pure record-dependent strategies. Moreover, since every strict equilibrium in a symmetric, continuum-population model is symmetric, we also assume all players use the same strategy.

Given a record-keeping system \( r \) and a strategy \( s \), we can define an update map \( f_{r,s} : M \to M \) as follows: First, let \( \phi(k', k'') := \mu(k') \mu(k'') \) denote the probability that a player with record \( k' \) meets a player with record \( k'' \). Next, let \( \rho_{r,s}(k', k'') \in \Delta(\mathbb{Z}) \) be the probability distribution over next-period records of a player with record \( k' \) who meets a player with record \( k'' \), when all players use strategy \( s \). Thus, \( \rho_{r,s}(k', k'')[k] = r(k', s(k', k''))[k] \). Then, for \( k \neq 0 \), \( f_{r,s}(\mu)[k] := \gamma \sum_{k', k''} \phi(k', k'') \rho_{r,s}(k', k'')[k] \), and \( f_{r,s}(\mu)[0] := 1 - \gamma + \gamma \sum_{k', k''} \phi(k', k'') \rho_{r,s}(k', k'')[0] \). A steady-state under \( r, s \) is a state \( \mu \) such that \( f_{r,s}(\mu) = \mu \).

**Theorem 1.** Under any record-keeping system with bounded-support updates and any record-dependent strategy, a steady state exists.

The proof is in A.1 of the Appendix; all other omitted proofs can be found in either the Appendix (A) or the Online Appendix (OA). In outline, we relabel records so that to depend on the action that her partner played. In the implementation-errors interpretation of stochastic records, this requires that the signals have a product structure in the sense of Fudenberg, Levine, and Maskin (1994).
two players with different ages can never share the same record, let \( R(t) \) be the set of feasible records for a player of age \( t \), and let \( \bar{M} = \left\{ \mu \in M : \sum_{k \in R(t)} \mu(k) \leq \gamma^t \forall t \in \mathbb{Z} \right\} \).

We first show \( \bar{M} \) is compact in the sup norm: intuitively, bounded-support updates and geometrically distributed lifetimes imply that most players have records in the finite set \( \bigcup_{t \leq T} R(t) \) for bounded \( T \), so \( \bar{M} \) resembles a finite-dimensional space. We then show that \( f \) maps \( \bar{M} \) to itself and is continuous in the sup norm and note that \( \bar{M} \) is convex, so we can appeal to a fixed point theorem.

Throughout the paper, our focus will be on equilibrium steady states. But note that Theorem 1 does not assert that the steady state for a given strategy is unique, and indeed we will see examples where it is not, as is the case under the strategy Grim2. Intuitively, this multiplicity corresponds to different initial conditions at time \( t = -\infty \).

It remains to define equilibrium. Given a record-keeping system \( r \), strategy \( s \), and state \( \mu \), define the flow payoff of a player with record \( k \) as

\[
\pi_k = \sum_{k'} \mu(k') u(s(k, k'), s(k', k)).
\]

Next, denote the probability that a player with record \( k \) today has record \( k' \) \( t \) periods from now by \( \rho^t_{r,s}(k)[k'] \): this is defined recursively by

\[
\rho^1_{r,s}(k)[k'] = \sum_{k''} \mu(k'') \rho_{r,s}(k, k'')[k']
\]

and, for \( t > 1 \),

\[
\rho^t_{r,s}(k)[k'] = \sum_{k''} \rho^{t-1}_{r,s}(k)[k''] \rho_{r,s}(k'')[k'].
\]

The continuation value of a player with record \( k \) is then given by

\[
V_k = (1 - \gamma) \sum_{t=0}^{\infty} \gamma^t \sum_{k'} \rho^t_{r,s}(k)[k'] \pi_{k'}.
\]

Note that we have normalized continuation payoffs by \( (1 - \gamma) \) to express them in per-
period terms. A pair \((s, \mu)\) is an equilibrium if \(\mu\) is a steady-state under \(r, s\) and, for each own record \(k\) and opponent’s record \(k'\), we have

\[
s(k, k') \in \arg \max_{a \in A} \left[ (1 - \gamma)u(a, s(k', k)) + \gamma \sum_{k''} r(k, a)[k'']V_{k''} \right].
\]

In particular, a player’s objective is to maximize her expected undiscounted lifetime payoff. An equilibrium is strict if the argmax is unique for all pairs of records \((k, k')\).

Equilibrium existence follows immediately from Theorem 1.

**Corollary 1.** Under any record-keeping system with bounded increments, an equilibrium exists.

**Proof.** Fix a symmetric stage game Nash equilibrium \(\alpha^*\), and let \(s\) recommend \(\alpha^*\) at every record pair \((k, k')\). Then \((s, \mu)\) is an equilibrium for any steady state \(\mu\).  

### 4 Cooperation in the Prisoner’s Dilemma

For the rest of the paper we specialize to the case where the stage game is the prisoner’s dilemma, given by Figure 1.\(^{13}\) Note that *Always Defect* (i.e., \(s(k, k') = D\) for all \(k, k'\)) is a strict equilibrium for any parameter values. We call an equilibrium non-trivial if \(C\) is played with positive probability (i.e., \(s(k, k') = C\) for some \(k, k'\) with \(\mu(k), \mu(k') > 0\)).

This section derives two basic necessary conditions for the existence of a non-trivial equilibrium: the prisoner’s dilemma stage game must be strictly supermodular, and it must not be possible for players to evade the record-keeping technology.

\(^{13}\)The normalization in the figure is without loss: given a symmetric prisoner dilemma with payoff matrix \[
\begin{array}{ccc}
C & R & S \\
D & T & C \\
\end{array}
\]
we can subtract \(P\) from all entries and divide by \(R - P\), so that \(g = (T - R)/(R - P)\) and \(l = (P - S)/(R - P)\). If the stochastic nature of record updating comes from a probability \(\varepsilon\) that attempts to play \(C\) result in \(D\), then the expected payoff matrix is slightly different, for example the payoff to \((C, C)\) is \((1 - \varepsilon)^2R + \varepsilon(1 - \varepsilon)(T + S) + \varepsilon^2P\). In this case we denote the corresponding normalized payoffs by \(g(\varepsilon)\) and \(l(\varepsilon)\) respectively, where \(g(0) = g\) and \(l(0) = l\).
4.1 Supermodular Payoffs

We show that in our model, as in Takahashi (2010) and Heller and Mohlin (2018), payoffs must be strictly supermodular for cooperation to be sustained in equilibrium. Recall that the prisoner’s dilemma is strictly supermodular if \( g < l \), so the benefit of defecting is greater when the opponent defects. Conversely, the stage game is strictly submodular when \( g > l \). A leading example of the prisoner’s dilemma is reciprocal gift giving, where each player can pay a cost \( c > 0 \) to give her partner a benefit \( b > c \). In this case, a player receives the same static gain from playing \( D \) instead of \( C \) regardless of the play of her opponent, so \( g = l \), and the game is neither strictly supermodular nor strictly submodular. Bertrand competition (with two price levels) is supermodular, while Cournot competition (with two quantity levels) is submodular whenever marginal revenue is decreasing in the opponent’s quantity. Our first lemma notes a simple consequence of strict equilibrium. Note that this result holds for any record-keeping system.

**Lemma 1.** At any record \( k \) in any strict equilibrium:

1. If \( g = l \) then either the player is
   - an unconditional defector who plays \( D \) regardless of the opponent’s record
   - an unconditional cooperator who plays \( C \) regardless of the opponent’s record.

2. If \( g > l \) then either the player is
   - an unconditional defector who plays \( D \) regardless of the opponent’s record
   - an unconditional cooperator who plays \( C \) regardless of the opponent’s record
   - an anti-reciprocator who plays \( C \) against opponents who play \( D \) and plays \( D \) against opponents who play \( C \).

3. If \( g < l \) then either the player is
   - an unconditional defector who plays \( D \) regardless of the opponent’s record
• an unconditional cooperator who plays C regardless of the opponent’s record
• a reciprocator who plays C against opponents who play C and plays D against opponents who play D.

Proof. Fix a strict equilibrium. Because the records use only first-order information, each player’s continuation payoff depends only on her current record and this period’s play, so the optimal action in each period depends only on the player’s record and the action prescribed by her opponent’s record. There are four forms such a strategy could take; the lemma shows how which forms are possible in equilibrium depends on how $g$ compares to $l$.

1. If $g = l$, the current period gain from playing $D$ instead of $C$ is independent of the opponent’s action, so $C$ is the strict best response against $C$ iff it is also a strict best response to $D$, in which case the player is an unconditional cooperator. Similarly, $D$ is the best response to $C$ iff the player is an unconditional defector.

2. If $g > l$ and $C$ is the strict best response to $C$, we have

$$(1 - \gamma)(u(C, D) - u(D, D)) > (1 - \gamma)(u(C, C) - u(D, C)) > \gamma(V^C_k - V^D_k),$$

so $C$ is also the best response to $D$. This shows there are no reciprocators, so every player is either an anti-reciprocator, an unconditional cooperator, or an unconditional defector.

3. If $g < l$, a similar argument shows that if $D$ is optimal against $C$ it is optimal against $D$, so there are no anti-reciprocators.

\[\square\]

Theorem 2. The payoffs must be strictly supermodular for C to be played in any strict equilibrium.

Proof. Suppose that $g \geq l$. Fix a number $k$, and suppose two players who both have record $k$ meet each other. By symmetry, they play either $(C, C)$ or $(D, D)$. In the
former case, $C$ is the strict best response to $C$, and therefore is also the strict best response to $D$. In the latter case, $D$ is the strict best response to $D$, and therefore is also the strict best response to $C$. Thus there are no anti-reciprocators, and each player with record $k$ is either an unconditional cooperator or an unconditional defector. Since this holds for all $k$, the distribution of opposing actions faced by any player is independent of her record. This implies $D$ is always optimal.

For the rest of the paper we restrict to the supermodular case where $g < l$. Here it is possible for some records to be reciprocators, which is what will allow equilibria that support some cooperation.

4.2 Impossibility of Evading the Record-Keeping System

In some environments players may be able to prevent their interactions from being recorded. We say that an equilibrium is *evasion-proof* if it remains an equilibrium when players can, at the start of any period, evade the recording system and so ensure that their record tomorrow is the same as it is today. One special case of this is when, as in the next section, a player’s record is simply the number of times she has played $D$. Here evading the recording technology is equivalent to preventing some or all of the $D$’s from being recorded in the player’s record.

**Theorem 3.** Any equilibrium in which $C$ is played with positive probability is not *evasion-proof*.

*Proof.* Let $\bar{\pi}$ be the supremum of the flow payoffs earned at any record. We claim first that if an equilibrium prescribes cooperation at any record (whether or not it is evasion-proof), then no player with an unconditional defector record can earn flow payoffs within $1 - \gamma$ of $\bar{\pi}$. This follows from the fact that there are no anti-reciprocators, so at any record a player who deviates to $D$ will receive flow payoff no less than that of players who are prescribed to play $D$. Hence, at any record where cooperation is prescribed, deviating to always playing $D$ increases a player’s instantaneous payoff.
by at least \((1 - \gamma)g\) and reduces her continuation payoff by at most \(\frac{1 - \gamma}{\gamma}g\), and thus constitutes a profitable deviation.

Now let \(k\) be a record where the flow payoff is within \(\frac{1 - \gamma}{\gamma}g\) of \(\bar{\pi}\). Then any evasion-proof equilibrium must prescribe unconditional defection at record \(k\), as playing \(D\) and evading the recording technology to keep one’s record fixed at \(k\) increases one’s instantaneous payoff by at least \((1 - \gamma)g\) and reduces one’s continuation payoff by at most \(\frac{1 - \gamma}{\gamma}g\).

Combining these two observations implies that players must defect at every record in any evasion-proof equilibrium.

Now suppose that, rather than evading the record-keeping system entirely so that her record remains fixed, a player can instead take her action, see what updated record the system assigns her, and then choose whether to “erase” this update and return to record \(k\). This ability to selectively erase records is even more powerful than the ability to evade recording altogether, as a player always has the option of erasing any possible update. Hence, cooperation is again impossible when players can erase records rather than avoiding them.\(^{14}\)

\section{Simple Records and Robust Equilibria}

\subsection{Counting \(D\)’s}

For the rest of the paper, we restrict attention to the following record-keeping system, which we call “noisy recording of \(D\)’s”: Newborn players have record 0. Whenever a player plays \(D\), her record increases by 1. Whenever a player plays \(C\), her record remains constant with probability \(1 - \varepsilon\), and her record increases by 1 with probability \(\varepsilon\).

\(^{14}\)The ability to erase individual records is also more powerful than the ability to re-enter the game under a pseudonym, thus completely clearing one’s record. In the latter case, for some information structures cooperation is possible via strategies where new players must “establish a reputation” for a number of periods before anyone will cooperate with them. See Friedman and Resnick (2001) for an analysis of “pseudonym-proofness” in a random-matching model with perfect monitoring.
Thus, a player’s record is simply a count of the number of times she has been recorded as having played $D$.\footnote{\label{ftn:errors}One could also consider a version of this information system with “two-sided errors,” where a player’s record remains constant with some probability $\varepsilon' > 0$ when she plays $D$. This would complicate notation without significantly affecting our results.}

One motivation for considering this information structure is that counting $D$’s is a simple record-keeping system that can support cooperation. Moreover, the equally simple system that counts $C$’s instead of $D$’s does not allow equilibria with any cooperation at all, as we show in the appendix. Roughly speaking, this is because a player with a very high $C$ count would always play $D$, which by supermodularity implies that she must receive a low payoff, but then there is no incentive to acquire $C$’s. Information structures that count both $C$’s and $D$’s can support cooperation, but are more complicated than counting only $D$’s.

More importantly, we think counting $D$’s is a natural information structure when players can forge records of past interactions but cannot hide true records. In particular, if players can create records of fake interactions where they played $C$, there cannot be an equilibrium where players are rewarded for having played $C$ in the past. Under the plausible assumption that players are also not punished for having played $C$ in the past, equilibrium strategies cannot depend on the number of times a player has played $C$ at all. Thus, we consider record-keeping systems that only count $D$’s.

Even with this minimalist record-keeping system, there can be equilibria that we regard as implausible and unintuitive. For this reason the next two subsections develop refinements on the equilibrium strategies. Here is an example of one type of equilibrium that we wish to rule out: Players with even records (including new players with record 0) are reciprocators and play $C$ if and only if their partner’s record is also even. Players with odd records are unconditional defectors and play $D$ against all partners.

The next result uses strategies that generalize this “even-odd” form. Throughout the paper, we say that \textit{limit efficiency is attainable} if, for every $\eta > 0$, there exists $\bar{\varepsilon} < 1$ and a function $\bar{\gamma} : (0, 1) \to (0, 1)$ such that, whenever $\varepsilon < \bar{\varepsilon}$ and $\gamma > \bar{\gamma}(\varepsilon)$, there...
exists an equilibrium with $V_0 > 1 - \eta$. (That is, limit efficiency refers to the iterated limit where first $\gamma \to 1$ and then $\varepsilon \to 0$.)

**Proposition 1.** Limit efficiency is attainable in strict equilibrium whenever the prisoner’s dilemma is mild ($g < 1$).

The proof of this result is in OA.2. Note that it leaves open whether and how limit efficiency might be possible when the prisoner’s dilemma is severe ($g \geq 1$).

Let us explain why even-odd supports limit efficiency in the case where $g < 1 < l$. If all players use even-odd and the noise in records is small, then most players will be reciprocators, because odd records tend to change quickly. Thus, players’ values are higher at even records, and so regardless of the parameters the prescribed play at odd records is optimal, because it both maximizes short-run payoffs and leads to a transition to a record with a higher payoff. Players with even records face different incentives, depending on their partner’s record. When the partner has an even record (and is thus expected to play $C$), the one-shot gain $g$ from defecting is less than the expected next-period loss of approximately 1 (when $\gamma \approx 1$ and $\varepsilon \approx 0$; note that with probability close to 1 the player’s record in two periods will be even again whether or not she deviates today), so playing $C$ as prescribed is optimal. When instead the partner has an odd record (and thus plays $D$), the one-shot gain from defecting is now $l$, which is greater than the next-period loss of 1, so playing $D$ is optimal.

Now we state a partial converse to Proposition 1.

16Note that $V_0$, the per-period expected payoff of a newborn agent, is also equal to average payoff in the population in every period. This follows because the expected fraction of a player’s lifetime spent at record $k$ is equal to fraction of the population with record $k$ (and there is no discounting, so both $V_0$ and the population-average payoff are given by $\sum_k \mu(k)\pi_k$).

17If $g < l < 1$ then a one-period punishment for playing $D$ is “too harsh,” in that reciprocators then prefer to cooperate against defectors as well as cooperators. In this case, the construction must be modified so that, “on average,” each play of $D$ leads to a number $\rho < 1$ of punishment periods, where $\rho \in (g,l)$. For example, if players are reciprocators with record 0 or 1 mod 3, and unconditional defectors with record 2 mod 3, the average punishment is $\rho = .5$.

18The result can be generalized to state that, for any record-keeping system, Always Defect is the unique strict equilibrium without unconditional cooperators that is robust to forging $C$’s. With the current record keeping-system that counts only $D$’s, all equilibria are trivially robust to forging $C$’s, since a player’s $C$ count is completely irrelevant for her future play and payoffs.
Proposition 2. If the prisoner’s dilemma is severe \((g \geq 1)\) then Always Defect is the unique strict equilibrium such that there is no record at which a player plays \(C\) for all opposing records (i.e., there are no unconditional cooperators).

The introduction gave an intuition for Proposition 2 based on the “snowballing” of \(D\)’s. A more detailed argument shows that, in any non-trivial equilibrium without unconditional cooperators, we must have \(\mu_R \geq g\), where \(\mu_R\) is the fraction of reciprocators in the population (and thus no such equilibrium can exist when \(g > 1\)). Let \(D_k\) be the occupation measure of \(D\)’s faced by a player with record \(k\).\(^{19}\) Without unconditional cooperators, only \((C,C)\) and \((D,D)\) are played on path, so if a player with record \(k\) is a conditional cooperator, \(D_k\) must increase by at least \(g\) whenever she plays \(D\) (otherwise, she would never play \(C\)). Note that, when she matches with a defector, she both faces \(D\) today and plays \(D\) today. Therefore,

\[
D_k \geq (1 - \mu^R) (1 + g). \tag{1}
\]

Moreover, if a player with record \(k\) is an unconditional defector, she faces \(D\) today for sure, and hence \(D_k \geq D_{k+1} \geq \ldots \geq D_{k'}\), where \(k'\) is the least integer greater than \(k\) such that a player with record \(k'\) is a reciprocator. Therefore, Equation 1 must also hold for unconditional defectors, and hence must hold for everyone.

This implies that the total share of \((D,D)\) outcomes is at least \((1 - \mu^R) (1 + g)\). On the other hand, if we assume that all of the reciprocators cooperate with each other (which is not imposed in the actual proof), then when two players match the outcome is \((D,D)\) if and only if at least one of them is an unconditional defector. Hence, we have

\[
1 - (\mu^R)^2 \geq (1 - \mu^R) (1 + g) \iff \mu^R \geq g.
\]

As we will see in Section 6, cooperation in the severe prisoner’s dilemma is nonetheless possible under strategies with unconditional cooperators, such as \textit{GrimKL}.

\(^{19}\)This is the “expected discounted number of \(D\)’s,” defined formally in A.3.
5.2 Forgery-Proofness

Motivated by the idea that players can forge records of past interactions, we restrict attention to equilibria that are robust to adding $D$’s to one’s record.\footnote{In their lending model, Bhaskar and Thomas (2018) use a similar restriction to rule out equilibria where borrowers with two defaults are treated better than borrowers with only one.}

**Definition 2.** An equilibrium is **forgery-proof** if $V_k \geq V_{k'}$ for every $k \leq k'$.

For example, even-odd violates forgery-proofness because a player with an odd record can avoid punishment by forging one interaction where she plays $D$. Note that even-odd has the flavor of the strategy “perfect tit-for-tat,” which says to cooperate today if yesterday both partners played $C$ or both partners played $D$, and to play $D$ if exactly one partner played $C$. Moreover, like even-odd here, perfect tit-for-tat can be a strict equilibrium in the prisoner’s dilemma with fixed pairs and imperfectly observed actions.\footnote{“Ordinary” tit-for-tat is not an equilibrium with fixed pairs, imperfectly observed actions, and generic payoffs.} Both even-odd and perfect tit-for-tat are “forgiving” in the sense that the punishment phases are of finite length. Forgiveness could also be achieved with “repenting” strategies, where it is understood that if one player defects against the other, the defector will cooperate for a few rounds while their partner defects until the books are balanced and both players resume reciprocation. Although forgiving strategies have appealing properties in settings where records cannot be manipulated, these strategies cannot support cooperation in environments where records can be forged.

5.3 Coordination-Proofness

We now impose an additional robustness property, which we call **coordination-proofness**. The idea is that the equilibrium should not use “miscoordination” within a match as a threat to support social cooperation, because such an equilibrium will fall apart if matched partners do not manage to coordinate successfully.
To formalize this property, let \( \hat{u}_k (a, a') \) be the augmented payoff of a player with record \( k \) at action profile \( (a, a') \), given by

\[
\hat{u}_k (a, a') = (1 - \gamma) u(a, a') + \gamma \sum_{k'} r(k, a)[k']V_{k'}.
\]

Given our restriction to strict equilibria, a player with record \( k \) is a reciprocator iff

\[
\hat{u}_k (C, C) - \hat{u}_k (D, C) > 0 > \hat{u}_k (C, D) - \hat{u}_k (D, D).
\]

In other words, the augmented stage game played between any two reciprocators is a coordination game, where both \((C, C)\) and \((D, D)\) are stage-game Nash equilibria. It is straightforward to see that \((C, C)\) is always the Pareto dominant equilibrium of this augmented game: since the gain to defection is always positive, \( \hat{u}_k (C, C) - \hat{u}_k (D, C) > 0 \) implies \( V_k > V_{k+1} \), so for any reciprocator the \((C, C)\) equilibrium yields higher continuation payoffs as well as higher stage-game payoffs. This observation motivates the following definition:

**Definition 3.** An equilibrium is **coordination-proof** if whenever two reciprocators match, they play \((C, C)\).

An example of a strategy profile that satisfies forgery-proofness but violates coordination-proofness is the following: When players with records \( k \) and \( k' \) match, they play \((C, C)\) if \( k + k' < 10 \), and they play \((D, D)\) otherwise. This violates coordination-proofness because any player with record \( k \leq 9 \) is a reciprocator (assuming parameters are such that this strategy profile is an equilibrium), but two players with record \( 9 \) play \( D \) when matched with each other. The spirit of coordination-proofness is that this equilibrium is “unreasonable” because, if two partners with record \( 9 \) could agree to play \((C, C)\) instead of \((D, D)\) (holding fixed everyone else’s strategy), both would have a strict incentive to abide by the agreement.

---

22Because two matched players will never face each other again, the coordination problem here is much simpler than that in dynamic games with a fixed set of partners, as in e.g. Bernheim and Ray (1989), Farrell and Maskin (1989), and Chassang and Takahashi (2011).
Note that coordination-proofness implies that every reciprocator plays \( C \) when matched with another reciprocator or an unconditional cooperator, and plays \( D \) when matched with a defector. In particular, all reciprocators play \( C \) against the same set of opposing records. Therefore, under coordination-proofness, a strategy profile is completed characterized by a description of which records are reciprocators, which are unconditional cooperators, and which are defectors (formally, a mapping \( \sigma : \mathbb{Z}_+ \to \{R, UC, D\} \)).

This characterization simplifies the equilibrium conditions as follows: Fix an equilibrium \((\sigma, \mu)\). Recall that \( \mu_k \) denotes the population share with record \( k \). For each class of records \( z \in \{R, UC, D\} \), let \( \mathbb{K}_z \subset \mathbb{Z}_+ \) be the set of records in class \( z \): that is, \( \mathbb{K}_z = \{k \in \mathbb{Z}_+ : \sigma(k) = z\} \). Then, let \( \mu^z \) denote the share of the population in class \( z \), given by \( \mu^z = \sum_{k \in \mathbb{K}_z} \mu_k \). Finally, we will also use the term cooperator for all players who are either reciprocators or unconditional cooperators (i.e., anyone who is not a defector), and we denote the population share of cooperators by \( \mu^C = \mu^R + \mu^{UC} = 1 - \mu^D \).

The following result establishes how each player’s record stochastically evolves from one period to the next (assuming the player survives to the next period).

**Lemma 2.** Consider a steady state with share of cooperators \( \mu^C \).

1. Before being matched in the current period, the probability that a reciprocator with record \( k \) has record \( k \) at the end of the period is \( (1 - \varepsilon)\mu^C \), and the probability that she has record \( k + 1 \) is \( 1 - (1 - \varepsilon)\mu^C \).

2. Before being matched in the current period, the probability that an unconditional cooperator with record \( k \) has record \( k \) at the end of the period is \( 1 - \varepsilon \) and the probability that she has record \( k + 1 \) is \( \varepsilon \).

3. Before being matched in the current period, the probability that a defector with record \( k \) has record \( k + 1 \) at the end of the period is \( 1 \).

**Proof.** Consider a reciprocator with record \( k \). With probability \( \mu^C \), she will face another reciprocator in the upcoming period. In this case, she will play \( C \), and her record
will remain the same with probability $1 - \varepsilon$ and increase to $k + 1$ with probability $\varepsilon$. With probability $1 - \mu^C$, she will face a defector. In this case, she will play $D$, and her record will increase to $k + 1$ with probability 1. Thus, the probability that the reciprocator’s record remains the same is the probability she faces another reciprocator times the probability that her action of $C$ is correctly recorded, which equals $(1 - \varepsilon)\mu^C$, while the probability that her record increases to $k + 1$ is the complementary probability of $1 - (1 - \varepsilon)\mu^C$.

An unconditional cooperator with record $k$ plays $C$ regardless of her opponent’s record, which causes her record to remain the same with probability $1 - \varepsilon$ and to increase to $k + 1$ with probability $\varepsilon$.

A defector always plays $D$, which always increases her record by 1. ■

We now derive the incentive constraints. We use the following notation: for a player with record $k$, $(C|C)_k$ is the condition that $C$ is the best response to $C$, $(D|D)_k$ is the condition that $D$ is the best response to $D$, $(C|D)_k$ is the condition that $C$ is the best response to $D$, and $(D|C)_k$ is the condition that $D$ is the best response to $C$. Note that, if a player with record $k$ is a reciprocator, $(C|C)_k$ and $(D|D)_k$ must both hold; if a player with record $k$ is an unconditional cooperator, $(C|D)_k$ must hold (which also implies $(C|C)_k$, by supermodularity); and if a player with record $k$ is a defector, $(D|C)_k$ must hold (which also implies $(D|D)_k$).

**Lemma 3.**

1. The $(C|C)_k$ constraint is

   $$\gamma(1 - \varepsilon)(V_k - V_{k+1}) > (1 - \gamma)g.$$  

   The $(D|C)_k$ constraint is the opposite inequality.

2. The $(D|D)_k$ constraint is

   $$\gamma(1 - \varepsilon)(V_k - V_{k+1}) < (1 - \gamma)l.$$
The \((C|D)_k\) constraint is the opposite inequality

**Proof.** Consider a player with record \(k\). When she plays \(C\), her expected continuation payoff is \((1 - \varepsilon)V_k + \varepsilon V_{k+1}\), while it is instead \(V_{k+1}\) when she plays \(D\). Thus, the \((C|C)_k\) constraint is

\[
1 - \gamma + \gamma(1 - \varepsilon)V_k + \gamma \varepsilon V_{k+1} > (1 - \gamma)(1 + g) + \gamma V_{k+1},
\]

which is equivalent to the expression in Part 1 of the result. Likewise, the \((D|D)_k\) constraint is

\[
\gamma V_{k+1} > -(1 - \gamma)l + \gamma(1 - \varepsilon)V_k + \gamma \varepsilon V_{k+1},
\]

which is equivalent to the expression in Part 2. Clearly, \((D|C)_k\) is the opposite of \((C|C)_k\), and \((C|D)_k\) is the opposite of \((D|D)_k\).

\[\blacksquare\]

6 **GrimKL Strategies**

6.1 Combining Forgery-Proofness and Coordination-Proofness

The main result of this subsection characterizes the unique family of strategies that can be equilibria that satisfy both forgery-proofness and coordination-proofness. This is the family of GrimKL strategies, which is defined for non-negative integers \(K\) and \(L\) so that a player is a reciprocator for the first \(K\) records, \(0 \leq k \leq K - 1\), an unconditional cooperator for the next \(L\) records, \(K \leq k \leq K + L - 1\), and a defector for all other records, \(k \geq K + L\). Note that GrimKL with \(K = L = 0\) is the familiar Always Defect strategy, and GrimKL with \(L = 0\) is GrimK: here a player is a reciprocator for the first \(K\) records, \(0 \leq k \leq K - 1\), and a defector for all other records, \(k \geq K\), but never an unconditional cooperator.\(^{23}\)

Figure 2 depicts the steady state record shares, \(\mu_k\), and the value functions, \(V_k\), in a GrimKL equilibrium for \(\gamma = .95, \varepsilon = .05, g = .5, \text{and } l = 1\); here \(K = 4\) and

\(^{23}\)Note that Always Defect is a special case of GrimK as well as GrimKL.
Figure 2: Steady state record shares and valuation functions for Grim41.

$L = 1$. In this equilibrium, $\mu^R \approx .492$, $\mu^{UC} \approx .260$, and $\mu^K \approx .247$. We note that there is no non-trivial GrimK equilibrium for these parameters, which can be shown using the characterization of GrimK equilibria in A.6. Intuitively, the fact that players with $k = 4$ always play $C$ implies that transitions from record 4 to record 5 are unlikely, which keeps the total share of cooperators high enough for a non-trivial GrimKL equilibrium to exist, whereas with GrimK strategies transitions to the defector region occur too quickly to support positive steady-state cooperation.

We now present the characterization result.

**Theorem 4.** Any strict equilibrium that satisfies forgery-proofness and coordination-proofness corresponds to GrimKL for some $K, L$. Moreover, if for some $K, L$ GrimKL is a strict equilibrium, it satisfies forgery-proofness and coordination-proofness.

The result can be understood in three steps. First, forgery-proofness implies that there is a critical record $\bar{k}$ such that all records $k \geq \bar{k}$ are defectors. Intuitively, this follows because, if there were infinitely many cooperator records, a player could profitably deviate by always playing $D$ and then inflating her record to the next cooperator record. Second, this critical record $\bar{k}$ can be chosen so that all records $k < \bar{k}$ are either reciprocators or unconditional cooperators, because whenever there is a defector record that is followed by a cooperator record, a player at the defector record could profitably
deviate by inflating her record to the next cooperator record. The third and last step is to classify the first $\bar{k}$ records as reciprocators or unconditional cooperators. Note that a player’s opponent plays the same way towards her whether she is a reciprocator or an unconditional cooperator. Thus, all incentives to play $C$ for players with records $k < \bar{k}$ come from avoiding the “punishment” of reaching record $\bar{k}$ and triggering an increase in the fraction of partners who will play $D$. Since the survival probability $\gamma$ is less than 1, this penalty looms larger the closer a player’s record is to $\bar{k}$. Hence, players with larger records are willing to incur greater costs to prevent their records from rising further. This implies that there is a critical record $k^* < \bar{k}$ such that only players with records greater than $k^*$ are willing to play $C$ at a cost of $l$ (while players with records less than $k^*$ are willing to play $C$ at a cost of $g$, but not at a cost of $l$). We conclude that it is those players with lower records who must be reciprocators, which yields a $GrimKL$ profile with $K = k^*$ and $L = \bar{k} - k^*$.

The following two subsections will analyze the maximum level of cooperation that can be supported by the family of $GrimKL$ strategies. We first define the maximal equilibrium level of cooperation for both the $GrimKL$ and $GrimK$ strategy families. Let $\mu_{KL}^C(\gamma, \varepsilon)$ be the maximal share of cooperators that is attained in any $GrimKL$ equilibrium for parameters $\gamma$ and $\varepsilon$:

$$\mu_{KL}^C(\gamma, \varepsilon) = \sup\{\mu^C : \mu^C \text{ is the share of cooperators in a } GrimKL \text{ equilibrium}\}.$$

Likewise, let $\mu_K^C(\gamma, \varepsilon)$ be the maximal share of cooperators that is attained in any $GrimK$ equilibrium for parameters $\gamma$ and $\varepsilon$:

$$\mu_K^C(\gamma, \varepsilon) = \sup\{\mu^C : \mu^C \text{ is the share of cooperators in a } GrimK \text{ equilibrium}\}.$$

### 6.2 GrimK Strategies

We first focus on the case of $GrimK$, where unconditional cooperators are not present.
Theorem 5.

1. For $g < 1$ and $l > g/(1 - g)$,

$$\lim_{\varepsilon \to 0} \lim_{\gamma \to 1} \mu_{K}^C(\gamma, \varepsilon) = \frac{l}{1 + l}.$$

2. For $g < 1$ and $l \leq g/(1 - g)$,

$$\lim_{\varepsilon \to 0} \lim_{\gamma \to 1} \mu_{K}^C(\gamma, \varepsilon) = 0.$$

3. For $g \geq 1$, $\mu_{K}^C(\gamma, \varepsilon) = 0$ for all $(\gamma, \varepsilon) \in (0, 1) \times (0, 1)$.

Part 3 of Theorem 5 says that no cooperation is possible in a GrimK equilibrium whenever $g \geq 1$, regardless of $\gamma$ and $\varepsilon$. Thus, Always Defect (i.e. Grim0) is the unique equilibrium in this family in a severe prisoner’s dilemma. This is a consequence of Proposition 1, since with GrimK there are no unconditional cooperators.

Part 1 of Theorem 5 says that in a mild prisoner’s dilemma with $g < 1$ and $l > g/(1 - g)$, GrimK strategies can support strictly positive cooperation, and moreover that the maximal level of cooperation tends to $l/(1 + l)$ in the iterated limit where $\gamma$ approaches 1 and then $\varepsilon$ approaches 0. Part 2 says that for $g < 1$ and $l < g/(1 - g)$ the level of cooperation approaches 0 in the iterated limit. The reason a higher $l$ makes cooperation easier to sustain in a GrimK equilibrium is that it makes it easier to satisfy the ($D|D$) constraints for higher values of $\mu^C$.24

To understand this result, note that the ($D|D$)$_{K-1}$ constraint may be written as

$$\mu^C < \frac{1}{\gamma(1 - \varepsilon)} \frac{l}{1 + l}.$$

24This is essentially the same reason that a high value of $l$ is necessary for contagion strategies to form an equilibrium in Kandori (1992).
This comes from combining the recursive equation for $V_{K-1}$,

$$V_{K-1} = (1 - \gamma)\mu^C + \gamma\mu^C (1 - \varepsilon) V_{K-1}$$

(which follows because, once a player with record $K-1$ gains another $D$ in her record, she receives payoff 0 forever), with the $(D|D)_{K-1}$ constraint

$$0 > (1 - \gamma)(-l) + \gamma(1 - \varepsilon)V_{K-1}.$$ 

As $\gamma \to 1$ and $\varepsilon \to 0$, Equation 2 yields the upper bound $\mu^C < \frac{l}{1+l}$.

On the other hand, since $GrimK$ does not involve unconditional cooperators, by the discussion following Proposition 2 we must have $\mu^C > g$. There exists such a value of $\mu^C$ that also satisfies the upper bound $\mu^C < \frac{l}{1+l}$ if and only if $l > \frac{g}{1-g}$.

We note that $Grim2$ provides a simple example of a strategy profile with multiple associated steady states. For example, if $\gamma = .8$ and $\varepsilon = .01$ then there are three steady-state values of $\mu^C$ under $Grim2$, given by $\mu^C \approx .6738$, $\mu^C \approx .8536$, and $\mu^C \approx .9979$. Additionally, for $g < .171$ and $l > 3.77$, $Grim2$ is an equilibrium for all three of these values of $\mu^C$.

### 6.3 General $GrimKL$ Strategies

We now turn to the full family of $GrimKL$ strategies, where $L$ can be greater than 0. Let $b : \mathbb{R}_+ \to \mathbb{R}$ be the function given by

$$b(g) = \min \left\{ \frac{1 + g}{|\ln(1 + g) - 1|}, 21.9223 - 3.57143g \right\}. \overset{25}{=}$

**Theorem 6.**

1. For $l > \max\{g(g + 1), b(g)\}$, $\lim_{\varepsilon \to 0} \liminf_{\gamma \to 1} \mathcal{P}_K^{\gamma \varepsilon} = 1$.

2. For $l \leq g(g + 1)$, $\lim_{\gamma \to 1} \mathcal{P}_K^{\gamma \varepsilon} = 0$ for all $\varepsilon \in (0, 1)$.

\[^{25}21.9223 - 3.57143g \text{ is the approximate form of } (3e^\phi - 2 - 2g)/(\phi - 1) \text{ when } \phi = 1.56.\]
3. For $g \geq 1$ and $l \leq g(g + 1)$, $\overline{C}_{KL}(\gamma, \varepsilon) = 0$ for all $(\gamma, \varepsilon) \in (0, 1) \times (0, 1)$.

Remark 1: Interpreting the parametric conditions of Theorem 6. Part 3 of Theorem 6 says that, when $g \geq 1$ and $l \leq g(g + 1)$ so that the prisoner’s dilemma is both severe and mildly supermodular, Always Defect is the unique equilibrium in the GrimKL family. Part 2 says that whenever the prisoner’s dilemma is mildly supermodular, the maximal level of cooperation approaches 0 as $\gamma \to 1$, for any value of $\varepsilon$, regardless of whether the prisoner’s dilemma is severe.

In contrast, Part 1 of Theorem 6 says that when the prisoner’s dilemma is sufficiently supermodular, the maximal level of cooperation with GrimKL strategies tends to 1 in the iterated limit. In particular, for sufficiently severe prisoner’s dilemmas with $g(g + 1) > b(g)$ (i.e., $g \geq 2.858$), limit efficiency is achieved if $l > g(g + 1)$. For smaller values of $g$, we use the stronger requirement that $l > b(g)$ to guarantee limit efficiency. We do not know whether this condition is necessary for limit efficiency; it comes from the fact that $K$ and $L$ must be integers, which affects the feasibility of generating particular values of $\mu^C$, $\mu^R$, $\mu^{UC}$, as is discussed in A.7. Our results do not say much about cooperation in the iterated limit when $l > g(g + 1)$ holds, but $l > b(g)$ does not. However, if integer problems could be ignored, then $l > g(g + 1)$ would again suffice for limit efficiency. If this were the case, for any parameters such that the GrimK family permits cooperation in the iterated limit, the GrimKL family would enable cooperation in the iterated limit with strictly smaller values of $l$ because $g(g + 1) < g/(1 - g)$ for all $g < 1$. Figure 3 presents a graphical illustration of the performance of GrimKL and GrimK in the iterated limit.

Remark 2: Why limit efficiency requires $l > g(g + 1)$. The inequality $l > g(g + 1)$ comes from combining the constraints $\mu^R > 1 - \frac{1}{1 + g}$ and $\mu^R < 1 - \frac{g}{l}$. We now explain each of these in turn.

The condition $\mu^R > 1 - \frac{1}{1 + g}$ is necessary for the flow payoff of an unconditional cooperator to exceed the flow payoff of a defector, which in turn is necessary for GrimKL to be an equilibrium. (Otherwise, unconditional cooperators would pre-
Figure 3: The performance of the maximally efficient GrimKL and GrimK strategies in the iterated limit where $\gamma$ approaches 1 and then $\varepsilon$ approaches 0. GrimKL attains limit efficiency in the green regions and has limit cooperation of 0 in the red region. Our results do not pin down the limit performance of GrimKL in the blue regions. Solid shading indicates that GrimK attains strictly positive limit cooperation; vertical lines indicate that GrimK has limit cooperation of 0.

In particular, unconditional cooperators earn a higher flow payoff than defectors if and only if $\mu^C - (1 - \mu^C)l > \mu^{UC}(1 + g)$, and this is equivalent to $\mu^R(1 + l) + \mu^{UC}(l - g) > l$. The left-hand side of this inequality is increasing in $\mu^{UC}$, so a necessary condition is that it holds when $\mu^{UC}$ takes on its highest possible value of $1 - \mu^R$. Thus, substituting $\mu^{UC} = 1 - \mu^R$, the inequality becomes $\mu^R(1 + l) + (1 - \mu^R)(l - g) \geq l$, which is equivalent to $\mu^R \geq 1 - \frac{1}{1+g}$.

The condition $\mu^R < 1 - \frac{g}{l}$ is necessary for $(C|C)_0$ and $(D|D)_{K-1}$ to be satisfied simultaneously. In particular, the probability that a player survives to reach record $K - 1$ is the ratio of the benefit of playing $C$ with record 0 to the benefit of playing $C$ with record $K - 1$, and it is also approximately equal to $1 - \mu^R$, the share of players with record $\geq K$.\textsuperscript{26} Thus,

$$\frac{\text{benefit of playing } C \text{ at } k = 0}{\text{benefit of playing } C \text{ at } k = K - 1} \approx 1 - \mu^R.$$

\textsuperscript{26}This follows because $1 - \mu^R$ is the probability that a player survives to reach record $L$, and when $\gamma$ is large relative to $\varepsilon$ these two probabilities are almost the same.
On the other hand, we have

\[
\frac{\text{cost of playing } C \text{ vs. } C}{\text{cost of playing } C \text{ vs. } D} = \frac{g}{l}.
\]

Now, for it to be optimal both to play \( C \) vs. \( C \) with record \( k = 0 \) and to play \( D \) vs. \( D \) with record \( k = K - 1 \), it must be the case that

\[
\frac{\text{benefit of playing } C \text{ at } k = 0}{\text{cost of playing } C \text{ vs. } C} > \frac{\text{benefit of playing } C \text{ at } k = K - 1}{\text{cost of playing } C \text{ vs. } D}.
\]

Combining these observations, we have

\[
1 - \mu^R > \frac{g}{l} \iff \mu^R < 1 - \frac{g}{l}.
\]

Note also that the constraint \( \mu^R < 1 - \frac{g}{l} \) implies that, in any equilibrium that supports \( \mu^C \approx 1 \), we must have \( \mu^{UC} > \frac{g}{l} \). Thus, unless \( l \) is much larger than \( g \), there must be a substantial share of unconditional cooperators in any efficient equilibrium.

**Remark 3: The role of unconditional cooperators.** Theorem 6 demonstrates the power of unconditional cooperation in sustaining equilibrium cooperation. Intuitively, unconditional cooperators help support cooperation in two ways. First, there is the relatively straightforward effect that turning those reciprocators who are most tempted to play \( C \) against \( D \) into unconditional cooperators relaxes the \((D|D)\) constraint. This explains why \( \text{GrimKL} \) can sometimes support almost full cooperation, while \( \text{GrimK} \) can never support \( \mu^C > \frac{1}{1+1} \).

Second, there is the novel effect that the presence of unconditional cooperators increases the steady-state share of cooperators by slowing the transition from cooperator records to defector records, and in particular cuts off the snowballing \( D \)'s that can destroy the possibility of positive steady-state cooperation in severe prisoner’s dilemmas. To see that this effect is distinct from the benefit of relaxing the \((D|D)\) constraint, note that the \((C|C)\) constraint in Proposition 4 cannot be satisfied if \( g \geq 1 \). That is, if \( g \geq 1 \) then \( \text{GrimK} \) strategies cannot support positive cooperation even in the relaxed prob-
lem where players’ incentives when their opponents play D are completely ignored.\textsuperscript{27} Thus, this benefit of unconditional cooperators has nothing to do with relaxing the (D|D) constraint.

Remark 4: Proving Theorem 6. The proof of Theorem 6 is more involved than that of Theorem 5. The main additional complication is that the cost of increasing one’s record by 1 is very similar for a player with record $K - 1$ and a player with record $K$, which makes it hard to find $K$ and $L$ that simultaneously satisfy the $(D|D)_{K-1}$ and $(C|D)_K$ constraints. This issue is also the source of the integer problems that lead to the requirement that $l > b(g)$ in Part 1 of Theorem 6. We provide a detailed proof outline at the beginning of A.7.

7 Conclusion

We conclude with some observations about extensions and alternative models.

Other types of equilibria. With the classes of strategies we have discussed so far, it is necessary that either $g < 1$ or $l > g(g + 1)$ for any cooperation to occur in equilibrium. This is not true in general if forgery-proofness and coordination-proofness are not imposed. Consider the family of strategies defined for four non-negative integers, $K_1$, $K_2$, $K_3$, and $K_4$, as follows: A player is a reciprocator for the first $K_1$ records, a defector for the next $K_2$ records, a reciprocator again for the next $K_3$ records, an unconditional cooperator for the next $K_4$ periods, and a defector for all higher records. We show in OA.10 that strategies in this family can sometimes enable cooperation in equilibrium even when $g > 1$ and $l < g(g + 1)$. Our intuition for this result is that, by splitting the set of defector records into two classes, these strategies accelerate the punishment for acquiring $D$'s and thus strengthen incentives for the first class of reciprocators. However, players in the first defector class would gain by forging $D$'s.

Stochastic transitions. We have restricted attention to equilibria with determinis-

\textsuperscript{27}Moreover, the proof of the more general Proposition 2 uses only the $(C|C)$ constraints and not the $(D|D)$ constraints.
tic record transitions. Stochastic transitions resemble public randomization and can improve on GrimK and GrimKL strategies in at least two ways (assuming the transition probability can be appropriately tuned). First, a stochastic version of GrimK can attain full limit efficiency when \( g < 1 \): the intuition is that making transitions from the reciprocator state to the defector state stochastic relaxes the \((D|D)\) constraint, which was responsible for the \( \mu^C \leq \frac{1}{1+l} \) upper bound for deterministic strategies, as well as the requirement that \( l > g/(1-g) \).\(^{28}\) Second, a stochastic version of GrimKL can avoid the integer problems that arose in Theorem 6, and thus relax the condition for limit efficiency from \( l > \max\{g(g+1), b(g)\} \) to \( l > g(g+1) \).\(^{29}\) We demonstrate these results in OA.11, using transition probabilities between records that depend on the parameters \( \varepsilon \) and \( \gamma \).\(^{30}\)

**Higher-order information.** We have also restricted to record-keeping systems that use only first-order information. If the update of a player’s record can depend on her opponent’s action as well as her own, supporting cooperation becomes much easier: for example, if a player’s record increases only when she plays \( D \) versus \( C \), this cuts off the “snowballing” problem, and we prove in OA.12 that limit efficiency can be attained by GrimK strategies for any values of \( g \) and \( l \), even those with \( g > l \). This shows that the key limitation of first-order information is that, when a player sees that her opponent has played a certain number of \( D \)’s, she cannot tell if these plays correspond to “selfish” deviations or “deserved” punishments.\(^{31}\)

**More general interaction structures.** Our model assumes that each player is matched

\(^{28}\)We prove this with a strategy with only two records, where record-0 players are reciprocators and record-1 players are defectors.

\(^{29}\)We prove this with a strategy with three records, where record-0 players are reciprocators, record-1 players are unconditional cooperators, and record-2 players are defectors. As with GrimKL, the defectors remain defectors forever: the difference is that playing \( D \) causes one’s record to increase with probability less than 1. We do not know whether more complex stochastic transition rules could allow limit efficiency for a wider range of parameters.

\(^{30}\)These results consider the problem of designing an optimal record-keeping system within some class, while the rest of this paper studies a particular simple and natural record-keeping system. See Bhaskar and Thomas (2018) for an information-design perspective in a lending game model.

\(^{31}\)In a related result, Heller and Mohlin (2018) show that, in their model, full cooperation is always possible when players observe only their partners’ past actions when facing \( C \).
to play the game every period. The same model describes the steady states when some constant non-zero share of players is selected at random to play each period, with $\gamma$ now interpreted as the probability of surviving for one more interaction. If different players are matched to play the game at different frequencies, the steady-state equations would be the same, but different players would face different incentive constraints. This extension seems interesting but we do not cover it here, except to note that, since our equilibria are strict, they are robust to small differences in interaction frequencies.

**Different stage games.** Our model of random-matching with first-order information could also be used to study stage games other than the prisoner’s dilemma. It is also easy to adapt our model to settings with multiple populations of players. For example, suppose a population of player 1’s and a population of player 2’s repeatedly play the product choice game, where only player 1 faces binding moral hazard at the efficient action profile (and player 2 wants to match player 1’s action). In this game, efficient payoffs can always be supported in the limit where first $\delta \to 1$ and then $\varepsilon \to 0$, using the following adaptation of GrimK strategies: in each match, both partners play $C$ if player 1’s record is $k < K$, and both play $D$ if player 1’s record is $k \geq K$. This example suggests our model might be especially tractable in settings with 1-sided moral hazard.
References


A.1 Proof of Theorem 1

Let $B$ be a bound on the update supports. Without loss, relabel records so that two players with different ages can never share the same record. Then let $R(t)$ be the set of feasible records for a player of age $t$, and let $\bar{M} = \{ \mu \in M : \sum_{k \in R(t)} \mu(k) \leq \gamma^t \forall t \in \mathbb{Z} \}$. Clearly, $\bar{M}$ is convex. Section OA.1 establishes the following three claims:

**Claim 1.** $\bar{M}$ is compact in the $\sup$ norm topology.

**Claim 2.** $f_{r,s}$ maps $\bar{M}$ to itself.

**Claim 3.** $f_{r,s}$ is continuous in the $\sup$ norm topology.

Since every normed space is a locally convex Hausdorff space, the theorem follows from Corollary 17.56 (page 583) of Aliprantis and Border (2006). The key point is that
the combination of bounded-support updates and geometric player lifetimes means that the tails of the feasible population states must be thin. This explains why \( \bar{M} \) is compact, and also plays a key role in ensuring that \( f \) is continuous.

### A.2 Impossibility for Counting C’s

We show cooperation cannot be supported when a player’s record is the number of times she has played \( C \) (where a play of \( C \) is recorded as \( D \) with probability \( \varepsilon \)). We continue to assume \( g < l \), as otherwise Theorem 2 already implies that cooperation is impossible.

**Proposition 3.** When records count \( C \)’s, the unique strict equilibrium is Always Defect.

*Proof.* Note that, in any strict equilibrium, if records \( k \) and \( k + 1 \) are both reached with positive probability, then \( V_k \leq V_{k+1} \). For if \( V_k > V_{k+1} \), then a player with record \( k \) always plays \( D \) (since this maximizes both her flow payoff and her continuation payoff), but since plays of \( D \) are always recorded accurately this implies no player ever reaches record \( k + 1 \).

We now show that there exists \( K^* \) such that a player with record \( k > K^* \) always defects. To see this, let \( V^* = \sup_k V_k \). For a player with record \( k \), the gain in continuation value from playing \( C \) is at most \( V^* - V_k \), while the gain in flow utility from playing \( D \) is at least \( (1 - \gamma)g \). Hence, whenever \( V_k > V^* - (1 - \gamma)g \), a player with record \( k \) always defects. Since \( V_k \) is non-decreasing, there exists \( K^* \) such that a player with record \( k > K^* \) always defects.

By supermodularity, everyone other than unconditional cooperators play \( D \) against a player with record \( k > K^* \). Therefore, \( V_{K^*+1} = V^* \), and this is precisely the payoff a player receives from always playing \( D \) while facing \( C \) from unconditional cooperators and \( D \) from everyone else. Note that this is also a lower bound on the payoff attained by a player who always plays \( D \) from the beginning of the game. Therefore, \( V_0 \geq V^* \).
which since $V_k$ is non-decreasing implies $V_0 = V^*$. But this implies that $V_k$ is constant, so playing $D$ is always optimal. \[\square\]

The above argument uses the fact that plays of $D$ are perfectly recorded to establish that $V_k$ is non-decreasing. However, as we show in Section OA.3, the same conclusion applies even if playing $D$ causes the record to increase with probability $1 - \varepsilon' < 1$.

### A.3 Proof of Proposition 2

For $k, t \in \mathbb{N}$, let $\pi_{k,t}$ be the probability that a player with record $k$ will play $C$ in $t$ periods. Since only $(C,C)$ and $(D,D)$ are played on-path, this equals the probability that a player with record $k$ will face $C$ (and receive payoff 1) in $t$ periods from now; otherwise, she receives payoff 0 that period, so $V_k = (1 - \gamma) \sum_{t=0}^{\infty} \gamma^t \pi_{k,t}$.

Let $D_k$ be the occupation measure of $D$'s faced by a player with record $k$, given by $D_k = (1 - \gamma) \sum_{t=0}^{\infty} \gamma^t (1 - \pi_{k,t})$.

Note that, for each $k$, $V_k = 1 - D_k$, and in particular $V_k - V_{k+1} = D_{k+1} - D_k$. In addition, note that the steady-state fraction of pairs that play $(C,C)$ is $\sum_{k \in \mathbb{N}} \mu_k V_k = \sum_{k \in \mathbb{N}} \mu_k (1 - D_k) = 1 - \sum_{k \in \mathbb{N}} \mu_k D_k$. Hence, the steady-state fraction of pairs that play $(D,D)$ is $\sum_{k=0}^{\infty} D_k \mu_k$. Let $\mathbb{K}_R$ denote the set of reciprocator records, and let $\mu^R = \sum_{k \in \mathbb{K}_R} \mu_k$ be the share of reciprocators. Let $\mu^R_k$ be the share of the population that plays $C$ against a player with record $k$. Note that, since only $(C,C)$ and $(D,D)$ are played on path, $\mu^R_k$ equals the probability that a player with record $k$ plays $C$. In particular $\mu^R_k > 0$ iff $k$ is a reciprocator, and $\mu^R_k \leq \mu^R$ for all $k \in \mathbb{N}$. Note also that the fraction of pairs that play $(C,C)$ equals $\sum_{k \in \mathbb{K}_R} \mu_k \mu^R_k$. Combining these observations with the fact that only $(C,C)$ and $(D,D)$ are played), we have

$$1 - \sum_{k \in \mathbb{K}_R} \mu_k \mu^R_k = \sum_{k=0}^{\infty} D_k \mu_k.$$

Now, suppose a player with record $k$ is a reciprocator: $k \in \mathbb{K}_R$. The corresponding
\( (C|C)_k \) constraint is
\[
V_k - V_{k+1} > \frac{1}{1-\varepsilon} \frac{1-\gamma}{\gamma} g.
\]
As \( V_k = 1 - D_k \), this is equivalent to
\[
D_{k+1} - D_k > \frac{1}{1-\varepsilon} \frac{1-\gamma}{\gamma} g. \tag{3}
\]

On the other hand, by definition of \( D_k \) we have
\[
D_k = (1 - \mu^R_k) [1 - \gamma + \gamma D_{k+1}] + \mu^R_k \gamma [(1 - \varepsilon) D_k + \varepsilon D_{k+1}].
\]
This follows because a reciprocator faces \( D \) with probability \( 1 - \mu^R_k \), while she gets an extra \( D \) in her record with probability \( 1 - \mu^R_k (1 - \varepsilon) \). Grouping terms, we obtain
\[
(1 - \gamma) D_k = (1 - \gamma) (1 - \mu^R_k) + \gamma (1 - \mu^R_k (1 - \varepsilon)) (D_{k+1} - D_k)
\]
\[
> (1 - \gamma) (1 - \mu^R_k) + \gamma (1 - \mu^R_k (1 - \varepsilon)) \left( \frac{1}{1-\varepsilon} \frac{1-\gamma}{\gamma} g \right) \quad \text{by 3}
\]
\[
= (1 - \gamma) (1 - \mu^R_k) + (1 - \gamma) \left( \frac{1}{1-\varepsilon} - \frac{\mu^R_k}{\mu^R_k} \right) g
\]
\[\iff\]
\[
D_k > 1 + \frac{g}{1-\varepsilon} - (1 + g) \mu^R_k. \tag{4}
\]

Next, we claim that, if \( k \not\in \mathbb{K}_R \), then
\[
D_k > 1 + \frac{g}{1-\varepsilon} - (1 + g) \bar{\mu}^R.
\]
(That is, (4) holds with \( \mu^R_k \) replaced by \( \bar{\mu}^R \).) To see this, note that \( V_k \) non-increasing implies \( D_k \) non-decreasing. So it suffices to establish (5) for \( k = 0 \). In particular, since \( \mu^R_k \leq \bar{\mu}^R \) for all \( k \in \mathbb{N} \), it suffices to establish that a player with record 0 must be a reciprocator (under the maintained hypothesis that \textit{someone} is a reciprocator). This follows because, if a player with record 0 is a defector, then given that there are no

42
unconditional cooperators in the population he must have a lower continuation payoff than the first record $k$ that is a reciprocator (as the flow payoff is 0 until the record reaches this $k$). But this contradicts $V_k$ non-increasing.

We thus have (4) for every $k \in \mathbb{K}_R$ and (5) for every $k \notin \mathbb{K}_R$. Since $\sum_{k \in \mathbb{N}} \mu_k = 1$ and $\sum_{k \in \mathbb{K}_R} \mu_k = 1 - \bar{\mu}^R$, we have

$$\sum_{k=0}^{\infty} D_k \mu_k > 1 + \frac{g}{1 - \varepsilon} - (1 + g) \left( \sum_{k \in \mathbb{K}_R} \mu_k \mu^R_{-k} + (1 - \bar{\mu}^R) \bar{\mu}^R \right).$$

Combining this with (A.3), we have

$$1 - \sum_{k \in \mathbb{K}_R} \mu_k \mu^R_{-k} > 1 + \frac{g}{1 - \varepsilon} - (1 + g) \left( \sum_{k \in \mathbb{K}_R} \mu_k \mu^R_{-k} + (1 - \bar{\mu}^R) \bar{\mu}^R \right) \iff g \sum_{k \in \mathbb{K}_R} \mu_k \mu^R_{-k} > \frac{g}{1 - \varepsilon} - (1 + g) \left( 1 - \bar{\mu}^R \right) \bar{\mu}^R.$$

Since $\mu^R_k \leq \bar{\mu}^R$ for all $k \in \mathbb{N}$ and $\sum_{k \in \mathbb{K}_R} \mu_k = \bar{\mu}^R$, this implies

$$g \left( \bar{\mu}^R \right)^2 > \frac{g}{1 - \varepsilon} - (1 + g) \left( 1 - \bar{\mu}^R \right) \bar{\mu}^R \iff$$

$$- \left( \mu^R \right)^2 + \mu^R (1 + g) - \frac{g}{1 - \varepsilon} > 0.$$

If $g \geq 1$ this inequality cannot be satisfied for any $\mu^R \in [0, 1]$, so there cannot be any reciprocators, and the unique equilibrium is *Always Defect*.

**A.4 Lemmas about Coordination-Proof Equilibria**

Define $i_k$ to be the total inflow of players into record $k$. This equals $1 - \gamma$ for $k = 0$ since all newborn players have record $k = 0$, while $i_k$ equals the share of players that move into record $k$ from record $k - 1$ in a period for $k > 0$. Define $\tau_k$ to be the total
share of players that transition from record $k$ to record $k + 1$ in a period. Finally, define $\delta_k$ to be the total share of players in record $k$ that die in a period. The steady state equation for record $k$ is

$$\delta_k + \tau_k = i_k,$$

which guarantees that the total inflow of players into record $k$ precisely balances with the total outflow of players from record $k$.

**Lemma 4.** Consider an equilibrium with total share of cooperators $\mu^C$.

1. $\delta_k = (1 - \gamma)\mu_k$.

2. $i_0 = 1 - \gamma$ and $i_{k+1} = \tau_k$ for all $k$.

**Lemma 5.** Consider a steady state with total share of cooperators $\mu^C$.

1. If $k$ is a reciprocator record, then $\tau_k = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_k$.

2. If $k$ is an unconditional cooperator record, then $\tau_k = \gamma\varepsilon\mu_k$.

3. If $k$ is a defector record, then $\tau_k = \gamma\mu_k$.

Part 1 of Lemma 4 follows because each player dies in a given period with probability $\gamma$, independently of every other living player. Part 2 is a consequence of the count $D$’s record-keeping system. Lemma 5 comes from Lemma 2 and a player’s play and survival being independent.

It will also be useful to compute the probability that a player with record $k$ increases her record to $k + 1$ before dying. When $k$ is a defector record, this probability simply equals $\gamma$, the probability of survival to the next period. When $k$ is an unconditional cooperator record, this value is $\sum_{n=0}^{\infty}(\gamma(1 - \varepsilon))^n\gamma\varepsilon$, since from Lemma 2 the probability that a record $k$ player becomes a record $k + 1$ player exactly $n$ periods from now is $(\gamma(1 - \varepsilon))^n\gamma\varepsilon$. Straightforward algebra shows that

$$\sum_{n=0}^{\infty}(\gamma(1 - \varepsilon))^n\gamma\varepsilon = \frac{\gamma\varepsilon}{1 - \gamma(1 - \varepsilon)} := \alpha(\gamma, \varepsilon).$$
Similarly, when $k$ is a reciprocator record, the probability that a player with record $k$ increases her record to $k+1$ before dying is $\beta(\gamma, \varepsilon, \mu^C)$, where $\beta : (0, 1) \times (0, 1) \times [0, 1] \to (0, 1)$ is given by

$$\beta(\gamma, \varepsilon, \mu^C) = \frac{\gamma(1 - (1 - \varepsilon)\mu^C)}{1 - \gamma(1 - \varepsilon)\mu^C}.$$ \[32\]

These probabilities play a key role in deriving the continuation value function.

**Lemma 6.** Consider a steady state with total share of cooperators $\mu^C$. Consider the value function of a player with record $k$.

1. If $k$ is a reciprocator record, $V_k = (1 - \beta(\gamma, \varepsilon, \mu^C))\mu^C + \beta(\gamma, \varepsilon, \mu^C)V_{k+1}$.

2. If $k$ is an unconditional cooperator record, $V_k = (1 - \alpha(\gamma, \varepsilon))(\mu^C - \mu^{DL}) + \alpha(\gamma, \varepsilon)V_{k+1}$.

3. If $k$ is a defector record, $V_k = (1 - \gamma)\mu^{UC}(1 + g) + \gamma V_{k+1}$.

**Proof.** The expected flow payoff of a reciprocator is $\mu^C$. Combining this with Lemma 2 yields $V_k = (1 - \gamma)\mu^C + \gamma(1 - \varepsilon)V_k + \gamma(1 - (1 - \varepsilon)\mu^C)V_{k+1}$. Solving for $V_k$ gives Part 1 of the result.

The expected flow payoff of an unconditional cooperator is $\mu^C - \mu^{DL}$. Combining this with Lemma 2 yields $V_k = (1 - \gamma)(\mu^C - \mu^{DL}) + \gamma(1 - \varepsilon)V_k + \gamma \varepsilon V_{k+1}$. Solving for $V_k$ gives Part 2 of the result.

The expected flow payoff of a defector is $\mu^{UC}(1 + g)$. Combining this with Lemma 2 yields $V_k = (1 - \gamma)\mu^{UC}(1 + g) + \gamma V_{k+1}$. Solving for $V_k$ gives Part 3 of the result. ■

### A.5 Proof of Theorem 4

Fix a coordination proof strict equilibrium. Denote the expected flow payoff to reciprocators by $\pi_R$, the expected flow payoff to unconditional cooperators by $\pi_{UC}$,

\[32\text{This comes from } \beta(\gamma, \varepsilon, \mu^C) = \sum_{n=0}^{\infty} (\gamma(1 - \varepsilon)\mu^C)^n \gamma(1 - (1 - \varepsilon)\mu^C).\]
and the expected flow payoff to defectors by $\pi_D$. By construction, $\pi_R = \mu_R + \mu_{UC}$, $\pi_{UC} = \mu_R + \mu_{UC} - \mu_D l$, and $\pi_D = \mu_{UC}(1 + g)$.

The proof that any equilibrium that satisfies forgery-proofness and coordination-proofness corresponds to GrimKL for some $K, L$ proceeds by establishing the following four lemmas, the proofs of which are in OA.4.

**Lemma 7.** In every non-trivial equilibrium, $\pi_R > \pi_D$. In every equilibrium with unconditional cooperators, $\pi_R > \pi_{UC} > \pi_D$.

$\pi_R > \pi_{UC}$ is an immediate consequence of unconditional cooperators’ playing $C$ against defectors, while reciprocators play $D$. Additionally, every equilibrium requires $\pi_R > \pi_D$, because otherwise reciprocators would strictly prefer to play $D$, as this would strictly increase their flow payoff and weakly increase their continuation value. For the same reason, $\pi_{UC} > \pi_D$ holds in an equilibrium with unconditional cooperators.

Let $z$ denote an arbitrary class from $\{R, UC, D\}$.

**Lemma 8.** Suppose there are infinitely many records at which a player is in class $z$. In an equilibrium that satisfies forgery-proofness, $V_k \geq \pi_z$ for all $k$.

This reason for this is that, were $V_k < \pi_z$ for some $k$, a player with record $k$ could do strictly better by forging a record in class $z$ every period.

**Lemma 9.** In any equilibrium that satisfies forgery-proofness and coordination-proofness, there exists a record $\bar{k}$ such that all records $k \geq \bar{k}$ are defectors and all records $k < \bar{k}$ are either reciprocators or unconditional cooperators.

This result holds because a defector record can never proceed a cooperator record. Otherwise, the defector would forge their record to skip to the next cooperator record. Moreover, all reciprocator records must precede all unconditional cooperator records.

**Lemma 10.** In any equilibrium that satisfies forgery-proofness and coordination-proofness, there exists a record $k^*$ such that all records $k < k^*$ are reciprocators and all records $k \geq k^*$ are either unconditional cooperators or defectors.
Intuitively, since survival is uncertain it is better to obtain the higher flow payoff from reciprocation earlier than the lower flow payoff from unconditional cooperation, so if a \( UC \) record \( k \) was followed by an \( R \) record \( k+1 \), a player could gain by playing \( R \) at record \( k \) and \( UC \) at \( k+1 \).

These lemmas show that only \( GrimKL \) can satisfy both forgery-proofness and coordination-proofness. For the converse, note that \( Grimk KL \) satisfies Coordination Proofness. Moreover, if \( GrimKL \) is an equilibrium, it must satisfy \( V_{k-1} \geq V_k \) for all \( k \leq K + L \), as otherwise some reciprocator or unconditional cooperator would prefer to defect. Since in addition \( V_k = \pi_D \) for all \( k \geq K + L \), it follows that \( V_k \) is everywhere non-increasing, so forgery-proofness is satisfied.

### A.6 Proof of Theorem 5

Parts 1 and 2 of Theorem 5 come from the following proposition.

**Proposition 4.** There is a \( GrimK \) equilibrium with share of cooperators \( \mu^C \) if and only if the following conditions hold:

1. Feasibility:
   \[
   \mu^C = 1 - \beta(\gamma, \varepsilon, \mu^C)^K.
   \]

2. Incentives:
   \[
   (C|C)_0: \mu^C \in \left( \frac{1 + g - \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2}, \frac{1 + g + \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2} \right),
   \]
   \[
   (D|D)_{K-1}: \mu^C < \frac{1}{\gamma(1 - \varepsilon) \left( 1 + \frac{l}{l+1} \right)}.
   \]

The proof of Proposition 4 is in OA.5. The feasibility constraint comes from solving for the steady state record shares \( \mu_k \) and using the fact that \( \mu^C = \sum_{k=0}^{K-1} \mu_k \). The \( (C|C)_0 \) and \( (D|D)_{K-1} \) incentive constraints come from computing the value functions.
for the GrimK strategy using Lemma 6 and then applying Lemma 3. It is without loss to restrict attention to the (C|C) constraint for record 0 and the (D|D) constraint for record K − 1, as these two incentive constraints imply all of the others, because $V_k - V_{k+1}$ is increasing in $k$ for $0 \leq k \leq K - 1$. Intuitively, the incentive for reciprocators to play D when they should play C is greatest when $k = 0$ as then they are farthest away from the punishment phase. Likewise, the incentive for reciprocators to play C when they should play D is greatest when they are closest to the punishment phase, which is when their record is $k = K - 1$.

To see how Part 2 of Theorem 5 comes from Proposition 4, note that

$$\lim_{\varepsilon \to 0} \frac{1 + g - \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2} = g$$

and

$$\lim_{\varepsilon \to 0} \lim_{\gamma \to 1} \frac{l}{\gamma(1 - \varepsilon)} \frac{l}{1 + l} = \frac{l}{1 + l}.$$ 

Thus, it is impossible to satisfy both the (C|C)$_0$ and (D|D)$_{K-1}$ constraints when $g > l/(1 + l)$, which is precisely when $l < g/(1 - g)$. OA.6 gives a slightly modified argument that handles the case where $l = g/(1 - g)$. To see how Part of 1 Theorem 5 comes from Proposition 4, note that, in addition to the above two limit results,

$$\lim_{\varepsilon \to 0} \frac{1 + g + \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2} = 1.$$ 

It follows that values of $\mu^C$ smaller than, but arbitrarily close to, $l/(1 + l)$ satisfy the (C|C)$_0$ and (D|D)$_{K-1}$ constraints in the iterated limit. Thus, the only issue is the feasibility of $\mu^C$ as a steady-state level of cooperation. Because $K$ must be an integer, some values of $\mu^C$ cannot be generated by any $K$, for given values of $\gamma$ and $\varepsilon$. The following result shows that this “integer problem” become irrelevant in the iterated limit. Intuitively, any value of $\mu^C \in [0, 1]$ can be approximated arbitrarily closely by a feasible share of cooperators in a GrimK steady state as $\gamma \to 1$. 

48
Lemma 11. Fix \( \varepsilon \in (0, 1) \). For all \( \Delta > 0 \), there exists \( \bar{\gamma} < 1 \) such that, for all \( \gamma > \bar{\gamma} \) and \( \mu^C \in [0, 1] \), there exists a \( \hat{\mu}^C \) satisfying \( |\hat{\mu}^C - \mu^C| < \Delta \) that satisfies the Feasibility constraint of Proposition 4 for some \( K \).

The proof of Lemma 11 is in OA.7.

A.7 Proof of Theorem 6

A.7.1 Proof Outline

Theorem 6 comes from the following characterization of GrimKL equilibria.

Proposition 5. There is a GrimKL equilibrium with total share of cooperators \( \mu^C \), share of reciprocators \( \mu^R \), and share of unconditional cooperators \( \mu^{UC} \) if and only if the following conditions hold:

1. Feasibility:

\[
\begin{align*}
\mu^C &= 1 - \alpha(\gamma, \varepsilon) L \beta(\gamma, \varepsilon, \mu^C)^K, \\
\mu^R &= 1 - \beta(\gamma, \varepsilon, \mu^C)^K, \\
\mu^{UC} &= (1 - \alpha(\gamma, \varepsilon) L) \beta(\gamma, \varepsilon, \mu^C)^K.
\end{align*}
\]

2. Incentives:

\[
\begin{align*}
(C|C)_0 : \frac{(1 - \varepsilon)(1 - \mu^C)}{1 - (1 - \varepsilon) \mu^C} \left[ \mu^R + \mu^{UC}(l - g) \right] > g, \\
(D|D)_{K - 1} : \frac{\gamma(1 - \varepsilon)(1 - \mu^C)}{1 - \gamma(1 - \varepsilon) \mu^C} \left[ \mu^R + \mu^{UC}(l - g) \right] + \mu^R l < l, \\
(C|D)_{K} : \frac{(1 - \varepsilon)(1 - \mu^C)}{1 - (1 - \varepsilon) \mu^C} \left[ \mu^R + \mu^{UC}(l - g) \right] + \mu^R l > l.
\end{align*}
\]

The proof, which is similar to that of Proposition 4, is in OA.8. In particular, the feasibility constraints come from calculating the steady state shares \( \mu_k \) for the
GrimKL strategy and then setting \( \mu^C = \sum_{k=0}^{K+L-1} \mu_k \), \( \mu^R = \sum_{k=0}^{K-1} \mu_k \), and \( \mu^{UC} = \sum_{k=K}^{K+L-1} \mu_k \). Moreover, the \((C|C)_0\), \((D|D)_{K-1}\), and \((C|D)_K\) incentive constraints come from computing the value functions for the GrimKL strategy using Lemma 6 and then applying Lemma 3.

Part 3 of Theorem 6 follows from Proposition 5 because \((C|C)_0\) and \((C|D)_K\) are impossible to satisfy simultaneously when \( g \geq 1 \) and \( l \leq g(g + 1) \), regardless of \( \gamma \) and \( \varepsilon \). Part 2 also follows from Proposition 5, because, as \( \gamma \) approaches 1, the left-hand side of the \((D|D)_{K-1}\) constraint becomes identical to the left-hand side of the \((C|D)_K\) constraint, so both constraints must hold with equality in this limit. Combining this with \((C|C)_0\) shows that \( \mu^R \leq 1 - g/l \) must be satisfied for all GrimKL equilibria with \( L > 0 \) in the \( \gamma \to 1 \) limit. However, neither the \((C|C)_0\) nor the \((D|D)_{K-1}\) constraints can be satisfied with \( \mu^R \leq 1 - g/l \) when \( l \leq g(g + 1) \). The formal proofs of Parts 2 and 3 of Theorem 6 are given in A.7.2.

Proving that Part 1 of Theorem 6 follows from Proposition 5 is more involved than proving that Part 1 of Theorem 5 follows from Proposition 4. The demonstration proceeds by first identifying a target level of cooperation for fixed \( \varepsilon \). The greatest difficulty in the proof involves showing that there are feasible profiles satisfying the incentive constraints where the level of cooperation actually attains this target as \( \gamma \) approaches 1. The proof then shows that that this level of cooperation approaches 1 as \( \varepsilon \) approaches 0.

A.7.2 Proof of Parts 2 and 3 of Theorem 6

Proof of Part 3 of Theorem 6. Fix \( \varepsilon \in (0,1) \), \( g \geq 1 \), and \( l \leq g(g + 1) \). Since

\[
\frac{(1 - \varepsilon)(1 - \mu^C)}{1 - (1 - \varepsilon)\mu^C} < 1
\]

for all \( \mu^C \in [0, 1] \), the \((C|C)_0\) constraint requires

\[
\mu^R + \mu^{UC}(l - g) > g,
\] (7)
and the \((C|D)_K\) constraint requires
\[
\mu^R(1 + l) + \mu^UC(l - g) > l. \tag{8}
\]

Since \(\mu^R \geq 0\), \(\mu^UC \geq 0\), and \(\mu^R + \mu^UC \leq 1\), Inequality 7 requires \(l > 2g\). Note that \(l > 2g\) and \(l \leq g(g + 1)\) cannot be jointly satisfied when \(g = 1\), and hence the \((C|C)_0\) constraint cannot be satisfied when \(g = 1\). As equilibria with unconditional cooperators but not reciprocators are impossible, it follows that, when \(g = 1\), there is no equilibrium with cooperators, so \(\overline{\mu}^C_{KL}(\gamma, \varepsilon) = 0\) when \(g = 1\). Now, consider \(g > 1\). Inequality 7 additionally requires \(\mu^R < \frac{l-2g}{l-g-1}\).

Because \((l-2g)/(l-g-1)\) is strictly increasing in \(l\) for \(l > 2g\), it follows that
\[
\mu^R < \frac{g(g + 1) - 2g}{g(g + 1) - g - 1} = \frac{g}{g + 1}, \tag{9}
\]

since \(l \leq g(g + 1)\) by assumption. Similarly, Inequality 8 requires
\[
\mu^R > \frac{g}{g + 1}. \tag{10}
\]

However, Inequalities 9 and 10 are mutually incompatible, so the \((C|C)_0\) and \((C|D)_K\) constraints cannot be jointly satisfied. Therefore, there is no \textit{GrimKL} equilibrium with both reciprocators and unconditional cooperators. As Inequality 7 cannot be satisfied with \(\mu^UC = 0\) and equilibria with unconditional cooperators but not reciprocators are impossible, it follows that there is no equilibrium with cooperators. Consequently, \(\overline{\mu}^C_{KL}(\gamma, \varepsilon) = 0\) when \(g > 1\).

\textbf{Proof of Part 2 of Theorem 6.} We show that \(\limsup_{\gamma \to 1} \overline{\mu}^C_{KL}(\gamma, \varepsilon) = 0\) for all \(\varepsilon \in (0, 1)\) when \(l \leq g(g + 1)\). Suppose otherwise that \(\limsup_{\gamma \to 1} \overline{\mu}^C_{KL}(\gamma, \varepsilon) > 0\). Then there is some \(\gamma_n \to 1\) and a sequence of associated equilibria with shares \((\mu^R_n, \mu^UC_n)\), such that \((\mu^R_n, \mu^UC_n)\) converges to \((\mu^R, \mu^UC)\) with \(\mu^C = \mu^R + \mu^UC > 0\). Theorem 5 implies that \(\limsup_{\gamma \to 1} \overline{\mu}^C_{K}(\gamma, \varepsilon) = 0\) for \(l \leq g(g + 1)\), so such a sequence must satisfy the \((C|C)_0\),
Taking the limit of these constraints as $n \to \infty$ shows that $(\mu^R, \mu^{UC})$ must satisfy the following “limit” constraints.

\begin{align*}
\text{Limit } (C|C)_0 : & \quad \frac{(1 - \varepsilon)(1 - \mu^C)}{1 - (1 - \varepsilon)\mu^C} (\mu^R + \mu^{UC} (l - g)) \geq g, \\
\text{Limit } (D|D)_{K-1} : & \quad \frac{(1 - \varepsilon)(1 - \mu^C)}{1 - (1 - \varepsilon)\mu^C} (\mu^R + \mu^{UC} (l - g)) + \mu^R l \leq l, \\
\text{Limit } (C|D)_K : & \quad \frac{(1 - \varepsilon)(1 - \mu^C)}{1 - (1 - \varepsilon)\mu^C} (\mu^R + \mu^{UC} (l - g)) + \mu^R l \geq l.
\end{align*}

The limit $(D|D)_{K-1}$ and $(C|D)_K$ constraints together imply that

\[ \frac{(1 - \varepsilon)(1 - \mu^C)}{1 - (1 - \varepsilon)\mu^C} (\mu^R + \mu^{UC} (l - g)) + \mu^R l = l. \]

Along with the limit $(C|C)_0$ constraint, this implies that $\mu^R \leq 1 - g/l$.

For any $\varepsilon \in (0, 1)$,

\[ \frac{(1 - \varepsilon)(1 - \mu^C)}{1 - (1 - \varepsilon)\mu^C} < 1 \]

for all $\mu^C \in [0, 1]$. Therefore, the Limit $(C|D)_K$ constraint requires

\begin{equation}
\mu^R (1 + l) + \mu^{UC} (l - g) > l.
\end{equation}

Since $0 \leq \mu^R \leq 1 - g/l$, $\mu^{UC} \geq 0$, and $\mu^R + \mu^{UC} \leq 1$, it follows that

\[ \mu^R (1 + l) + \mu^{UC} (l - g) \leq (1 - \frac{g}{l}) (1 + l) + \frac{g}{l} (l - g) = l + \frac{1}{l} (l - g (g + 1)). \]

Since $l \leq g (g + 1)$, this implies that $\mu^R (1 + l) + \mu^{C} (l - g) \leq l$, a contradiction. ■
A.7.3 Proof of Part 1 of Theorem 6

Fix $\mu_R \in (g/(1 + g), 1 - g/l]$. Consider the equation

$$
\frac{(1 - \varepsilon)(1 - \mu_C)}{1 - (1 - \varepsilon)\mu_C} \left[ (l - g)\mu_C + (1 + g - l)\mu_R \right] + l\mu_R = l
$$

(13)

and the function $h(\varepsilon, \mu_R)$ defined by

$$
h(\varepsilon, \mu_R) = \max\{\mu_C \in [0, 1] : \mu_C \text{ solves Equation 13}\}.
$$

If $h(\varepsilon, \mu_R)$ is well-defined, it gives the maximum level of cooperation for the given $\mu_R$ and $\varepsilon$ that satisfies the $\gamma \to 1$ “limit” constraints of Equation 11. Straightforward calculations show that, for any $\mu_R \in (g/(1 + g), 1 - g/l]$, $h(\varepsilon, \mu_R)$ is well-defined for sufficiently small and positive $\varepsilon$, and that

$$
\lim_{\varepsilon \to 0} \frac{1 - h(\varepsilon, \mu_R)}{\varepsilon} = \frac{l(1 - \mu_R)}{(1 + g)\mu_R - g}.
$$

An immediate implication of this is $\lim_{\varepsilon \to 0} h(\varepsilon, \mu_R) = 1$. Combining this with the following two lemmas proves Part 1 of Theorem 6.

Let $\kappa : (g/(1 + g), 1 - g/l] \to \mathbb{R}$ be the function given by

$$
\kappa(\mu_R) = \frac{l \ln(1 + \mu_R)(1 - \mu_R)}{l - g + (1 + g - l)\mu_R},
$$

and $\iota : (g/(1 + g), 1 - g/l] \to \mathbb{R}_+$ be the function given by

$$
\iota(\mu_R) = \frac{(1 + g)\mu_R - g + 1}{l - g + (1 + g - l)\mu_R}.
$$

**Lemma 12.** Fix $\mu_R \in (g/(1 + g), 1 - g/l]$. If $|1 + \kappa(\mu_R)| > \iota(\mu_R)$, then there exists some $\varepsilon > 0$, such that $\liminf_{\gamma \to 1} \overline{\mu}_{KL}^C(\gamma, \varepsilon) \geq h(\varepsilon, \mu_R)$ for $\varepsilon < \varepsilon$.

**Lemma 13.** Suppose that $l > g(g + 1)$. Some $\mu_R \in (g/(1 + g), 1 - g/l]$ satisfies $|1 + \kappa(\mu_R)| > \iota(\mu_R)$ if $l > \max\{g(g + 1), b(g)\}$.

OA.9.1 presents the proof of Lemma 12. It makes heavy use of the inverse function theorem and other tools of differential calculus to show that, when $|1 + \kappa(\mu_R)| >$
ι(μ^R), for sufficiently small ε, any neighborhood of (h(ε, μ^R), μ^R) can be approached by feasible profiles for sufficiently high γ. The proof of Lemma 13 is in OA.9.2.