Recall that $R(t)$ is the set of feasible records for a player of age $t$, and

$$\bar{M} = \left\{ \mu \in M : \sum_{k \in R(t)} \mu_k \leq \gamma^t \ \forall t \in \mathbb{N} \right\}.$$

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

**Proof of Claim 1.** Since $\bar{M}$ is a metric space under the sup norm topology, it suffices to show that it is sequentially compact. Consider a sequence $\{\mu^n\}_{n \in \mathbb{N}}$ of $\mu^n \in \bar{M}$. By a standard diagonalization argument, there exists some $\tilde{\mu} \in [0, 1]^\mathbb{Z}$ and some subsequence $\{\mu^{n_m}\}_{m \in \mathbb{N}}$ such that $\lim_{m \to \infty} \mu^{n_m} = \tilde{\mu}$ for all $k \in \mathbb{Z}$. This, along with the fact that

$$\sum_{t' \leq t} \sum_{k \in R(t')} \mu^{n_m}_k \geq 1 - \frac{\gamma^t}{1 - \gamma}$$

for all $t \in \mathbb{N}$ and $m \in \mathbb{N}$ implies that

$$\sum_{t' \leq t} \sum_{k \in R(t')} \tilde{\mu}_k \geq 1 - \frac{\gamma^t}{1 - \gamma}$$

for all $t \in \mathbb{N}$. Thus, $\sum_{k \in \mathbb{Z}} \tilde{\mu}_k = 1$ since $\gamma^t \to 0$ as $t \to \infty$, so $\tilde{\mu} \in M$. Additionally, since

$$\sum_{k \in R(t)} \mu^{n_m}_k \leq \gamma^t$$
for all \( t \in \mathbb{N} \) and \( m \in \mathbb{N} \), it follows that

\[
\sum_{k \in R(t)} \tilde{\mu}_k \leq \gamma^t
\]

for all \( t \in \mathbb{N} \), so \( \tilde{\mu} \in \bar{M} \).

Now, we show that \( \lim_{m \to \infty} \mu_{nm} = \tilde{\mu} \). Fix \( \varepsilon > 0 \). Let \( T \in \mathbb{N} \) be such that \( \gamma^T < \varepsilon \), and let \( M \in \mathbb{N} \) be such that \( |\mu_{nm}^k - \tilde{\mu}_k| < \varepsilon \) for all \( k \in R(t), t < T \) and \( m > M \). Thus,

\[
\sup_{k \in \mathbb{Z}} |\mu_{nm}^k - \tilde{\mu}_k| < \varepsilon
\]

for all \( m > M \).

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

**Proof of Claim 2.** By the definition of \( R(t) \), for all \( \mu \in M \) and all \( t \in \mathbb{N} \), we have

\[
\gamma \sum_{k \in R(t)} \mu_k = \sum_{k \in R(t+1)} f_{r,s}(\mu)[k].
\]

Hence, if \( \sum_{k \in R(t)} \mu_k \leq \gamma^t \) for all \( t \in \mathbb{N} \) then \( \sum_{k \in R(t+1)} f_{r,s}(\mu)[k] \leq \gamma^{t+1} \) for all \( t \in \mathbb{N} \) (and, trivially, \( \sum_{k \in R(0)} f_{r,s}(\mu)[k] \leq 1 \)). That is, \( f_{r,s} \) maps \( \bar{M} \) to itself.

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

**Proof of Claim 3.** Given a vector \( \mu \in \bar{M} \), define the vector \( \mu_{\mid T} \in \bar{M}_T \) by setting \( \mu_{\mid T} = \mu_\tau \) for \( t \in \bigcup_{t \leq T} R(t) \) and \( \mu_{\mid T} = 0 \) for \( t \notin \bigcup_{t \leq T} R(t) \). For each \( T \), define the function \( f_{r,s\mid T} : \bar{M} \to \bar{M}_T \) by setting \( f_{r,s\mid T}(\mu) = f_{r,s}(\mu)\mid T \) for each \( \mu \in \bar{M} \). Note that \( f_{r,s\mid T}(\mu) = f_{r,s\mid T}(\mu') \) whenever \( \mu_{\mid \bigcup_{t \leq T} R(t)} = \mu'_{\mid \bigcup_{t \leq T} R(t)} \). Thus, \( f_{r,s\mid T} \) can equivalently be viewed as a function from \( \bar{M}_T \) to \( \bar{M}_T \), and as a polynomial function it is continuous on \( \bar{M}_T \), and hence on \( \bar{M} \). Moreover, since the mass on records outside \( \bigcup_{t \leq T} R(t) \) goes to 0 as \( T \to \infty \), for any \( \varepsilon > 0 \) there exists \( T \) such that \( |f_{r,s}(\mu) - f_{r,s}(\mu)\mid T| < 2\varepsilon \) for all \( \mu \in \bar{M} \). Hence, \( f_{r,s} \) is also continuous on \( \bar{M} \).
OA.2 Proof of Proposition 1

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

Assume $g < 1$. Fix any rational number $\rho$ satisfying $g < \rho < \min\{l, 1\}$. Let $m$ and $n$ be integers such that $m \geq n > 0$ and $n/m = \rho$.

We consider a strategy with $m + n$ phases, where a player is in phase $j$ whenever her record equals $j - 1 \mod m + n$. The first $m$ phases, denoted $G_1$ through $G_m$, are good phases, and the remaining $n$ phases, denoted $B_{m+1}$ through $B_{m+n}$, are bad phases. A player is a reciprocator while in a good phase and a defector while in a bad phase.

We denote the share of players in phase $G_j$ by $\mu_{G_j}$ and the share of players in phase $B_j$ by $\mu_{B_j}$. Consequently, the total share of cooperators is $\mu^C = \sum_{j=1}^{m} \mu_{G_j}$ and the total share of defectors is $\mu^D = \sum_{j=m+1}^{m+n} \mu_{B_j} = 1 - \mu^C$.

We first prove that under this strategy the share of cooperators $\mu^C$ converges to 1 in the iterated limit where $\gamma$ approaches 1 and then $\varepsilon$ approaches 0.\footnote{Here and subsequently we do not track the dependence of endogenous objects like $\mu^C$ on $\gamma$ and $\varepsilon$ in the notation.} We then prove that the strategy does in fact give strict equilibria. This result holds for all $\rho < 1$, so for any $g$ and $l$ such that $g < \min\{l, 1\}$, there is a strategy that obtains limit efficiency. This proves Proposition 1.

Lemma 1. $\lim_{\varepsilon \to 0} \lim_{\gamma \to 1} \mu^C = 1$.

Before proving the lemma, we give a heuristic argument. As $\gamma \to 1$, the mass $\mu^C$ of reciprocators will be approximately equally distributed among the first $m$ phases, while the mass $1 - \mu^C$ of defectors will be approximately equally distributed among the next $n$ phases. Following $\gamma \to 1$, as $\varepsilon \to 0$, the flow from phase $m+n$ to phase 1 (the “inflow into cooperation”) is approximately $(1 - \mu^C)/n$, while the flow from phase $m$ to phase $m+1$ (the “inflow into defection”) is approximately $(1 - \mu^C)(\mu^C/m)$. If these flows were equal for some $\mu^C < 1$, this would imply $\mu^C = m/n < 1$. But this contradicts the
fact that \( m \geq n \) as there are more good phases than bad phases. Therefore, a steady state requires that \( \mu^C = 1 \) in the iterated limit.

**Proof of Lemma 1.** Since the first \( m \) phases all correspond to reciprocator records, Lemmas 4 and 5 imply that \( i_{G_j} = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_{G_{j-1}} \) and \( \tau_{G_j} = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_{G_j} \) for all \( 1 < j \leq m \). By Equation 7, \( \mu_{G_j} = \beta(\gamma, \varepsilon, \mu^C)\mu_{G_{j-1}} \), so induction gives

\[
\mu_{G_j} = \beta(\gamma, \varepsilon, \mu^C)^{j-1}\mu_{G_1} \tag{1}
\]

for \( 1 \leq i \leq m \).

Since the phase \( m \) corresponds to reciprocator records and phase \( m+1 \) corresponds to defector records, Lemmas 4 and 5 give \( i_{B_{m+1}} = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_{G_{m}} \) and \( \tau_{B_{m+1}} = \gamma\mu_{B_{m+1}} \), so Equation 7 implies

\[
\mu_{B_{m+1}} = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_{G_{m}}
= \beta(\gamma, \varepsilon, \mu^C)^m(1 - \gamma(1 - \varepsilon)\mu^C)\mu_{G_1}. \tag{2}
\]

Since the last \( n \) phases all correspond to defector records, Lemmas 4 and 5 give \( i_{B_j} = \gamma\mu_{B_{j-1}} \) and \( \tau_{B_j} = \gamma\mu_{B_j} \) for \( m < j \leq m + n \). Thus, Equation 7 implies that \( \mu_{B_j} = \gamma\mu_{B_{j-1}} \), so induction, combined with Equation 2, gives

\[
\mu_{B_{m+n}} = \gamma^{n-1}\mu_{B_{m+1}}
= \gamma^{n-1}\beta(\gamma, \varepsilon, \mu^C)^m(1 - \gamma(1 - \varepsilon)\mu^C)\mu_{G_1}. \tag{3}
\]

Finally, since phase 1 corresponds to reciprocator records and phase \( m + n \) corresponds to defector records, Lemmas 4 and 5 give \( i_{G_1} = 1 - \gamma + \gamma\mu_{B_{m+n}} \) and \( \tau_{G_1} = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_{G_1} \), so Equations 7 and 3 imply

\[
\mu_{G_1}^C = \frac{1 - \gamma}{1 - \gamma(1 - \varepsilon)\mu^C} + \frac{\gamma}{1 - \gamma(1 - \varepsilon)\mu^C}\mu_{B_{m+n}}
= 1 - \beta(\gamma, \varepsilon, \mu^C) + \gamma^n\beta(\gamma, \varepsilon, \mu^C)^m\mu_{G_1}.
\]
Solving this for $\mu_{G_1}$ gives

$$\mu_{G_1} = \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m}. \quad (4)$$

Equations 1 and 4 together imply that

$$\sum_{j=1}^{m} \mu_{G_j} = \sum_{j=1}^{m} \beta(\gamma, \varepsilon, \mu^C)^{j-1} \left( \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m} \right) = \frac{1 - \beta(\gamma, \varepsilon, \mu^C)^m}{1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m}.$$  

Therefore,

$$\mu^C = \frac{1 - \beta(\gamma, \varepsilon, \mu^C)^m}{1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m}. \quad (5)$$

Consider the function $f : [0, 1] \times (0, 1) \times [0, 1] \to \mathbb{R}$ given by

$$f(\gamma, \varepsilon, \mu^C) = \begin{cases} 
\frac{1 - \beta(\gamma, \varepsilon, \mu^C)^m}{1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m} & \text{if } \gamma < 1 \\
1 + \rho(1 - (1 - \varepsilon)\mu^C) & \text{if } \gamma = 1
\end{cases}.$$  

This function extends $(1 - \beta(\gamma, \varepsilon, \mu^C)^m)/(1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m)$ to $\gamma = 1$ and, for any fixed value of $\varepsilon \in (0, 1)$, is a continuous function of $(\gamma, \mu^C) \in [0, 1] \times [0, 1]$, which can be shown using L'Hôpital’s rule.

Therefore, for fixed $\varepsilon$, any limit point of any sequence of steady state $\mu^C$ as $\gamma \to 1$ must satisfy

$$\mu^C = f(1, \varepsilon, \mu^C) = \frac{1}{1 + \rho(1 - (1 - \varepsilon)\mu^C)}.$$  

The only such $\mu^C \in [0, 1]$ that satisfies this equation for $\varepsilon$ is

$$\mu^C(\varepsilon) = \frac{1 + \rho - \sqrt{(1 + \rho)^2 - 4(1 - \varepsilon)\rho}}{2(1 - \varepsilon)\rho},$$

so $\lim_{\gamma \to 1} \mu^C = \mu^C(\varepsilon)$ for all $\varepsilon$. Lemma 1 follows since $\lim_{\varepsilon \to 0} \mu^C(\varepsilon) = 1$.  

\[\blacksquare\]
The next lemma formalizes the idea that, on average, adding a $D$ to one’s record while in a good phase leads to an extra $\rho$ periods of punishment. Let $V_{G_j}$ denote the value function of a player in phase $G_j$ and $V_{B_j}$ denote the value function of a player in phase $B_j$.

**Lemma 2.** The following iterated limits hold:

$$\lim_\varepsilon \lim_\gamma \frac{V_{G_j} - V_{G_{j+1}}}{1 - \gamma} = \rho \text{ for } 1 \leq j < m \text{ and } \lim_\varepsilon \lim_\gamma \frac{V_{G_m} - V_{B_{m+1}}}{1 - \gamma} = \rho.$$

By Lemma 3, this implies that the $(C|C)$ and $(D|D)$ constraints for a player in a good phase are satisfied in the iterated limit when $g < \rho < l$. The incentives for players in bad phases are trivial, because the value function in the next phase is larger than the value function in the current phase whenever the current phase is bad. This implies that playing $D$ while in a bad phase maximizes both the flow payoff and the continuation payoff. Therefore, Lemma 3, along with Lemma 1, suffices to prove Proposition 1.

**Proof of Lemma 2.** By Lemma 6,

$$\frac{V_{G_j} - V_{G_{j+1}}}{1 - \gamma} = \frac{\mu^C - V_{G_{j+1}}}{1 - \gamma(1 - \varepsilon)\mu^C}$$

for all $1 \leq j < m$ and

$$\frac{V_{G_m} - V_{B_{m+1}}}{1 - \gamma} = \frac{\mu^C - V_{B_{m+1}}}{1 - \gamma(1 - \varepsilon)\mu^C}.$$

Lemma 6 also implies that the value functions in all phases converge to the same value as $\gamma \to 1$ for fixed $\varepsilon$. Therefore, it suffices to show that

$$\lim_\varepsilon \lim_\gamma \frac{\mu^C - V_{G_1}}{1 - \gamma(1 - \varepsilon)\mu^C} = \rho.$$

Since the first $m$ phases all correspond to reciprocator records, combining Lemma
6 with induction gives

\[ V_{G_1} = (1 - \beta(\gamma, \varepsilon, \mu^C)^m)\mu^C + \beta(\gamma, \varepsilon, \mu^C)^m V_{B_{m+1}}. \]

Likewise, since the last \( n \) phases all correspond to defector records, Lemma 6, along with the fact that \( \mu^{UC} = 0 \) in the strategies considered, implies

\[ V_{B_{m+1}} = \gamma^n V_{G_1}. \]

Combining these equations gives

\[ V_{G_1} = (1 - \beta(\gamma, \varepsilon, \mu^C)^m)\mu^C + \gamma^n \beta(\gamma, \varepsilon, \mu^C) V_{G_1}. \]

and solving this for \( V_{G_1} \) renders

\[ V_{G_1} = \left( \frac{1 - \beta(\gamma, \varepsilon, \mu^C)^m}{1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m} \right) \mu^C. \]

Therefore,

\[
\frac{\mu^C - V_{G_1}}{1 - \gamma(1 - \varepsilon)\mu^C} = \left( \frac{(1 - \gamma^n)\beta(\gamma, \varepsilon, \mu^C)^m}{(1 - \gamma(1 - \varepsilon)\mu^C)(1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m)} \right) \mu^C.
\]

Since \( \lim_{\varepsilon \to 0} \lim_{\gamma \to 1} \mu^C = 1 \) by Lemma 1, we need only show that

\[
\lim_{\varepsilon \to 0} \lim_{\gamma \to 1} \frac{(1 - \gamma^n)\beta(\gamma, \varepsilon, \mu^C)^m}{(1 - \gamma(1 - \varepsilon)\mu^C)(1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m)} = \rho.
\]

(6)

Consider the function \( \tilde{f} : [0, 1] \times (0, 1) \times [0, 1] \to \mathbb{R} \) given by

\[
\tilde{f}(\gamma, \varepsilon, \mu^C) = \begin{cases} 
\frac{(1 - \gamma^n)\beta(\gamma, \varepsilon, \mu^C)^m}{(1 - \gamma(1 - \varepsilon)\mu^C)(1 - \gamma^n \beta(\gamma, \varepsilon, \mu^C)^m)} & \text{if } \gamma < 1 \\
\frac{\rho}{1 + \rho(1 - (1 - \varepsilon)\mu^C)} & \text{if } \gamma = 1
\end{cases}
\]

For any fixed value of \( \varepsilon \), this function is a continuous function of \( (\gamma, \mu^C) \in [0, 1] \times [0, 1] \),
which can be shown using L'Hôpital's rule. Therefore,

$$\lim_{\epsilon \to 0} \lim_{\gamma \to 1} \frac{(1 - \gamma^n)\beta(\gamma, \epsilon, \mu^C)^m}{(1 - \gamma(1 - \epsilon)\mu^C)(1 - \gamma^n\beta(\gamma, \epsilon, \mu^C)^m)} = \lim_{\epsilon \to 0} \lim_{\gamma \to 1} \tilde{f}(\gamma, \epsilon, \mu^C)$$

$$= \lim_{\epsilon \to 0} \frac{\rho}{1 + \rho(1 - (1 - \epsilon)\overline{\mu}^C(\epsilon))} = \rho,$$

where the second and third equalities follow from the facts established in the proof of Lemma 1 that \(\lim_{\gamma \to 1} \mu^C = \overline{\mu}^C(\epsilon)\) and \(\lim_{\epsilon \to 0} \overline{\mu}^C(\epsilon) = 1\). Thus, the limit in Equation 6 is satisfied, and Lemma 2 follows.

\[\square\]

OA.3 Extension of Proposition 3

**Proposition 1.** When records count C’s, the unique strict equilibrium remains Always Defect even if plays of D are mis-recorded as C with probability \(\epsilon' > 0\).

**Proof.** It suffices to show that \(V_k \leq V_{k+1}\) for all \(k \in \mathbb{N}\) and apply the proof of Proposition 3. So suppose \(V_k > V_{k+1}\) for some \(k\). Then a player with record \(k\) always plays D: this maximizes both her flow payoff and her continuation payoff. Since there are no anti-reciprocators, this implies that only unconditional cooperators play C against a player with record \(k\).

Now, let \(\pi_k\) denote the flow payoff of a player with record \(k\), i.e. a player who always plays D while facing C from unconditional cooperators and D from everyone else. We claim that \(\pi_k \leq \pi_{k'}\) for every record \(k'\), and in particular \(\pi_k \leq V_{k+1}\). To see this, note that a player who always plays D gets at least \(\pi_k\) in every period (as unconditional cooperators always play C against her, and depending on her record maybe other players do, too). So if \(\pi_k > V_{k'}\) then a player with record \(k'\) would do strictly better to play D forever.

We now have \(V_k > V_{k+1}\) and \(\pi_k \leq V_{k+1}\). But \(V_k\) is a weighted average of \(\pi_k\) and \(V_{k+1}\), a contradiction.  

\[\square\]
OA.4 Erasure-Proofness

Proposition 2. For any record-keeping system, the only erasure-proof equilibrium is Always Defect.

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 erasure-proof), then no player with an unconditional defector record can earn flow payoffs within $\frac{1-\gamma}{\gamma} g$ 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 erasure-proof equilibrium must prescribe unconditional defection at record $k$, as playing $D$ and erasing the record update 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 erasure-proof equilibrium. ■

OA.5 Proofs of Results for Theorem 3

Lemma 3. In every non-trivial equilibrium, $\pi_R > \pi_D$. In every equilibrium with unconditional cooperators, $\pi_R > \pi_{UC} > \pi_D$.

Proof of Lemma 3. $\pi_R > \pi_{UC}$ follows from $\mu^D > 0$. To see that $\mu^D > 0$, note that, if $\mu^D = 0$, then every player would face only cooperators for the duration of her lifetime regardless of her history of play. However, then every player would defect in every period, a contradiction.
Next, if $\pi_D \geq \pi_{UC}$, then unconditional cooperators would receive the lowest flow payoff of any class. Since $V_k$ is a convex combination of $\pi_{k'}$ for $k' \geq k$ and a player’s record remains constant when she plays $C$, this implies that a player at any unconditional cooperator record would do strictly better by playing $D$ until her record changes, a contradiction.

Hence, $\pi_R > \pi_{UC} > \pi_D$ in any equilibrium with unconditional cooperators. A similar argument implies $\pi_R > \pi_D$ in any non-trivial equilibrium without unconditional cooperators. □

**Lemma 7.** In any equilibrium that satisfies forgery-proofness, there exists a record $\bar{k}$ such that a record $k$ is an unconditional defector record iff $k \geq \bar{k}$.

**Proof of Lemma 7.** First, we establish that there must be some cutoff record after which a player is always an unconditional defector. Note that for any record $k$ at which a player is not an unconditional defector, the $(C|C)_k$ constraint requires that

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

Thus, if in a forgery-proof equilibrium where the value function is everywhere non-increasing, there were infinitely many records at which a player is not an unconditional defector, there would be some $k$ for which $V_k < 0$, which is impossible in equilibrium.

We now establish that unconditional defector records can only be followed by other unconditional defector records. The reason for this is that otherwise, there would be some unconditional defector record $k$ at which a player would strictly prefer to inflate her record until she reaches the next record at which some cooperation occurs, which would violate forgery proofness. □

**Lemma 8.** 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 unconditional defectors.
Proof of Lemma 8. From Lemma 7, there exists some record $\bar{k}$ such that all records $k \geq \bar{k}$ are unconditional defector records and all records $k < \bar{k}$ are either reciprocator or unconditional cooperator records. Suppose that there are $m$ records at which a player is a reciprocator and $\bar{k} - m$ records at which a player is an unconditional cooperator. We must show that the first $m$ records, $0 \leq k \leq m - 1$, correspond to $R$ while the next $\bar{k} - m$ records, $m \leq k \leq \bar{k} - 1$, correspond to $UC$. Note that this is vacuously true if $m = 0$ or $m = \bar{k}$. We now show that it is true for $0 < m < \bar{k}$. Suppose towards a contradiction that there exists some $k$ satisfying $0 \leq k < \bar{k} - 1$ which corresponds to $UC$ and is such that $k + 1$ corresponds to $R$. By the familiar recursive expressions for a player’s expected payoff as a function of their record, a record $k + 1$ player’s expected payoff given by the strategy profile is $V_{k+1} = (1 - \beta(\gamma, \varepsilon, \mu^C))\pi_R + \beta(\gamma, \varepsilon, \mu^C)V_{k+2}$, and a record $k$ player’s expected payoff given by the strategy profile is

$$V_k = (1 - \alpha(\gamma, \varepsilon))\pi_{UC} + \alpha(\gamma, \varepsilon)V_{k+1}$$

$$= \alpha(\gamma, \varepsilon)(1 - \beta(\gamma, \varepsilon, \mu^C))\pi_R + (1 - \alpha(\gamma, \varepsilon))\pi_{UC} + \alpha(\gamma, \varepsilon)\beta(\gamma, \varepsilon, \mu^C)V_{k+2},$$

where $V_{k+2}$ is the player’s expected payoff upon reaching record $k + 2$. However, suppose instead that the player changed their strategy so that she plays according to $R$ at record $k$ and according to $UC$ at record $k + 1$, but otherwise keeps her strategy the same. Then the player’s expected payoff upon reaching record $k + 1$, which we denote by $\tilde{V}_{k+1}$, would be $\tilde{V}_{k+1} = (1 - \alpha(\gamma, \varepsilon))\pi_{UC} + \alpha(\gamma, \varepsilon)V_{k+2}$, and the player’s expected payoff upon reaching record $k$, which we denote by $\tilde{V}_k$, would be

$$\tilde{V}_k = (1 - \beta(\gamma, \varepsilon, \mu^C))\pi_R + \beta(\gamma, \varepsilon, \mu^C)V_{k+1}$$

$$= (1 - \beta(\gamma, \varepsilon, \mu^C))\pi_R + (1 - \alpha(\gamma, \varepsilon))\beta(\gamma, \varepsilon, \mu^C)\pi_{UC} + \alpha(\gamma, \varepsilon)\beta(\gamma, \varepsilon, \mu^C)V_{k+2}.$$ 

Note that $\tilde{V}_k - V_k = (1 - \alpha(\gamma, \varepsilon))(1 - \beta(\gamma, \varepsilon, \mu^C))(\pi_R - \pi_{UC}) > 0$ where the inequality follows because $\pi_R > \pi_{UC}$ by Lemma 3. Thus the profile is not an equilibrium. ■
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:

   \[ (C|C)_0 : \mu^C \in \left( \frac{1 + g - \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon}g}}{2}, \frac{1 + g + \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon}g}}{2} \right), \]

   \[ (D|D)_{K-1} : \mu^C < \frac{1}{\gamma(1 - \varepsilon)l}. \]

Lemma 4. In a GrimK steady state with total share of cooperators $\mu^C$,

\[ \mu_k = \begin{cases} 
\beta(\gamma, \varepsilon, \mu^C)^k(1 - \beta(\gamma, \varepsilon, \mu^C)) & \text{if } 0 \leq k \leq K - 1 \\
\gamma^{k-K} \beta(\gamma, \varepsilon, \mu^C)^K(1 - \gamma) & \text{if } k \geq K 
\end{cases} \]  

Moreover, $\mu^C$ satisfies the equation

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

Proof. Since $i_0 = 1 - \gamma$ and $\tau_0 = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_0$, Equation 7 implies $\mu_0 = \frac{1 - \gamma}{1 - \gamma(1 - \varepsilon)\mu^C} = 1 - \beta(\gamma, \varepsilon, \mu^C)$. Moreover, by Lemmas 4 and 5, $i_{k+1} = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_k$ and $\tau_{k+1} = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_{k+1}$ for $0 \leq k \leq K - 1$. Thus, Equation 7 implies $\mu_{k+1} = \beta(\gamma, \varepsilon, \mu^C)\mu_k$ for $0 \leq k \leq K - 1$. By induction, $\mu_k = \beta(\gamma, \varepsilon, \mu^C)^k(1 - \beta(\gamma, \varepsilon, \mu^C))$ for $0 \leq k \leq K - 1$.

By Lemmas 4 and 5, $i_K = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_{K-1}$ and $\tau_K = \gamma\mu_K$, so Equation 7 implies $\mu_K = \gamma(1 - (1 - \varepsilon)\mu^C)\mu_{K-1} = \beta(\gamma, \varepsilon, \mu^C)^K(1 - \gamma)$. Likewise, by Lemmas 4 and
5, \(i_{k+1} = \gamma \mu_k\) and \(\tau_{k+1} = \gamma \mu_{k+1}\) for \(k \geq K\). Hence, Equation 7 implies \(\mu_{k+1} = \gamma \mu_k\) for \(k \geq K\). Combining this with the previously derived \(\mu_K = \beta(\gamma, \varepsilon, \mu_C)^K(1 - \gamma)\) and applying induction gives \(\mu_k = \gamma^{k-K} \beta(\gamma, \varepsilon, \mu_C)^K(1 - \gamma)\) for \(k \geq K\). This proves Equation 7.

To prove Equation 8, note that Equation 7 implies that

\[
\mu^C = \sum_{k=0}^{K-1} \mu_k = \sum_{k=0}^{K-1} \beta(\gamma, \varepsilon, \mu_C^C)^k (1 - \beta(\gamma, \varepsilon, \mu_C^C)) = 1 - \beta(\gamma, \varepsilon, \mu_C^C)^K.
\]

\[\blacksquare\]

**Lemma 5.** The value function of a player with record \(k\) is

\[
V_k = \begin{cases} 
(1 - \beta(\gamma, \varepsilon, \mu_C^C)^K)^{k-\mu_C^C} & \text{if } 0 \leq k \leq K - 1 \\
0 & \text{if } k \geq K
\end{cases}
\]

**Proof.** Players with record \(k \geq K\) are defectors and obtain a flow payoff of 0 in all future periods, so \(V_k = 0\) for \(k \geq K\). Combining this with \(V_k = (1 - \beta(\gamma, \varepsilon, \mu_C^C))(1 - \beta(\gamma, \varepsilon, \mu_C^C)^K) + \beta(\gamma, \varepsilon, \mu_C^C)V_{k+1}\) for \(0 \leq k \leq K - 1\) from Lemma 6 and solving inductively for \(V_k\) gives \(V_k = (1 - \beta(\gamma, \varepsilon, \mu_C^C)^K)^{k-\mu_C^C} \mu_C^C\) for \(0 \leq k \leq K - 1\). \[\blacksquare\]

**Lemma 6.**

1. The \((C|C)_0\) constraint is

\[
\frac{1 - \varepsilon}{1 - \gamma(1 - \varepsilon)^{\mu_C}} \beta(\gamma, \varepsilon, \mu_C) > g.
\]

2. The \((D|D)_{K-1}\) constraint is

\[
\mu_C < \frac{1}{\gamma(1 - \varepsilon)} \frac{l}{1 + l}.
\]
Proof. We first derive the \((C|C)_{0}\) constraint. From Lemma 9,

\[
\gamma(1 - \varepsilon) \frac{V_0 - V_1}{1 - \gamma} = \gamma(1 - \varepsilon) \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{1 - \gamma} \beta(\gamma, \varepsilon, \mu^C)^{K-1} \mu^C
\]

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

and the \((C|C)_{0}\) constraint is equivalent to

\[
\frac{1 - \varepsilon}{1 - \gamma(1 - \varepsilon) \mu^C} \beta(\gamma, \varepsilon, \mu^C)^{K} \mu^C > g.
\]

We now derive the \((D|D)_{K-1}\) constraint. From Lemma 9, \(V_{K-1} = (1 - \beta(\gamma, \varepsilon, \mu^C)) \mu^C\) and \(V_K = 0\). Therefore,

\[
\gamma(1 - \varepsilon) \frac{V_{K-1} - V_K}{1 - \gamma} = \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{1 - \gamma} \gamma(1 - \varepsilon) \mu^C
\]

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

Hence, the \((D|D)_{K-1}\) constraint is equivalent to

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

Manipulating this inequality yields (11.)

\[
\text{Corollary 1. When combined with the steady state condition Equation 8, (11) reduces to}
\]

\[
\mu^C \in \left( \frac{1 + g - \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2}, \frac{1 + g + \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2} \right).
\]

(12)

Proof. Equation 8 implies that \(\beta(\gamma, \varepsilon, \mu^C)^{K} = 1 - \mu^C\). Combining this with Inequality 11 gives

\[
(1 - \mu^C) \mu^C > \frac{g}{1 - \varepsilon} - g \mu^C.
\]

Solving this inequality for \(\mu^C\) provides the desired expression.
Proposition 4 follows from combining the feasibility constraint given by Equation 8 in Lemma 4 and the incentive constraints given by (11) in Lemma 6 and (12) in Corollary 1.

**OA.7 Proof of Part 2 of Theorem 4**

**Theorem 4** (Part 2). For $g < 1$ and $l \leq g/(1 - g)$,

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

*Proof.* The case $l < g/(1 - g)$ was already handled in A.6. Here we handle the case $l = g/(1 - g)$, or equivalently $l/(1 + l) = g$. We show that there exists some $\varepsilon > 0$ such that $\limsup_{\gamma \to 1} \pi_K^C(\gamma, \varepsilon) = 0$ for all $\varepsilon < \varepsilon$. Suppose that $\limsup_{\gamma \to 1} \pi_K^C(\gamma, \varepsilon) = \mu_K^C(\varepsilon) > 0$ for some $\varepsilon$. Then there is some $\gamma_n \to 1$ and a sequence of associated equilibria with share of cooperators $\mu_K^C(\gamma_n, \varepsilon)$ such that $\lim_{n \to \infty} \mu_K^C(\gamma_n, \varepsilon) = \mu_K^C(\varepsilon)$. Such a sequence must satisfy the $(C|C)_0$ and $(D|D)_{K-1}$ constraints for each corresponding $\gamma_n$. Taking the limit of these constraints as $n \to \infty$ shows that $\mu_K^C(\varepsilon)$ must satisfy the following “limit” constraints.

$$\text{Limit } (C|C)_0 : \mu_K^C(\varepsilon) \in \left[ \frac{1 + g - \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2}, \frac{1 + g + \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2} \right],$$

$$\text{Limit } (D|D)_{K-1} : \mu_K^C(\varepsilon) \leq \frac{1}{1 - \varepsilon} g.$$

We show that the function

$$q(\varepsilon) := \frac{1 + g - \sqrt{(1 + g)^2 - \frac{4}{1 - \varepsilon} g}}{2} - \frac{1}{1 - \varepsilon} g$$

is strictly positive for all sufficiently small, but non-zero, $\varepsilon$, which precludes $\mu_K^C(\varepsilon)$ satisfying the above “limit” constraints for such $\varepsilon$. To see that $q(\varepsilon)$ for sufficiently small, but non-zero, $\varepsilon$, note that $q(\varepsilon) = 0$, while the $\varepsilon$ derivative of $q$ evaluated at
\[ \varepsilon = 0 \text{ is } \frac{dq}{d\varepsilon}(0) = g\left(\frac{1}{1-g} - 1\right) > 0, \]

where the inequality comes from \( 0 < g < 1 \).

**OA.8 Proof of Lemma 9**

**Lemma 9.** Fix \( \varepsilon \in (0,1) \). For all \( \Delta > 0 \), there exists \( \gamma < 1 \) such that, for all \( \gamma > \gamma \) 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 \( K \).

Let \( \tilde{K} : (0,1) \times (0,1) \times (0,1) \to \mathbb{R}_+ \) be the function given by

\[
\tilde{K}(\gamma, \varepsilon, \mu^C) = \frac{\ln(1 - \mu^C)}{\ln(\beta(\gamma, \varepsilon, \mu^C))}. \tag{13}
\]

By construction, \( \tilde{K}(\gamma, \varepsilon, \mu^C) \) is the unique \( K \in \mathbb{R}_+ \) such that \( \mu^C = 1 - \beta(\gamma, \varepsilon, \mu^C)K \).

Let \( d : (0,1] \times (0,1) \times (0,1) \to \mathbb{R} \) be the function given by

\[
d(\gamma, \varepsilon, \mu^C) = \begin{cases} 
1 + \ln(1 - \mu^C)(1 - \mu^C) \frac{\partial \beta}{\partial \mu^C}(\gamma, \varepsilon, \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C)) & \text{if } \gamma < 1 \\
1 + \frac{(1-\varepsilon)\ln(1-\mu^C)(1-\mu^C)}{1-(1-\varepsilon)\mu^C} & \text{if } \gamma = 1
\end{cases}
\]

The \( \mu^C \) derivative of \( \tilde{K}(\gamma, \varepsilon, \mu^C) \) is related to \( d(\gamma, \varepsilon, \mu^C) \) in the following lemma.

**Lemma 7.** \( \tilde{K} : (0,1) \times (0,1) \times (0,1) \to \mathbb{R}_+ \) is differentiable in \( \mu^C \) with derivative given by

\[
\frac{\partial \tilde{K}}{\partial \mu^C}(\gamma, \varepsilon, \mu^C) = -\frac{d(\gamma, \varepsilon, \mu^C)}{(1 - \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C))}.
\]

**Proof of Lemma 7.** From Equation 13, it follows that \( \tilde{K}(\gamma, \varepsilon, \mu^C) \) is differentiable in
\( \mu^C \) with derivative given by

\[
\frac{\partial \tilde{K}}{\partial \mu^C} (\gamma, \varepsilon, \mu^C) = - \frac{\ln(\beta(\gamma, \varepsilon, \mu^C)) + \ln(1 - \mu^C) \frac{\partial \beta}{\partial \mu^C} (\gamma, \varepsilon, \mu^C)}{\ln(\beta(\gamma, \varepsilon, \mu^C))^2} = - \frac{1 + \ln(1 - \mu^C)(1 - \mu^C) \frac{\partial \beta}{\partial \mu^C} (\gamma, \varepsilon, \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C))}{(1 - \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C))} = - \frac{d(\gamma, \varepsilon, \mu^C)}{(1 - \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C))}. 
\]

The following two lemmas concern properties of \( d(\gamma, \varepsilon, \mu^C) \) that will be useful for the proof of Lemma 9.

**Lemma 8.** \( d : (0, 1] \times (0, 1) \times (0, 1) \rightarrow \mathbb{R} \) is well-defined and continuous.

**Proof of Lemma 8.** Since \( \beta(\gamma, \varepsilon, \mu^C) \) is differentiable and only takes values in \((0, 1)\), it follows that \( d(\gamma, \varepsilon, \mu^C) \) is well-defined. Moreover, since \( \beta(\gamma, \varepsilon, \mu^C) \) is continuously differentiable for all \((\gamma, \mu^C) \in (0, 1) \times (0, 1)\), \( d(\gamma, \varepsilon, \mu^C) \) is continuous for \( \gamma < 1 \). All that remains is to check that \( d(\gamma, \varepsilon, \mu^C) \) is continuous for \( \gamma = 1 \).

First, note that \( d(1, \varepsilon, \mu^C) \) is continuous in \( \mu^C \). Thus, we need only check the limit in which \( \gamma \) approaches 1, but never equals 1. Note that

\[
\frac{\partial \beta}{\partial \mu^C} (\gamma, \varepsilon, \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C)) = - \frac{\gamma(1 - \varepsilon)(1 - \gamma)}{(1 - \gamma(1 - \varepsilon)\mu^C)^2} \ln(\beta(\gamma, \varepsilon, \mu^C)) \bigg( \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{\ln(\beta(\gamma, \varepsilon, \mu^C))} \bigg). \tag{14}
\]

It is clear that

\[
\lim_{\gamma \neq 1} \frac{\gamma(1 - \varepsilon)}{(1 - \gamma(1 - \varepsilon)\mu^C)} = \frac{1 - \varepsilon}{(1 - (1 - \varepsilon)\mu^C)} \tag{15}
\]

for all \( \mu^C \in (0, 1) \). For \( \gamma \) close to 1,

\[
\ln(\beta(\gamma, \varepsilon, \mu^C)) = \beta(\gamma, \varepsilon, \mu^C) - 1 + O((\beta(\gamma, \varepsilon, \mu^C) - 1)^2).
\]
Thus,
\[
\lim_{\gamma, \mu \to (1, \mu^C)} \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{\ln(\beta(\gamma, \varepsilon, \mu^C))} = -1
\]
for all \( \mu^C \in (0, 1) \). Equations 14, 15, and 16 together imply that \( d(\gamma, \varepsilon, \mu^C) \) is continuous for \( \gamma = 1 \). \hfill \blacksquare

**Lemma 9.** For any fixed \( \varepsilon \), \( d(1, \varepsilon, \mu^C) \) has at most two zeros in \( \mu^C \in (0, 1) \).

**Proof of Lemma 9.** It suffices to show that
\[
\frac{\ln(1 - \mu^C)(1 - \mu^C)}{1 - (1 - \varepsilon)\mu^C}
\]
is single-peaked in \( \mu^C \in (0, 1) \). Note that
\[
\frac{\partial}{\partial \mu^C} \left[ \frac{\ln(1 - \mu^C)(1 - \mu^C)}{1 - (1 - \varepsilon)\mu^C} \right] = \frac{(1 - \varepsilon)\mu^C - \varepsilon \ln(1 - \mu^C) - 1}{(1 - (1 - \varepsilon)\mu^C)^2}.
\]
The single-peakedness of \( \ln(1 - \mu^C)(1 - \mu^C)/(1 - (1 - \varepsilon)\mu^C) \) follows from \( (1 - \varepsilon)\mu^C - \varepsilon \ln(1 - \mu^C) - 1 \) being increasing in \( \mu^C \). \hfill \blacksquare

With these preliminaries established, we now present the proof of Lemma 9.

**Proof of Lemma 9.** Fix \( \varepsilon \in (0, 1) \). Lemma 9 says \( d(1, \varepsilon, \mu^C) \) has at most two zeros for \( \mu^C \in (0, 1) \). Because of this, there exists \( \mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \mu_6 \in (0, 1) \) satisfying
\[
0 < \mu_1 < \mu_2 < \mu_3 < \mu_4 < \mu_5 < \mu_6 < 1 \text{ such that }
\]
\[
\min\{|\mu^C - \mu_1|, |\mu^C - \mu_2|, |\mu^C - \mu_3|, |\mu^C - \mu_4|, |\mu^C - \mu_5|, |\mu^C - \mu_6|\} < \Delta/2
\]
for all \( \mu^C \in [0, 1] \), and \( d(1, \mu^C) \) is non-zero on the intervals \([\mu_1, \mu_2]\), \([\mu_3, \mu_4]\), and \([\mu_5, \mu_6]\). Equation 17 says that the interval endpoints can be chosen so that \( M \) is no farther than \( \Delta/2 \) from any \( \mu^C \in [0, 1] \), while the second condition implies that
\[
|d(1, \mu^C)| > 0
\]
for all $\mu^C \in M$.

Lemma 8 says $d(\gamma, \varepsilon, \mu^C)$ is continuous for $(\gamma, \mu^C) \in (0, 1) \times (0, 1)$. Hence, $d(\gamma, \varepsilon, \mu^C)$ is uniformly continuous for $(\gamma, \mu^C) \in [\gamma, 1] \times M$ for any $\gamma > 0$. Equation 18 then implies that there exists some $\lambda > 0$ and $\tilde{\gamma} \in (0, 1)$ such that $|d(\gamma, \varepsilon, \mu^C)/(1 - \mu^C)| > \lambda$ for all $\gamma > \tilde{\gamma}$ and $\mu^C \in M$.

Define $\eta \in (0, 1)$ to be

$$\eta = \min \left\{ \frac{\mu_2 - \mu_1}{2}, \frac{\mu_4 - \mu_3}{2}, \frac{\mu_6 - \mu_5}{2}, \frac{\Delta}{2} \right\}.$$  

Because $\lim_{\gamma \to 1} \min_{\mu^C \in [0, 1]} \beta(\gamma, \varepsilon, \mu^C) = 1$, there exists $\gamma' \in (0, 1)$ such that $|\ln(\beta(\gamma, \varepsilon, \mu^C))| < \lambda \eta$ for all $\gamma > \gamma'$ and $\mu^C \in M$.

Moreover, $\lim_{\gamma \to 1} \min_{\mu^C \in [0, 1]} \beta(\gamma, \varepsilon, \mu^C) = 1$ implies that there exists $\hat{\gamma} \in (0, 1)$ such that $\tilde{K}(\gamma, \varepsilon, \mu^C) \geq 1$ for all $\gamma > \hat{\gamma}$ and $\mu^C \in M$.

Let $\gamma = \max\{\tilde{\gamma}, \gamma', \hat{\gamma}\}$. Thus, $|d(\gamma, \varepsilon, \mu^C)/(1 - \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C))| > 1/\eta$ and $\tilde{K}(\gamma, \varepsilon, \mu^C) \geq 1$ for all $\gamma > \gamma$ and $\mu^C \in M$. For the remainder of the proof, fix $\gamma \in (\gamma, 1)$. We now show that, for a given $\mu^C \in M$, there exists some $\tilde{\mu}^C \in M$ and non-negative integer $\hat{K}$ such that $|\tilde{\mu}^C - \mu^C| < \Delta/2$ and $\tilde{\mu}^C = 1 - \beta(\gamma, \varepsilon, \tilde{\mu})\hat{K}$. This, when combined with Equation 17, completes the proof.

Fix $\mu^C \in M$. Suppose for concreteness that $\mu^C \in [\mu_1, \mu_2]$. An identical argument handles the case when $\mu^C \in [\mu_3, \mu_4] \cup [\mu_5, \mu_6]$. By construction, $\eta$ is weakly smaller than both $(\mu_2 - \mu_1)/2$ and $\Delta/2$. Therefore, there is some $\tilde{\mu}^C \in [\mu_1, \mu_2]$ such that $\eta \leq |\tilde{\mu}^C - \mu^C| \leq \Delta/2$. Because $|d(\gamma, \varepsilon, \mu^C)/(1 - \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C))| > 1/\eta$ for all $\mu^C \in M$, it follows from Lemma 7 that $|\partial/\partial \mu^C(\gamma, \varepsilon, \mu^C)| > 1/\eta$ for all $\mu^C \in M$. Hence, $|\tilde{K}(\gamma, \varepsilon, \tilde{\mu}^C) - \tilde{K}(\gamma, \varepsilon, \mu^C)| > 1$. It thus follows that there exists some $\tilde{\mu}^C$ between $\mu^C$ and $\tilde{\mu}^C$ and some non-negative integer $\hat{K}$ between $\tilde{K}(\gamma, \varepsilon, \mu^C)$ and $\tilde{K}(\gamma, \varepsilon, \tilde{\mu}^C)$ such that $\tilde{K}(\gamma, \varepsilon, \tilde{\mu}^C) = \hat{K}$. Thus, $|\tilde{\mu}^C - \mu^C| < \Delta/2$ and $\tilde{\mu}^C = 1 - \beta(\gamma, \varepsilon, \tilde{\mu}^C)\hat{K}$. ■
Proof of Proposition 5

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:

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

2. Incentives:

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

Lemma 10. In a GrimKL steady state with total share of cooperators $\mu^C$,

\[
\mu_k = \begin{cases} 
\beta(\gamma, \varepsilon, \mu^C)^k(1 - \beta(\gamma, \varepsilon, \mu^C)) & \text{if } 0 \leq k \leq K - 1 \\
\alpha(\gamma, \varepsilon)^{K-k} \beta(\gamma, \varepsilon, \mu^C)^K(1 - \alpha(\gamma, \varepsilon)) & \text{if } K \leq k \leq K + L - 1 \\
\gamma^{k-K-L} \alpha(\gamma, \varepsilon)^L \beta(\gamma, \varepsilon, \mu^C)^K(1 - \gamma) & \text{if } k \geq K
\end{cases}
\]

Moreover, $\mu^C$ satisfies the equation

\[
\mu^C = 1 - \alpha(\gamma, \varepsilon)^L \beta(\gamma, \varepsilon, \mu^C)^K,
\]

$\mu^R$ satisfies the equation

\[
\mu^R = 1 - \beta(\gamma, \varepsilon, \mu^C)^K,
\]
and $\mu_{UC}$ satisfies the equation

$$
\mu_{UC} = (1 - \alpha(\gamma, \varepsilon)^L) \beta(\gamma, \varepsilon, \mu_C)^K.
$$

Proof. We establish the first part of this result. Since $i_0 = 1 - \gamma$ and $\tau_0 = \gamma(1 - (1 - \varepsilon)\mu_C)\mu_0$, Equation 7 implies

$$
\mu_0 = \frac{1 - \gamma}{1 - \gamma(1 - \varepsilon)\mu_C} = 1 - \beta(\gamma, \varepsilon, \mu_C).
$$

Moreover, by Lemmas 4 and 5, $i_{k+1} = \gamma(1 - (1 - \varepsilon)\mu_C)\mu_k$ and $\tau_{k+1} = \gamma(1 - (1 - \varepsilon)\mu_C)\mu_{k+1}$ for $0 \leq k \leq K - 1$. Thus, Equation 7 implies $\mu_{k+1} = \beta(\gamma, \varepsilon, \mu_C)\mu_k$ for $0 \leq k \leq K - 1$.

By induction, $\mu_k = \beta(\gamma, \varepsilon, \mu_C)^k(1 - \beta(\gamma, \varepsilon, \mu_C))$ for $0 \leq k \leq K - 1$.

By Lemmas 4 and 5, $i_K = \gamma(1 - (1 - \varepsilon)\mu_C)\mu_{K-1}$ and $\tau_K = \gamma\varepsilon\mu_K$, so Equation 7 implies

$$
\mu_K = \frac{\gamma(1 - (1 - \varepsilon)\mu_C)(1 - \beta(\gamma, \varepsilon, \mu_C))}{1 - \gamma(1 - \varepsilon)\mu_C} = \beta(\gamma, \varepsilon, \mu_C)^{K-1}(1 - \alpha(\gamma, \varepsilon)).
$$

Likewise, by Lemmas 4 and 5, $i_{k+1} = \gamma\varepsilon\mu_k$ and $\tau_{k+1} = \gamma\varepsilon\mu_{k+1}$ for $K \leq k \leq K + L - 1$. Hence, Equation 7 implies $\mu_{k+1} = \alpha(\gamma, \varepsilon)\mu_k$ for $K \leq k \leq K + L - 1$. Combining this with the previously derived $\mu_K = \beta(\gamma, \varepsilon, \mu_C)^K(1 - \alpha(\gamma, \varepsilon))$ and applying induction gives $\mu_K = \alpha(\gamma, \varepsilon)^{K-1}\beta(\gamma, \varepsilon, \mu_C)^K(1 - \alpha(\gamma, \varepsilon))$ for $K \leq k \leq K + L - 1$.

By Lemmas 4 and 5, $i_{K+L} = \gamma\varepsilon\mu_{K+L-1}$ and $\tau_{K+L} = \gamma\mu_{K+L}$, so Equation 7 implies

$$
\mu_{K+L} = \gamma\varepsilon\mu_{K+L-1} = \alpha(\gamma, \varepsilon)^L\beta(\gamma, \varepsilon)^K(1 - \gamma).
$$

Likewise, by Lemmas 4 and 5, $i_{k+1} = \gamma\mu_k$ and $\tau_{k+1} = \gamma\mu_{k+1}$ for $k \geq K + L$. 21
Hence, Equation 7 implies $\mu_{k+1} = \gamma \mu_k$ for $k \geq K + L$. Combining this with the previously derived $\mu_{K+L} = \alpha(\gamma, \varepsilon)^L \beta(\gamma, \varepsilon, \mu^C)^K (1 - \gamma)$ and applying induction gives $\mu_k = \gamma^{k-K-L} \alpha(\gamma, \varepsilon)^L \beta(\gamma, \varepsilon, \mu^C)^K (1 - \gamma)$ for $k \geq K + L$.

Now we establish the second part of the result. Using Equation 7, it follows that

$$
\mu^R = \sum_{k=0}^{K-1} \mu_k = \sum_{k=0}^{K-1} \beta(\gamma, \varepsilon, \mu^C)^k (1 - \beta(\gamma, \varepsilon, \mu^C))
= 1 - \beta(\gamma, \varepsilon, \mu^C)^K,
$$

$$
\mu^{UC} = \sum_{k=K}^{K+L-1} \mu_k = \sum_{k=K}^{K+L-1} \alpha(\gamma, \varepsilon)^{k-K} \beta(\gamma, \varepsilon, \mu^C)^K (1 - \alpha(\gamma, \varepsilon))
= (1 - \alpha(\gamma, \varepsilon)^L) \beta(\gamma, \varepsilon, \mu^C)^K,
$$

and

$$
\mu^C = \mu^R + \mu^{UC} = 1 - \alpha(\gamma, \varepsilon)^L \beta(\gamma, \varepsilon, \mu^C)^K,
$$

which establishes Equation 8.

**Lemma 11.** The value function of a player with record $k$ is

$$
V_k = \begin{cases} 
(1 - \beta(\gamma, \varepsilon, \mu^C)^K)\mu^C + \beta(\gamma, \varepsilon, \mu^C)^K (1 - \alpha(\gamma, \varepsilon)) L (\mu^C - \mu^D) 
+ \beta(\gamma, \varepsilon, \mu^C)^K \alpha(\gamma, \varepsilon)^L \mu^{UC} (1 + g) & \text{if } 0 \leq k \leq K - 1 \\
(1 - \alpha(\gamma, \varepsilon)^{K+L-k}) (\mu^C - \mu^D) + \alpha(\gamma, \varepsilon)^{K+L-k} \mu^{UC} (1 + g) & \text{if } K \leq k \leq K + L - 1 \\
\mu^{UC} (1 + g) & \text{if } k \geq K + L 
\end{cases}
$$

*Proof.* Players with record $k \geq K+L$ are defectors and obtain a flow payoff of $\mu^{UC} (1+g)$ in all future periods, so $V_k = \mu^{UC} (1+g)$ for $k \geq K + L$. Combining this with $V_k = (1 - \alpha(\gamma, \varepsilon)) (\mu^C - \mu^D) + \alpha(\gamma, \varepsilon) V_{k+1}$ for $0 \leq k \leq K + L - 1$ from Lemma 6 and solving inductively for $V_k$ gives

$$
V_k = (1 - \alpha(\gamma, \varepsilon)^{K+L-k}) (\mu^C - \mu^D) + \alpha(\gamma, \varepsilon)^{K+L-k} \mu^{UC} (1 + g)
$$
for $K \leq k \leq K + L - 1$. Finally, combining this with $V_k = (1 - \beta(\gamma, \varepsilon, \mu^C))\mu^C + \beta(\gamma, \varepsilon, \mu^C)V_{k+1}$ for $0 \leq k \leq K - 1$ from Lemma 6 and solving inductively for $V_k$ gives

$$V_k = (1 - \beta(\gamma, \varepsilon, \mu^C)^{K-k})\mu^C + \beta(\gamma, \varepsilon, \mu^C)^{K-k}(1 - \alpha(\gamma, \varepsilon)^L)(\mu^C - \mu^D l) + \beta(\gamma, \varepsilon, \mu^C)^{K-k}\alpha(\gamma, \varepsilon)^L\mu^{UC}(1 + g)$$

for $K \leq k \leq K + L - 1$. ■

**Lemma 12.**

1. The $(C|C)_0$ constraint is

$$\frac{1 - \varepsilon}{1 - (1 - \varepsilon)\mu^C}\beta(\gamma, \varepsilon, \mu^C)^K\left[\mu^C - (1 - \alpha(\gamma, \varepsilon)^L)(\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L\mu^{UC}(1 + g)\right] > g.$$  

2. The $(D|D)_{K-1}$ constraint is

$$\frac{\gamma(1 - \varepsilon)}{1 - \gamma(1 - \varepsilon)\mu^C}\left[\mu^C - (1 - \alpha(\gamma, \varepsilon)^L)(\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L\mu^{UC}(1 + g)\right] < l.$$  

3. The $(C|D)_K$ constraint is

$$\frac{1 - \varepsilon}{\varepsilon}\alpha(\gamma, \varepsilon)^L\left[\mu^C - \mu^D l - \mu^{UC}(1 + g)\right] > l.$$  

**Proof.** We first derive the $(C|C)_0$ constraint. From Lemma 11,

$$V_0 - V_1 = (1 - \beta(\gamma, \varepsilon, \mu^C))\beta(\gamma, \varepsilon, \mu^C)^{K-1}\left[\mu^C - (1 - \alpha(\gamma, \varepsilon)^L)(\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L\mu^{UC}(1 + g)\right].$$
Therefore,
\[
\gamma(1 - \varepsilon) \frac{V_0 - V_1}{1 - \gamma} = \gamma(1 - \varepsilon) \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{1 - \gamma} \beta(\gamma, \varepsilon, \mu^C) K^{-1} \left[ \mu^C - (1 - \alpha(\gamma, \varepsilon)^L) (\mu^C - \mu^D l) \right.
\]
\[
- \alpha(\gamma, \varepsilon)^L \mu^U C (1 + g) \left. \right]
\]
\[
= (1 - \varepsilon) \frac{\gamma}{1 - \gamma(1 - \varepsilon) \mu^C} \beta(\gamma, \varepsilon, \mu^C) K^{-1} \left[ \mu^C - (1 - \alpha(\gamma, \varepsilon)^L) (\mu^C - \mu^D l) \right.
\]
\[
- \alpha(\gamma, \varepsilon)^L \mu^U C (1 + g) \left. \right]
\]
\[
= \frac{1 - \varepsilon}{1 - (1 - \varepsilon) \mu^C} \beta(\gamma, \varepsilon, \mu^C)^K \left[ \mu^C - (1 - \alpha(\gamma, \varepsilon)^L) (\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L \mu^U C (1 + g) \right].
\]

Hence, the $(C|C)_0$ constraint is equivalent to
\[
\frac{1 - \varepsilon}{1 - (1 - \varepsilon) \mu^C} \beta(\gamma, \varepsilon, \mu^C)^K \left[ \mu^C - (1 - \alpha(\gamma, \varepsilon)^L) (\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L \mu^U C (1 + g) \right] > g.
\]

We now derive the $(D|D)_{K-1}$ constraint. From Lemma 11,
\[
V_{K-1} - V_K = (1 - \beta(\gamma, \varepsilon, \mu^C)) \left[ \mu^C - (1 - \alpha(\gamma, \varepsilon)^L) (\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L \mu^U C (1 + g) \right].
\]

Therefore,
\[
\gamma(1 - \varepsilon) \frac{V_{K-1} - V_K}{1 - \gamma} = \gamma(1 - \varepsilon) \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{1 - \gamma} \left[ \mu^C - (1 - \alpha(\gamma, \varepsilon)^L) (\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L \mu^U C (1 + g) \right]
\]
\[
= \gamma(1 - \varepsilon) \frac{1 - \beta(\gamma, \varepsilon, \mu^C)}{1 - \gamma} \mu^C \left[ \mu^C - (1 - \alpha(\gamma, \varepsilon)^L) (\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L \mu^U C (1 + g) \right].
\]

Hence, the $(D|D)_{K-1}$ constraint is equivalent to
\[
\frac{\gamma(1 - \varepsilon)}{1 - \gamma(1 - \varepsilon) \mu^C} \left[ \mu^C - (1 - \alpha(\gamma, \varepsilon)^L) (\mu^C - \mu^D l) - \alpha(\gamma, \varepsilon)^L \mu^U C (1 + g) \right] < l.
\]

We now derive the $(C|D)_K$ constraint. From Lemma 11,
\[
V_K - V_{K+1} = (1 - \alpha(\gamma, \varepsilon)) \alpha(\gamma, \varepsilon)^L l^{-1} \left[ \mu^C - \mu^D l - \mu^U C (1 + g) \right].
\]
Therefore,
\[
\gamma(1 - \varepsilon) \frac{V_{K} - V_{K+1}}{1 - \gamma} = \gamma(1 - \varepsilon) \frac{1 - \alpha(\gamma, \varepsilon)}{1 - \gamma} \alpha(\gamma, \varepsilon)^{L-1} \left[ \mu^{C} - \mu^{D} l - \mu^{UC} (1 + g) \right] \\
= (1 - \varepsilon) \frac{\gamma}{1 - \gamma(1 - \varepsilon)} \alpha(\gamma, \varepsilon)^{L-1} \left[ \mu^{C} - \mu^{D} l - \mu^{UC} (1 + g) \right] \\
= \frac{1 - \varepsilon}{\varepsilon} \alpha(\gamma, \varepsilon)^{L} \left[ \mu^{C} - \mu^{D} l - \mu^{UC} (1 + g) \right].
\]

Hence, the \((C|D)_{K}\) constraint is equivalent to
\[
\frac{1 - \varepsilon}{\varepsilon} \alpha(\gamma, \varepsilon)^{L} \left[ \mu^{C} - \mu^{D} l - \mu^{UC} (1 + g) \right] > l.
\]

\[\blacksquare\]

Corollary 2. When combined with the steady state conditions from Lemma 10,

1. The \((C|C)_{0}\) constraint reduces to
\[
\frac{(1 - \varepsilon)(1 - \mu^{C})}{1 - (1 - \varepsilon)\mu^{C}} \left[ \mu^{R} + \mu^{UC} (l - g) \right] > g.
\]

2. The \((D|D)_{K-1}\) constraint reduces to
\[
\gamma(1 - \varepsilon) \frac{(1 - \mu^{C})}{1 - \gamma(1 - \varepsilon)\mu^{C}} \left[ \mu^{R} + \mu^{UC} (l - g) \right] + \mu^{R} l < l.
\]

3. The \((C|D)_{K}\) constraint reduces to
\[
\frac{(1 - \varepsilon)(1 - \mu^{C})}{1 - (1 - \varepsilon)\mu^{C}} \left[ \mu^{R} + \mu^{UC} (l - g) \right] + \mu^{R} l > l.
\]

Proof. The steady state condition from Lemma 10 implies that \(\beta(\gamma, \varepsilon, \mu^{C})^{K} = 1 - \mu^{R}\), \(1 - \alpha(\gamma, \varepsilon)^{L} = \mu^{UC} / (1 - \mu^{R})\), and \(\alpha(\gamma, \varepsilon)^{L} = (1 - \mu^{C}) / (1 - \mu^{R})\). Imposing these conditions on the \((C|C)_{0}\), \((D|D)_{K-1}\), and \((C|D)_{K}\) constraints in Lemma 12 and manipulating the various inequalities gives the inequalities in Corollary 5. \[\blacksquare\]
Let \( \rho : [0, 1] \times (0, 1) \times [0, 1] \to [0, 1) \) be the function given by

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

Equation 14 can be equivalently written as

\[
\rho(1, \varepsilon, \mu^C) \left[ (l - g)\mu^C + (1 + g - l)\mu^R \right] + l\mu^R = l.
\]

Setting \( \mu^C = h(\varepsilon, \mu^R) \) in the above equation and solving for \( \rho(1, \varepsilon, h(\varepsilon, \mu^R)) \) gives

\[
\rho(1, \varepsilon, h(\varepsilon, \mu^R)) = \frac{l(1 - \mu^R)}{(l - g)h(\varepsilon, \mu^R) + (1 + g - l)\mu^R}
\]

for all \( \varepsilon \) such that \( h(\varepsilon, \mu^R) \) is well-defined. Since \( \lim_{\varepsilon \to 0} h(\varepsilon, \mu^R) = 1 \), an immediate corollary follows.

**Corollary 3.** For every \( \mu^R \in (g/(1 + g), 1 - g/l) \),

\[
\lim_{\varepsilon \to 0} \rho(1, \varepsilon, h(\varepsilon, \mu^R)) = \frac{l(1 - \mu^R)}{l - g + (1 + g - l)\mu^R}.
\]

**OA.10.1 Proof of Lemma 10**

**Lemma 10.** 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 \( \liminf_{\gamma \to 1} \overline{\mu}^C_{KL}(\gamma, \varepsilon) \geq h(\varepsilon, \mu^R) \) for \( \varepsilon < \varepsilon \).

Define the function \( I : [0, 1] \times (0, 1) \times [0, 1] \times [0, 1] \to \mathbb{R} \) by

\[
I(\gamma, \varepsilon, \mu^C, \mu^R) = \rho(\gamma, \varepsilon, \mu^C)((l - g)\mu^C + (1 + g - l)\mu^R) + l\mu^R.
\]

The \((D|D)_{K-1}\) constraint is equivalent to \( I(\gamma, \varepsilon, \mu^C, \mu^R) < l \), and the \((C|D)_{K}\) constraint
is equivalent to $I(1, \varepsilon, \mu^C, \mu^R) > l$. The $(C|C)_0$ constraint holds whenever the $(C|D)_K$ constraint holds and $\mu^R \leq 1 - g/l$, which is true for the profiles we consider.

**Lemma 13.** Fix $\mu^R \in (g/(1 + g), 1 - g/l]$. There exists $\varepsilon > 0$ such that

$$\frac{\partial I}{\partial \mu^C}(1, \varepsilon, h(\varepsilon, \mu^R), \mu^R) < 0 < \frac{\partial I}{\partial \mu^R}(1, \varepsilon, h(\varepsilon, \mu^R), \mu^R)$$

for all $\varepsilon < \varepsilon$.

**Proof of Lemma 13.** Note that

$$\frac{\partial I}{\partial \mu^R}(1, \varepsilon, h(\varepsilon, \mu^R), \mu^R) = \rho(1, \varepsilon, h(\varepsilon, \mu^R))(1 + g - l) + l > \rho(1, \varepsilon, h(\varepsilon, \mu^R))(1 + g) > 0,$$

since $0 < \rho(1, \varepsilon, h(\varepsilon, \mu^R)) < 1$.

Moreover,

$$\frac{\partial I}{\partial \mu^C}(1, \varepsilon, h(\varepsilon, \mu^R), \mu^R) = - \left( \frac{1}{1 - h(\varepsilon, \mu^R)} \right) \left( \frac{\varepsilon}{1 - (1 - \varepsilon) h(\varepsilon, \mu^R)} \right)$$

$$\rho(1, \varepsilon, h(\varepsilon, \mu^R))(1 + g - l)\mu^R + (l - g)h(\varepsilon, \mu^R)$$

$$+ \rho(1, \varepsilon, h(\varepsilon, \mu^R))(l - g)$$

$$= - \left( \frac{1}{1 - h(\varepsilon, \mu^R)} \right) \left( \frac{\varepsilon}{1 - (1 - \varepsilon) h(\varepsilon, \mu^R)} \right) l(1 - \mu^R)$$

$$+ \rho(1, \varepsilon, h(\varepsilon, \mu^R))(l - g).$$

Since $\lim_{\varepsilon \to 0} h(\varepsilon, \mu^R) = 1$ and

$$\lim_{\varepsilon \to 0} \frac{\varepsilon}{1 - (1 - \varepsilon) h(\varepsilon, \mu^R)} = \frac{(1 + g)\mu^R - g}{(1 + g - l)\mu^R + l - g},$$

it follows that

$$\lim_{\varepsilon \to 0} \frac{\partial I}{\partial \mu^C}(1, \varepsilon, h(\varepsilon, \mu^R), \mu^R) = -\infty.$$
Thus, there exists some $\varepsilon > 0$ such that

$$\frac{\partial I}{\partial \mu^C}(1, \varepsilon, h(\varepsilon, \mu^R), \mu^R) < 0$$

for all $\varepsilon < \varepsilon$. ■

Let $\tilde{K}: (0,1) \times (0,1) \times (0,1) \times (0,1) \rightarrow \mathbb{R}$ be the function given by

$$\tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R) = \frac{\ln(1 - \mu^R)}{\ln(\beta(\gamma, \varepsilon, \mu^C))}, \quad (19)$$

and $\tilde{L}: (0,1) \times (0,1) \times (0,1) \times (0,1) \rightarrow \mathbb{R}_+$ be the function given by

$$\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R) = \frac{\ln(1 - \mu^C) - \ln(1 - \mu^R)}{\ln(\alpha(\gamma, \varepsilon))}. \quad (20)$$

Note that $\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R) \geq 0$ whenever $\mu^C \geq \mu^R$, which is the case of interest. By construction, $\tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R)$ and $\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)$ are the unique $(K, L) \in \mathbb{R}_+ \times \mathbb{R}_+$ such that the feasibility constraints in Proposition 5 are satisfied.

Differentiating Equations 19 and 20 gives the following result.

**Lemma 14.** $\tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R)$ and $\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)$ are differentiable in $(\mu^C, \mu^R) \in (0,1) \times (0,1)$ with partial derivatives

$$\frac{\partial \tilde{K}}{\partial \mu^C}(\gamma, \varepsilon, \mu^C, \mu^R) = -\frac{\ln(1 - \mu^R) \frac{\partial \beta}{\partial \mu}(\gamma, \varepsilon, \mu^C)}{\ln(\beta(\gamma, \varepsilon, \mu^C))^2 \beta(\gamma, \varepsilon, \mu^C)},$$

$$\frac{\partial \tilde{L}}{\partial \mu^C}(\gamma, \varepsilon, \mu^C, \mu^R) = -\frac{1}{(1 - \mu^C) \ln(\alpha(\gamma, \varepsilon))},$$

$$\frac{\partial \tilde{K}}{\partial \mu^R}(\gamma, \varepsilon, \mu^C, \mu^R) = -\frac{1}{(1 - \mu^R) \ln(\beta(\gamma, \varepsilon, \mu^C))},$$

$$\frac{\partial \tilde{L}}{\partial \mu^R}(\gamma, \varepsilon, \mu^C, \mu^R) = \frac{1}{(1 - \mu^R) \ln(\alpha(\gamma, \varepsilon))}.$$
of $\tilde{K}$ and $\tilde{L}$. That is,

$$J(\gamma, \varepsilon, \mu^C, \mu^R) = \begin{bmatrix} \frac{\partial \tilde{K}}{\partial \mu^C}(\gamma, \varepsilon, \mu^C, \mu^R) & \frac{\partial \tilde{L}}{\partial \mu^C}(\gamma, \varepsilon, \mu^C, \mu^R) \\ \frac{\partial \tilde{K}}{\partial \mu^R}(\gamma, \varepsilon, \mu^C, \mu^R) & \frac{\partial \tilde{L}}{\partial \mu^R}(\gamma, \varepsilon, \mu^C, \mu^R) \end{bmatrix}$$

$$= \begin{bmatrix} -\ln(1-\mu^R) & \frac{\ln(1-\mu^R)}{\ln(\beta(\gamma, \varepsilon, \mu^C))} \\ \ln(1-\mu^C) & -\frac{1}{(1-\mu^C) \ln(\alpha(\gamma, \varepsilon))} \end{bmatrix} \cdot \begin{bmatrix} \frac{\ln(1-\mu^R)}{\ln(\beta(\gamma, \varepsilon, \mu^C))} \beta(\gamma, \varepsilon, \mu^C) \\ \ln(1-\mu^C) \rho(1, \varepsilon, \mu^C) \end{bmatrix}.$$

Let $\zeta : [0, 1] \times (0, 1) \times (0, 1) \times (0, 1) \to \mathbb{R}$ be the function given by

$$\zeta(\gamma, \varepsilon, \mu^C, \mu^R) = \begin{cases} \ln(1-\mu^R) \frac{(1-\mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C))}{\beta(\gamma, \varepsilon, \mu^C) \ln(\alpha(\gamma, \varepsilon))} & \text{if } \gamma < 1 \\ \ln(1-\mu^R) \rho(1, \varepsilon, \mu^C) & \text{if } \gamma = 1 \end{cases}.$$

The following lemma comes from direct calculation.

**Lemma 15.**

1. The determinant of $J(\gamma, \varepsilon, \mu^C, \mu^R)$ is

$$\det(J(\gamma, \varepsilon, \mu^C, \mu^R)) = -\frac{1}{(1-\mu^C) \ln(\alpha(\gamma, \varepsilon)) \ln(\beta(\gamma, \varepsilon, \mu^C))}.$$

2. When $J(\gamma, \varepsilon, \mu^C, \mu^R)$ is invertible, its inverse is

$$J(\gamma, \varepsilon, \mu^C, \mu^R)^{-1} = \begin{bmatrix} \frac{(1-\mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C))}{1+\zeta(\gamma, \varepsilon, \mu^C, \mu^R)} & \frac{(1-\mu^C) \ln(\alpha(\gamma, \varepsilon))}{1+\zeta(\gamma, \varepsilon, \mu^C, \mu^R)} \\ \frac{(1-\mu^R) \ln(\beta(\gamma, \varepsilon, \mu^C))}{1+\zeta(\gamma, \varepsilon, \mu^C, \mu^R)} & \frac{(1-\mu^R) \ln(\alpha(\gamma, \varepsilon))}{1+\zeta(\gamma, \varepsilon, \mu^C, \mu^R)} \end{bmatrix}.$$

We establish the continuity of $\zeta(\gamma, \varepsilon, \mu^R, \mu^C)$.

**Lemma 16.** For all $\varepsilon \in (0, 1)$, $\zeta(\gamma, \varepsilon, \mu^C, \mu^R)$ is continuous in $(\gamma, \mu^C, \mu^R)$.

**Proof of Lemma 16.** Clearly, $\zeta(\gamma, \varepsilon, \mu^C, \mu^R)$ is continuous whenever $\gamma < 1$. What remains is to show that it is continuous when $\gamma = 1$. Note that $\ln(1-\mu^R) \rho(1, \varepsilon, \mu^C)$ is continuous in $(\mu^C, \mu^R)$. Thus, we need only check the limit in which $\gamma$ approaches 1,
but never equals 1. Recall that
\[
\frac{\partial \beta}{\partial \mu}(\gamma, \varepsilon, \mu^C) \ln(\beta(\gamma, \varepsilon, \mu^C)) = -\frac{\gamma(1-\varepsilon)(1-\gamma)(1-\gamma(1-\varepsilon)\mu^C)^2}{\beta(\gamma, \varepsilon, \mu^C)(1-\gamma(1-\varepsilon)\mu^C)}.
\]

It is clear that
\[
\lim_{\gamma, \mu \to (1, \mu^C)} \frac{\gamma(1-\varepsilon)}{\beta(\gamma, \varepsilon, \mu)(1-\gamma(1-\varepsilon)\mu)} = \frac{1-\varepsilon}{(1-(1-\varepsilon)\mu^C)}
\]
for all $\mu^C \in (0, 1)$. For $\gamma$ close to 1,
\[
\ln(\beta(\gamma, \varepsilon, \mu^C)) = \beta(\gamma, \varepsilon, \mu^C) - 1 + O((\beta(\gamma, \varepsilon, \mu^C) - 1)^2).
\]
Thus,
\[
\lim_{\gamma, \mu \to (1, \mu^C)} \frac{1-\beta(\gamma, \varepsilon, \mu)}{\ln(\beta(\gamma, \varepsilon, \mu))} = -1
\]
for all $\mu^C \in (0, 1)$. Combining these results, it follows that
\[
\lim_{\gamma, \mu \to (1, \mu^C)} \frac{(1-\mu)^{\partial \beta}{\partial \mu}(\gamma, \varepsilon, \mu)}{\beta(\gamma, \varepsilon, \mu) \ln(\beta(\gamma, \varepsilon, \mu))} = \rho(1, \varepsilon, \mu^C)
\]
for all $\mu^C \in (0, 1)$. Hence, $\zeta(\gamma, \varepsilon, \mu^C, \mu^R)$ is continuous.

The following lemma concerns the extent to which, for small $\varepsilon$ and fixed $\hat{\mu}^R \in (g/(1+g), 1-g/l]$, profiles $(\mu^C, \mu^R)$ near $(h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)$ are close to feasible profiles. It combines Lemmas 15 and 16 with the inverse function theorem to obtain a bound on how far such $(\mu^C, \mu^R)$ are from feasible profiles when the corresponding value of $\tilde{L}$ is an integer. Moreover, the size of this bound is related to the magnitude of $1 + \zeta(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)$, which is close to $|1 + \kappa(\hat{\mu}^R)|$ for small $\varepsilon$.

**Lemma 17.** Fix $\hat{\mu}^R \in (g/(1+g), 1-g/l]$ and $\eta > 0$. If $|1 + \kappa(\hat{\mu}^R)| > \lambda$ for some
\( \lambda > 0 \), there exists some \( \bar{\varepsilon} > 0 \) such that, for all \( \varepsilon < \bar{\varepsilon} \), there exists some \( \bar{\gamma} < 1 \) and an open neighborhood of \((h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)\), \( M \), such that, for all \( \gamma > \bar{\gamma} \) and \((\mu^C, \mu^R) \in M\), whenever \( L = \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R) \) is an integer, there exists some feasible \( \tilde{\mu}^C \) and \( \tilde{\mu}^R \) such that

\[
0 \leq \tilde{\mu}^C - \mu^C < -\frac{1 + \eta}{\lambda} (1 - h(\varepsilon, \hat{\mu}^R)) \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R))),
\]
\[
0 \leq \tilde{\mu}^R - \mu^R < -\frac{1 + \eta}{\lambda} (1 - \hat{\mu}^R) \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R))).
\]

Proof of Lemma 17. We handle the case where \( 1 + \kappa(\hat{\mu}^R) > \lambda > 0 \). The case where \( 1 + \kappa(\hat{\mu}^R) < -\lambda < 0 \) can be handled analogously.

Note that

\[
1 + \zeta(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R) = 1 + \ln(1 - \hat{\mu}^R) \rho(1, \varepsilon, h(\varepsilon, \hat{\mu}^R)).
\]

Moreover,

\[
\lim_{\varepsilon \to 0} \ln(1 - \hat{\mu}^R) \rho(1, \varepsilon, h(\varepsilon, \hat{\mu}^R)) = \kappa(\hat{\mu}^R)
\]

by Lemma 3. Thus, when \( 1 + \kappa(\hat{\mu}^R) > \lambda \), there exists some \( \bar{\varepsilon} > 0 \) such that, for all \( \varepsilon < \bar{\varepsilon} \), there exists \( \bar{\gamma}_1 < 1 \) and an open neighborhood of \((h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)\), \( M_1 \), such that

\[
1 + \zeta(\gamma, \varepsilon, \mu^C, \mu^R) < -\lambda
\]

for all \( \gamma > \bar{\gamma}_1 \) and \((\mu^C, \mu^R) \in M_1\). By Lemma 15, \( J(\gamma, \varepsilon, \mu^C, \mu^R) \) is invertible for all such points. Thus, for a given \( \varepsilon < \bar{\varepsilon} \) and \( \gamma > \bar{\gamma}_1 \), the inverse function theorem implies the existence of differentiable functions of \((K, L)\), \( \tilde{\mu}^C \) and \( \tilde{\mu}^R \), that constitute a local inverse of \( \tilde{K} \) and \( \tilde{L} \) for \((\mu^C, \mu^R) \in M_1\). Additionally, the partial derivatives of these
functions are given by $J^{-1}$, so that

\[
\frac{\partial \tilde{\mu}^C}{\partial K}(\gamma, \varepsilon, K, L) = \frac{(1 - \tilde{\mu}^C(\gamma, \varepsilon, K, L)) \ln(\beta(\gamma, \varepsilon, \tilde{\mu}^C(\gamma, \varepsilon, K, L)))}{1 + \zeta(\gamma, \varepsilon, \tilde{\mu}^C(\gamma, \varepsilon, K, L), \tilde{\mu}^R(\gamma, \varepsilon, K, L))},
\]

\[
\frac{\partial \tilde{\mu}^R}{\partial K}(\gamma, \varepsilon, K, L) = \frac{(1 - \tilde{\mu}^R(\gamma, \varepsilon, K, L)) \ln(\beta(\gamma, \varepsilon, \tilde{\mu}^R(\gamma, \varepsilon, K, L)))}{1 + \zeta(\gamma, \varepsilon, \tilde{\mu}^C(\gamma, \varepsilon, K, L), \tilde{\mu}^R(\gamma, \varepsilon, K, L))},
\]

\[
\frac{\partial \tilde{\mu}^C}{\partial L}(\gamma, \varepsilon, K, L) = \frac{\zeta(\gamma, \varepsilon, \tilde{\mu}^C(\gamma, \varepsilon, K, L), \tilde{\mu}^R(\gamma, \varepsilon, K, L))(1 - \tilde{\mu}^R(\gamma, \varepsilon, K, L)) \ln(\alpha(\gamma, \varepsilon))}{1 + \zeta(\gamma, \varepsilon, \tilde{\mu}^C(\gamma, \varepsilon, K, L), \tilde{\mu}^R(\gamma, \varepsilon, K, L))},
\]

\[
\frac{\partial \tilde{\mu}^R}{\partial L}(\gamma, \varepsilon, K, L) = \frac{\zeta(\gamma, \varepsilon, \tilde{\mu}^C(\gamma, \varepsilon, K, L), \tilde{\mu}^R(\gamma, \varepsilon, K, L))(1 - \tilde{\mu}^R(\gamma, \varepsilon, K, L)) \ln(\alpha(\gamma, \varepsilon))}{1 + \zeta(\gamma, \varepsilon, \tilde{\mu}^C(\gamma, \varepsilon, K, L), \tilde{\mu}^R(\gamma, \varepsilon, K, L))},
\]

for any $(K, L)$ that equals $(\tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R), \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^C))$ for some $(\mu^C, \mu^R) \in M_1$.

There is a neighborhood of $(h(\varepsilon, \tilde{\mu}^R), \tilde{\mu}^R)$, $M_2$, such that

\[1 - \mu^C < \sqrt{1 + \eta(1 - h(\varepsilon, \tilde{\mu}^R))}\]

and

\[1 - \mu^R < \sqrt{1 + \eta(1 - \tilde{\mu}^R)}\]

for all $(\mu^C, \mu^R) \in M_2$. Moreover, because $\beta(\gamma, \varepsilon, \mu^C)$ is decreasing in $\mu^C$ and

\[\lim_{\gamma \to 1} \frac{\ln(\beta(\gamma, \varepsilon, \mu^C_1))}{\ln(\beta(\gamma, \varepsilon, \mu^C_2))} = \frac{1 - (1 - \varepsilon)\mu^C_2}{1 - (1 - \varepsilon)\mu^C_1}\]

for all $(\gamma, \varepsilon) \in (0, 1) \times (0, 1)$ and $\mu^C_1, \mu^C_2 \in [0, 1]$, we can take the neighborhood $M_2$ to be small enough so that

\[\ln(\beta(\gamma, \varepsilon, \mu^C)) > \sqrt{1 + \eta \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \tilde{\mu}^R)))}\]

for all $(\mu^C, \mu^R) \in M$ and $\gamma > \gamma_2$ for some sufficiently high $\gamma_2 < 1$.

Combining the expression for the partial derivatives of $\tilde{\mu}^C$ and $\tilde{\mu}^R$ with these in-
equalities gives
\[
\frac{1 + \eta}{\lambda} (1 - h(\varepsilon, \hat{\mu}^R)) \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R))) < \frac{\partial \hat{\mu}^C}{\partial K}(\gamma, \varepsilon, K, L) < 0,
\]
\[
\frac{1 + \eta}{\lambda} (1 - \hat{\mu}(\varepsilon, \hat{\mu}^R)) \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R))) < \frac{\partial \hat{\mu}^R}{\partial K}(\gamma, \varepsilon, K, L) < 0,
\]
\[
\frac{1 + \eta}{\lambda} (1 - h(\varepsilon, \hat{\mu}^R)) \ln(\alpha(\gamma, \varepsilon)) < \frac{\partial \hat{\mu}^C}{\partial L}(\gamma, \varepsilon, K, L) < 0,
\]
\[
\frac{(1 + \eta)(\lambda + 1)}{\lambda} (1 - \hat{\mu}^R) \ln(\alpha(\gamma, \varepsilon)) < \frac{\partial \hat{\mu}^R}{\partial L}(\gamma, \varepsilon, K, L) < 0,
\]
for all \(\gamma > \max\{\gamma_1, \gamma_2\}\) and any \((K, L)\) that equals \((\tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R), \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R))\)
for some \((\mu^C, \mu^R) \in M_1 \cap M_2\).

Along with the mean value theorem, these bounds on the partial derivatives of \(\hat{\mu}^C\) and \(\hat{\mu}^R\) imply that there exists some \(\overline{\gamma} < 1\) and some open neighborhood of \((h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)\), \(M\), such that
\[
0 \leq \hat{\mu}^C(\gamma, \varepsilon, \tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R)], \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)) - \mu^C < -\frac{1 + \eta}{\lambda} (1 - h(\varepsilon, \hat{\mu}^R)) \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R))),
\]
\[
0 \leq \hat{\mu}^R(\gamma, \varepsilon, \tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R)], \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)) - \mu^R < -\frac{1 + \eta}{\lambda} (1 - \hat{\mu}^R) \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R))),
\]
for all \(\gamma > \overline{\gamma}\) and \((\mu^C, \mu^R) \in M\).

Lemma 17 then follows by noting that \(\hat{\mu}^C(\gamma, \varepsilon, \tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R)], \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R))\) and \(\hat{\mu}^R(\gamma, \varepsilon, \tilde{K}(\gamma, \varepsilon, \mu^C, \mu^R)], \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R))\) is feasible whenever \(\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)\) is an integer.

Fix \(\hat{\mu}^R \in (g/(1 + g), 1 - g/l], \eta > 0,\) and \(\lambda > 0\). Let \(J^C_{\hat{\mu}^R, \eta, \lambda} : [0, 1] \times (0, 1) \times (0, 1) \to \mathbb{R}\) be the function given by
\[
J^C_{\hat{\mu}^R, \eta, \lambda}(\gamma, \varepsilon, \mu^C, \mu^R) = I\left(1, \varepsilon, \mu^C - \frac{1 + \eta}{\lambda} (1 - h(\varepsilon, \hat{\mu}^R)) \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R))), \mu^R\right), \tag{21}
\]
and \(J^D_{\hat{\mu}^R, \eta, \lambda} : [0, 1] \times (0, 1) \times (0, 1) \times (0, 1) \to \mathbb{R}\) be the function given by
\[
J^D_{\hat{\mu}^R, \eta, \lambda}(\gamma, \varepsilon, \mu^C, \mu^R) = I\left(\gamma, \varepsilon, \mu^C, \mu^R - \frac{1 + \eta}{\lambda} (1 - \hat{\mu}^R) \ln(\beta(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R)))\right). \tag{22}
\]
Combining Lemmas 13 and 17, it follows that, if $|1 + \kappa(\hat{\mu}^R)| > \lambda$, there exists some $\bar{c} > 0$ such that, for all $\varepsilon < \bar{c}$ and $\eta > 0$, there exists $\bar{\gamma} < 1$ and an open neighborhood of $(h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)$, $M$, such that, for all $\gamma > \bar{\gamma}$ and $(\mu^C, \mu^R) \in M$, whenever $L = \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)$ is a non-negative integer, the feasible profile $(\hat{\mu}^C, \hat{\mu}^R)$ described in Lemma 17 is such that $I(1, \varepsilon, \hat{\mu}^C, \hat{\mu}^R) \geq J_{\mu^C, \eta, \lambda}^{C}(\gamma, \varepsilon, \mu^C, \mu^R)$ and $I(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R) \leq J_{\hat{\mu}^R, \eta, \lambda}^{D}(\gamma, \varepsilon, \mu^C, \mu^R)$.

Next we give conditions under which the $\gamma$ partial derivatives of $J_{\hat{\mu}^R, \eta, \lambda}^{C}(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R)$ and $J_{\hat{\mu}^R, \eta, \lambda}^{D}(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R)$ evaluated at $(\gamma, \mu^C, \mu^R) = (1, \hat{\mu}^R)$ are both strictly negative, and are such that the $\gamma$ partial derivative of $J_{\hat{\mu}^R, \eta, \lambda}^{D}(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R)$ is strictly less than that of $J_{\hat{\mu}^R, \eta, \lambda}^{C}(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R)$. An implication of this is that, for all sufficiently high $\gamma$, there is a $(\mu^C, \mu^R)$ isocurve of $I(1, \gamma, \mu^C, \mu^R)$ in $M$ such that $J_{\hat{\mu}^R, \eta, \lambda}^{D}(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R) < 0 < J_{\hat{\mu}^R, \eta, \lambda}^{C}(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R)$ for all $(\mu^C, \mu^R)$ on the isocurve.

**Lemma 18.** Fix $\hat{\mu}^R \in (g/(1 + g), 1 - g/l)$. If there is some $\lambda$ such that $|1 + \kappa(\hat{\mu}^R)| > \lambda > \iota(\hat{\mu}^R)$, then there exists some $\eta > 0$ and $\bar{c} > 0$ such that, for all $\varepsilon < \bar{c}$,

$$0 < \frac{\partial J_{\hat{\mu}^R, \eta, \lambda}^{C}(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R)}{\partial \gamma} < \frac{\partial J_{\hat{\mu}^R, \eta, \lambda}^{D}(\gamma, \varepsilon, \hat{\mu}^C, \hat{\mu}^R)}{\partial \gamma}.$$

**Proof of Lemma 18.** Differentiating Equation 21, we find that

$${\frac{\partial J_{\hat{\mu}^R, \eta, \lambda}^{C}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)}{\partial \gamma} = -\frac{1 + \eta}{\lambda} \frac{1 - h(\varepsilon, \hat{\mu}^R)}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)} \frac{\partial I}{\partial \mu^C}(1, \varepsilon, \hat{\mu}^R)}{\partial \gamma} = \frac{l(1 - \hat{\mu}^R)}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)} \varepsilon \left( \frac{1 + \eta}{\lambda} \frac{1 - h(\varepsilon, \hat{\mu}^R)}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)} \right).$$
Differentiating Equation 22, we find that

\[
\frac{\partial J_{\hat{\mu}^R, \eta, \lambda}}{\partial \gamma}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R) = \frac{\partial I}{\partial \gamma}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R) - \frac{1 + \eta}{\lambda} \frac{1 - \hat{\mu}^R}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)} \frac{\partial I}{\partial \mu^R}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)
\]

\[
= \frac{l(1 - \hat{\mu}^R)}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)} \left(1 - \frac{1 + \eta}{1 - \varepsilon} \rho(1, \varepsilon, h(\varepsilon, \hat{\mu}^R))(1 + g - l) + 1 \right).
\]

Note that

\[
\lim_{\varepsilon \to 0} \frac{l(1 - \hat{\mu}^R)}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)} \frac{\partial J_{\hat{\mu}^R, \eta, \lambda}}{\partial \gamma}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R) = \frac{1 + \eta}{\lambda} \frac{(1 + g)\hat{\mu}^R - g}{(1 + g - l)\hat{\mu}^R + l - g},
\]

and

\[
\lim_{\varepsilon \to 0} \frac{l(1 - \hat{\mu}^R)}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)} \frac{\partial J_{\hat{\mu}^R, \eta, \lambda}}{\partial \gamma}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R) = 1 - \frac{1 + \eta}{\lambda} \left( \frac{1}{(1 + g - l)\hat{\mu}^R + l - g} \right).
\]

When \( \lambda > \iota(\hat{\mu}^R), \)

\[
1 - \frac{1}{\lambda} \left( \frac{1}{(1 + g - l)\hat{\mu}^R + l - g} \right) > \frac{1}{\lambda} \frac{(1 + g)\hat{\mu}^R - g}{(1 + g - l)\hat{\mu}^R + l - g} > 0,
\]

so there is some \( \eta > 0 \) such that

\[
1 - \frac{1 + \eta}{\lambda} \left( \frac{1}{(1 + g - l)\hat{\mu}^R + l - g} \right) > \frac{1 + \eta}{\lambda} \frac{(1 + g)\hat{\mu}^R - g}{(1 + g - l)\hat{\mu}^R + l - g} > 0.
\]

Thus, for such an \( \eta \), there exists some \( \varepsilon > 0 \) such that

\[
0 < \frac{\partial J_{\hat{\mu}^R, \eta, \lambda}}{\partial \gamma}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R) < \frac{\partial J_{\hat{\mu}^R, \eta, \lambda}}{\partial \gamma}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)
\]

for all \( \varepsilon < \varepsilon \).

\[\blacksquare\]

**Lemma 19.** Fix \( \hat{\mu}^R \in (g/(1 + g), 1 - g/l] \). There exists some \( \varepsilon > 0 \) such that, for all \( \varepsilon < \varepsilon \), the isocurves of \( \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R) \) and \( I(1, \varepsilon, \mu^C, \mu^R) \) are not tangent at \( (h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R) \).
Proof of Lemma 19. By Lemma 14, we the isocurve of \( \tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R) \) has slope

\[
\frac{d\mu^C}{d\mu^R} = - \frac{\partial \tilde{L}}{\partial \mu^R}(\gamma, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)
\]

\[
= \frac{1 - h(\varepsilon, \hat{\mu}^R)}{1 - \hat{\mu}^R}
\]

at \((h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)\).

Likewise, we find that the isocurve of \( I(1, \varepsilon, \mu^C, \mu^R) \) has slope

\[
\frac{d\mu^C}{d\mu^R} = - \frac{\partial I}{\partial \mu^R}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)
\]

\[
= \frac{\rho(1, \varepsilon, h(\varepsilon, \hat{\mu}^R))(1 + g - l) + l}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)l - (1 - h(\varepsilon, \hat{\mu}^R))(1 - \hat{\mu}^R)}\frac{(1 - h(\varepsilon, \hat{\mu}^R))(l - g)}{1 - \hat{\mu}^R}
\]

at \((h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R)\).

Since

\[
\lim_{\varepsilon \to 0} \frac{\rho(1, \varepsilon, h(\varepsilon, \hat{\mu}^R))(1 + g - l) + l}{1 - (1 - \varepsilon)h(\varepsilon, \hat{\mu}^R)l - (1 - h(\varepsilon, \hat{\mu}^R))(1 - \hat{\mu}^R)}\frac{(1 - h(\varepsilon, \hat{\mu}^R))(l - g)}{1 - \hat{\mu}^R} = \frac{1}{(1 + g)\hat{\mu}^R - g} > 1,
\]

the result follows.

Combining Lemmas 18 and 19 gives the following result.

**Lemma 20.** Fix \( \hat{\mu}^R \in (g/(1 + g), 1 - g/l] \). If \( |1 + \kappa(\hat{\mu}^R)| > \iota(\hat{\mu}^R) \), there exists some \( \varepsilon > 0 \) such that, for all \( \varepsilon < \varepsilon \) and all open neighborhoods of \((h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R), M, \) there exists \( \gamma < 1 \) such that, for all \( \gamma > \gamma \), there is a feasible \((\mu^C, \mu^R) \in M \) that satisfies the incentive constraints.

**Proof of Lemma 20.** By Lemma 18, there exists some \( \gamma < 1 \), sufficiently small neighborhood of \((\mu^C, \mu^R) = (h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R), M, \) and \( \eta_1, \eta_2 > 0 \) such that

\[
0 < \frac{\partial J^C_{\hat{\mu}^R, \eta, \lambda}}{\partial \gamma}(\gamma, \varepsilon, \mu^C, \mu^R) < \eta_1 < \eta_2 < \frac{\partial J^D_{\hat{\mu}^R, \eta, \lambda}}{\partial \gamma}(\gamma, \varepsilon, \mu^C, \mu^R)
\]
for all \((\mu^C, \mu^R) \in M\) and \(\gamma > \overline{\gamma}\). Therefore,

\[
J^C_{\mu^R, \eta, \lambda}(\gamma, \varepsilon, \mu^C, \mu^R) \geq J^C_{\mu^R, \eta, \lambda}(1, \varepsilon, \mu^C, \mu^R) - \eta_1(1 - \gamma)
\]

\[
= I(1, \varepsilon, \mu^C, \mu^R) - \eta_1(1 - \gamma)
\]

\[
J^D_{\mu^R, \eta, \lambda}(\gamma, \varepsilon, \mu^C, \mu^R) \leq J^D_{\mu^R, \eta, \lambda}(1, \varepsilon, \mu^C, \mu^R) - \eta_2(1 - \gamma)
\]

\[
= I(1, \varepsilon, \mu^C, \mu^R) - \eta_2(1 - \gamma)
\]

for all \((\mu^C, \mu^R) \in M\) and \(\gamma > \overline{\gamma}\). It thus follows that if there is some \((\mu^C, \mu^R) \in M\) such that \(\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)\) is a non-negative integer and that satisfies \(\eta_1(1 - \gamma) < I(1, \varepsilon, \mu^C, \mu^R) < \eta_2(1 - \gamma)\) and \(\mu^R \leq 1 - g/l\), then \((\hat{\mu}^C(\gamma, \varepsilon, \mu^C, \mu^R), \hat{\mu}^R(\gamma, \varepsilon, \mu^C, \mu^R))\) is both feasible and satisfies all of the incentive constraints for \(\gamma\).

All that remains is to show that, for all \(\gamma > \overline{\gamma}\), there exists some \((\mu^C, \mu^R) \in M\) for which these conditions are met. Because

\[
\frac{\partial I}{\partial \mu^C}(1, \varepsilon, h(\varepsilon, \hat{\mu}^R), \hat{\mu}^R) < 0,
\]

it follows that, for sufficiently large \(\gamma\), isocurves of the form \(I(1, \varepsilon, \mu^C, \mu^R) = (\eta_1 + \eta_2)/2(1 - \gamma)\) intersect \(M\) for every \(\mu^R\) in an open neighborhood of \(1 - g/l\). By Lemma 19, the isocurves of \(I(1, \varepsilon, \mu^C, \mu^R)\) and \(\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)\) are not tangent. Because the \(\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)\) isocurves do not depend on \(\gamma\) and

\[
\lim_{\gamma \to 1} \ln(\alpha(\gamma, \varepsilon)) = 0,
\]

it follows by Lemma 14 that there exists some \((\mu^C, \mu^R) \in M\) on the isocurve \(I(1, \varepsilon, \mu^C, \mu^R) = (\eta_1 + \eta_2)/2(1 - \gamma)\) that satisfies \(\mu^R \leq 1 - g/l\) and is such that \(\tilde{L}(\gamma, \varepsilon, \mu^C, \mu^R)\) is a non-negative integer for sufficiently large \(\gamma\).

Lemma 10 is an immediate consequence of Lemma 20.
OA.10.2 Proof of Lemma 11

Lemma 11. 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 \( l > \max\{g(g+1), b(g)\} \).

Lemma 11 is a consequence of the following lemma.

Lemma 21. Suppose \( l > g(g+1) \). Some \( \mu^R \in (g/(1+g), 1-g/l] \) satisfies 
\[ |1 + \kappa(\mu^R)| > \iota(\mu^R) \] if any of the following conditions hold.

1. \( g < e - 1 \) and 
   \[ l > \frac{1 + g}{1 - \ln(1 + g)}. \]

2. \( g > e - 1 \) and 
   \[ l > \frac{1 + g}{\ln(1 + g) - 1}. \]

3. For some \( \phi > 1, g < e^\phi - 1, l \geq e^\phi g, \) and 
   \[ l > \frac{3e^\phi - 2 - 2g}{\phi - 1}. \]

Proof of Lemma 21. We handle Cases 2 and 3. The proof for Case 1 is similar to that for Case 2.

Suppose that \( g > e - 1 \) and \( l > (1 + g)/(\ln(1 + g) - 1) \). Note that 
\[ \lim_{\mu^R \to \frac{g}{1+g}} |1 + \kappa(\mu^R)| - \iota(\mu^R) = \ln(1 + g) - 1 - \frac{1 + g}{l}. \]

Since \( l > (1 + g)/(\ln(1 + g) - 1) \), \( \ln(1 + g) - 1 - (1 + g)/l > 0 \), and the result follows.

Suppose that, for some \( \phi > 1, g < e^\phi - 1, l \geq e^\phi g \) and \( l > (3e^\phi g - 2 - 2g)/(\phi - 1) \).

Note that \( g/(1 + g) < 1 - e^{-\phi} \leq 1 - g/l \) and
\[ |1 + \kappa(1 - e^{-\phi})| - \iota(1 - e^{-\phi}) = \frac{|l(\phi - 1) - e^\phi + 1 + g|}{e^\phi - 1 - g + l} + \frac{-2e^\phi + 1 + g}{e^\phi - 1 - g + l}. \]
Since $l > (3e^\phi - 2 - 2g)/(\phi - 1)$, $|l(\phi - 1) - e^\phi + 1 + g| - 2e^\phi + 1 + g > 0$, and the result follows.

Applying the special case where $\phi = 1.56$ to Lemma 21 and noting that, for $\phi = 1.56$,

$$g \geq e^\phi - 1 \text{ or } e^\phi g > \frac{3e^\phi - 2 - 2g}{\phi - 1}$$

only when $(1+g)/|\ln(1+g)-1| < (3e^\phi - 2 - 2g)/(\phi-1)$ or $g(g+1) > (1+g)/|\ln(1+g)-1|$
gives Lemma 11.

**OA.11 Another Family of Strategies**

Fix positive integers $K_1$, $K_2$, $K_3$, $K_4$ and consider the following strategy: If $0 \leq k \leq K_1 - 1$, the player is a conditional cooperator; if $k_1 \leq K \leq k_1 + k_2 - 1$, the player is a defector, if $k_1 + k_2 \leq K \leq k_1 + k_2 + k_3 - 1$, the player is a conditional cooperator, if $k_1 + k_2 - k_3 \leq K \leq k_1 + k_2 + k_3 + k_4 - 1$, the player is an unconditional cooperator, and if $k \geq K_1 + K_2 + K_3 + K_4$, the player is a defector.

Combining Equations 7, 4, and 5 and using induction shows that the steady-state record shares are

$$\mu_k = \begin{cases} 
\beta(\gamma, \varepsilon, \mu^C)^k(1 - \beta(\gamma, \varepsilon, \mu^C)) & \text{if } 0 \leq k \leq K_1 - 1 \\
\gamma^{k-K_1}\beta(\gamma, \varepsilon, \mu^C)^{K_1}(1 - \gamma) & \text{if } 0 \leq k - K_1 \leq K_2 - 1 \\
\gamma^{K_2}\beta(\gamma, \varepsilon, \mu^C)^{k-K_2-K_3}\beta(\gamma, \varepsilon, \mu^C)^{K_1+K_3}(1 - \alpha(\gamma, \varepsilon)) & \text{if } 0 \leq k - K_1 - K_2 \leq K_3 - 1 \\
\gamma^{K_2}\alpha(\gamma, \varepsilon)^{K_4}\beta(\gamma, \varepsilon, \mu^C)^{K_1+K_3} & \text{if } k \geq K_1 + K_2 + K_3 + K_4
\end{cases}$$
Thus,

$$
\mu_{CC} = \sum_{k=0}^{k_1-1} \mu_k + \sum_{k=k_1+k_2}^{k_1+k_2+k_3-1} \mu_k
\]

$$

$$
= \left[ \sum_{k=0}^{k_1-1} \beta(\gamma, \epsilon, \mu_C)^k + \sum_{k=k_1+k_2}^{k_1+k_2+k_3-1} \gamma^{k_2} \beta(\gamma, \epsilon, \mu_C)^{k_2-1} \right] (1 - \beta(\gamma, \epsilon, \mu_C)) (1 - \beta(\gamma, \epsilon, \mu_C))
\]

$$

$$
= 1 - \beta(\gamma, \epsilon, \mu_C)^{k_1} + \gamma^{k_2} \beta(\gamma, \epsilon, \mu_C)^{k_1+k_2} (1 - \beta(\gamma, \epsilon, \mu_C)^{k_2})
\]

$$

(23)

$$
and

\[
\mu_{UC} = \sum_{k=k_1+k_2+k_3}^{k_1+k_2+k_3+k_4-1} \mu_k
\]

$$

$$
= \sum_{k=k_1+k_2+k_3}^{k_1+k_2+k_3+k_4-1} \gamma^{k_2} \alpha(\gamma, \epsilon)^{k_1-k_2-k_3} \beta(\gamma, \epsilon, \mu_C)^{k_1+k_3} (1 - \alpha(\gamma, \epsilon))
\]

$$

$$
= \gamma^{k_2} (1 - \alpha(\gamma, \epsilon)^{k_1}) \beta(\gamma, \epsilon, \mu_C)^{k_1+k_3}
\]

$$

(24)

Equations 23 and 24 give

$$
\mu_C = \mu_{CC} + \mu_{UC}
\]

$$

(25)

The only incentive constraints that need to be checked are \((C|C)_0\), \((D|D)_{K_1-1}\), \((C|C)_{K_1+K_2}\), \((D|D)_{K_1+K_2+K_3-1}\), and \((C|D)_{K_1+K_2+K_3}\). By Lemma 3, these are

$$
(C|C)_0 : \gamma(1 - \epsilon)(V_0 - V_1) > (1 - \gamma)g,
\]

$$

(26)

$$
(D|D)_{K_1-1} : \gamma(1 - \epsilon)(V_{K_1-1} - V_{K_1}) < (1 - \gamma)l,
\]

$$

\]

$$
(C|C)_{K_1+K_2} : \gamma(1 - \epsilon)(V_{K_1+K_2} - V_{K_1+K_2+1}) > (1 - \gamma)g,
\]

$$

\]

$$
(D|D)_{K_1+K_2+K_3-1} : \gamma(1 - \epsilon)(V_{K_1+K_2+K_3-1} - V_{K_1+K_2+K_3}) < (1 - \gamma)l,
\]

$$

\]

$$
(C|D)_{K_1+K_2+K_3} : \gamma(1 - \epsilon)(V_{K_1+K_2+K_3} - V_{K_1+K_2+K_3+1}) > (1 - \gamma)l.
\]

$$

To check these incentive constraints, it suffices to compute the relevant value func-
tions by performing the following calculations sequentially:

\[
\begin{align*}
V_{K_1+K_2+K_3+K_4} &= \mu_{UC}(1 + g), \\
V_{K_1+K_2+K_3+1} &= (1 - \alpha(\gamma, \varepsilon)^{K_4-1})(\mu_C - (1 - \mu_C)l) + \alpha(\gamma, \varepsilon)^{K_4-1}V_{K_1+K_2+K_3+K_4}, \\
V_{K_1+K_2+K_3} &= (1 - \alpha(\gamma, \varepsilon)^{K_4})(\mu_C - (1 - \mu_C)l) + \alpha(\gamma, \varepsilon)^{K_4}V_{K_1+K_2+K_3+K_4}, \\
V_{K_1+K_2+K_3-1} &= (1 - \beta(\gamma, \varepsilon, \mu_C))\mu_C + \beta(\gamma, \varepsilon, \mu_C)V_{K_1+K_2+K_3}, \\
V_{K_1+K_2+1} &= (1 - \beta(\gamma, \varepsilon, \mu_C)^{K_4-1})\mu_C + \beta(\gamma, \varepsilon, \mu_C)^{K_4-1}V_{K_1+K_2+K_3}, \\
V_{K_1+K_2} &= (1 - \beta(\gamma, \varepsilon, \mu_C)^{K_3-1})\mu_C + \beta(\gamma, \varepsilon, \mu_C)^{K_3-1}V_{K_1+K_2+K_3}, \\
V_{K_1+1} &= (1 - \gamma K_2)\mu_{UC}(1 + g) + \gamma K_2 V_{K_1+K_2}, \\
V_{K_1-1} &= (1 - \beta(\gamma, \varepsilon, \mu_C)\mu_C + \beta(\gamma, \varepsilon, \mu_C)\mu_C)\mu_C + \beta(\gamma, \varepsilon, \mu_C)\mu_C V_{K_1}, \\
V_0 &= (1 - \beta(\gamma, \varepsilon, \mu_C)^{K_1})\mu_C + \beta(\gamma, \varepsilon, \mu_C)^{K_1}V_{K_1}.
\end{align*}
\]  

The validity of these equations comes from combining Lemma 6 with recursion and the fact that every player with record \( k \geq K_1 + K_2 + K_3 + K_4 \) is a defector who faces a flow payoff of \( \mu_{UC}(1 + g) \) in every future period.

When \( g = 1.0001, \quad l = 2, \quad \gamma = .99999, \) and \( \varepsilon = .0000000001, \) we numerically verified that the strategy with \( K_1 = 1, \quad K_2 = 2, \quad K_3 = 1, \) and \( K_4 = 1 \) has a steady state satisfying Equations 24 and 25 with \( \mu_C \approx .999984 \) and \( \mu_{UC} \approx .378686. \) With these values, we calculated the relevant value functions using Equation 27 and showed that the constraints in Inequality 26 were satisfied. Thus, this strategy has an equilibrium even when \( g > 1 \) and \( l < g(g + 1). \)

\section*{OA.12 Stochastic Transitions}

\subsection*{OA.12.1 Stochastic GrimK}

This subsection shows that a stochastic version of GrimK strategies can support full limit cooperation whenever \( g < 1. \)
We use the following record-keeping system: There are two possible records, 0 and 1. Newborn players have record 0. When a player with record 0 plays $D$, her record transitions to 1 with probability $\chi$. When a player with record 0 plays $C$, her record transitions to 1 with probability $\varepsilon \chi$. Record 1 is absorbing.

We consider Grim1 strategies under this record-keeping system: A player plays $C$ if and only if both she and her opponent have record 0.

**Theorem 1.** Fix parameters $(g,l,\varepsilon,\gamma)$. There exists $\chi \in (0,1)$ such that GRIM1 is a strict equilibrium with steady-state cooperation share $\mu^C > 0$ if and only if the following conditions hold.

1. **Feasibility:**

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

2. **Incentives:**

   \[
   \begin{align*}
   (C|C)_0 : & \quad \mu^C \in \left[ \frac{1 + g - \sqrt{(1 + g)^2 - 4 \frac{g}{1 - \varepsilon}}}{2}, \frac{1 + g + \sqrt{(1 + g)^2 - 4 \frac{g}{1 - \varepsilon}}}{2} \right], \\
   (D|D)_0 : & \quad \mu^C \notin \left[ \frac{1 + l - \sqrt{(1 + l)^2 - 4 \frac{l}{1 - \varepsilon}}}{2}, \frac{1 + l + \sqrt{(1 + l)^2 - 4 \frac{l}{1 - \varepsilon}}}{2} \right].
   \end{align*}
   \]

Moreover, letting $\bar{\mu}^C(\gamma,\varepsilon)$ be the maximal level of $\mu^C$ that can be supported for any choice of $\chi$, the following hold:

1. If $g < 1$, then $\lim_{\varepsilon \to 0} \lim_{\gamma \to 1} \bar{\mu}^C(\gamma,\varepsilon) = 1$.

2. If $g \geq 1$, then $\bar{\mu}^C(\gamma,\varepsilon) = 0$.

Note that $g < l$ implies that any $\mu^C$ that satisfies $(D|D)_0$ also satisfies $(C|C)_0$. In particular, $(D|D)_0$ never rules out the greatest level of $\mu^C$ that satisfies $(C|C)_0$. In
addition, \((C|C)_0\) and \((D|D)_0\) are independent of \(\gamma\), and Feasibility is always satisfied when \(\gamma\) is sufficiently large. Combined with the fact that the right endpoint of the interval describing \((C|C)_0\) converges to 1 as \(\gamma \to 1\) whenever \(g < 1\) (and is always at least 1 whenever \(g \geq 1\)), these observations imply that second part of the theorem follows immediately from the first.

**Proof.** Let \(\mu^C\) be the population share with record 0. Let \(V_C\) be the continuation value of a player with record 0. Note that the continuation value of a player with record 1 is 0. Therefore, \(V_C = (1 - \gamma)\mu^C + \gamma[1 - \chi (1 - (1 - \varepsilon)\mu^C)]V_C\), which is equivalent to

\[
V_C = \frac{(1 - \gamma)\mu^C}{1 - \gamma + \gamma \chi (1 - (1 - \varepsilon)\mu^C)}.
\]

On the other hand, the steady-state equation for \(\mu^C\) is

\[
1 - \gamma = (1 - \gamma)\mu^C + \gamma \chi (1 - (1 - \varepsilon)\mu^C)\mu^C,
\]

which is equivalent to

\[
\mu^C = \frac{1 - \gamma}{1 - \gamma + \gamma \chi (1 - (1 - \varepsilon)\mu^C)}.
\]

These equations imply \(V_C = (\mu^C)^2\).

It will be helpful to solve (28) for \(\chi\):

\[
\chi = \frac{(1 - \gamma) (1 - \mu^C)}{\gamma (1 - (1 - \varepsilon)\mu^C)\mu^C}.
\]

Feasibility requires that this quantity is less than 1.

Next, note that \((C|C)_0\) is

\[
\gamma (1 - \varepsilon \chi)V_C > (1 - \gamma)g + \gamma (1 - \chi)V_C.
\]

The above results imply that this is equivalent to

\[
\chi > \frac{1 - \gamma}{\gamma} \frac{g}{1 - \varepsilon (\mu^C)^2}.
\]
Comparing this with (29), we see that $(C|C)_0$ holds iff
\[
\frac{(1 - \gamma)(1 - \mu^C)}{\gamma(1 - (1 - \varepsilon)\mu^C)\mu^C} > \frac{1 - \gamma}{\gamma} \frac{g}{1 - \varepsilon (\mu^C)^2}.
\]
Solving this gives
\[
\mu^C \in \left( \frac{1 + g - \sqrt{(1 + g)^2 - 4g}}{2}, \frac{1 + g + \sqrt{(1 + g)^2 - 4g}}{2} \right).
\]

Similarly, $(D|D)_0$ is $\gamma(1 - \varepsilon\chi)V_C < (1 - \gamma)l + \gamma(1 - \chi)V_C$. By the preceding results, this is equivalent to the corresponding constraint in Theorem 1.

Using (29) and solving the resulting inequality, $(D|D)_0$ holds iff the constraint in Theorem 1 holds.

\section*{OA.12.2 Stochastic GrimKL}

In this subsection, we show that a stochastic version of GrimKL strategies can support full limit cooperation whenever either $g < 1$ or $l > g(g + 1)$.

We use the following record-keeping system: There are three possible records, 0, 1, and 2. Newborn players have record 0. When a player with record 0 plays $D$, her record transitions to 1 with probability $\chi_1$, while her record transitions to 1 with probability $\varepsilon\chi_1$ when she plays $C$. When a player with record 1 plays $D$, her record transitions to 2 with probability $\chi_2$, while her record transitions to 2 with probability $\varepsilon\chi_2$ when she plays $C$. Record 2 is absorbing.

We consider GrimKL strategies under this record-keeping system, with $K = 1$ and $L = 1$: Players with record 0 are reciprocators, players with record 1 are unconditional cooperators, and players with record 2 are defectors.

\textbf{Theorem 2.} Fix parameters $(g, l, \varepsilon, \gamma)$. There exist $\chi_1 \in (0, 1)$ and $\chi_2 \in (0, 1)$ such that GrimKL with $K = 1$ and $L = 1$ is an equilibrium with steady-state cooperation shares $\mu^R$ and $\mu^{UC}$ if and only if
1. Feasibility:
\[ \max \left\{ \frac{(1-\gamma)(1-\mu_R)}{\gamma(1-(1-\epsilon)\mu_C)} \mu_R, \frac{(1-\gamma)\mu_D}{\gamma\epsilon \mu_{UC}} \right\} < 1. \]

2. Incentives:
\[ (C|C)_0 : \frac{(1-\epsilon)(1-\mu_C)}{1-(1-\epsilon)\mu_C} \left[ \mu_R + \mu_{UC}(l-g) \right] > g, \]
\[ (D|D)_0 : \frac{(1-\epsilon)(1-\mu_C)}{1-\gamma(1-\epsilon)\mu_C} \left[ \mu_R + \mu_{UC}(l-g) \right] < l, \]
\[ (C|D)_1 : \frac{(1-\epsilon)(1-\mu_C)}{\epsilon(1-\mu_R)} [\mu_R - g\mu_{UC} - l(1-\mu_C)] > l. \]

Moreover, letting \( \bar{\mu}^C(\gamma, \epsilon) \) be the maximal level of \( \mu_C \) that can be supported for any choice of \( \chi_1 \) and \( \chi_2 \), the following hold:

1. If \( g \geq 1 \) and \( l \geq g(1+g) \), then \( \bar{\mu}^C(\gamma, \epsilon) = 0 \).
2. If either \( g < 1 \) or \( l > g(1+g) \), then \( \lim_{\epsilon \to 0} \lim_{\gamma \to 1} \bar{\mu}^C(\gamma, \epsilon) = 1 \).

**Proof.** We first compute the value functions. For defectors, we have \( V^D = (1+g)\mu_{UC} \).

For unconditional cooperators, we have \( V^{UC} = (1-\gamma)(\mu_C - l\mu_D) + \gamma(1-\epsilon\chi_2)V^{UC} + \gamma\epsilon\chi_2V^D \), which is equivalent to
\[ V^{UC} = \frac{(1-\gamma)(\mu_C - l\mu_D) + \gamma\epsilon\chi_2(1+g)\mu_{UC}}{1-\gamma(1-\epsilon\chi_2)}. \]

For reciprocators, we have \( V^R = (1-\gamma)\mu_C + \gamma(1-\chi_1(1-(1-\epsilon)\mu_C))V^R + \gamma\chi_1(1-(1-\epsilon)\mu_C)V^{UC} \), which is equivalent to
\[ V^R = \frac{(1-\gamma)\mu_C + \gamma\chi_1(1-(1-\epsilon)\mu_C)V^{UC}}{1-\gamma + \gamma\chi_1(1-(1-\epsilon)\mu_C)}. \]

We next consider the steady-state equations. For \( \mu_R \), we have \( 1-\gamma = (1-\gamma + \ldots \)
\( \gamma \chi_1 (1 - (1 - \varepsilon) \mu^C) \mu^R \), or equivalently
\[
\mu^R = \frac{1 - \gamma}{1 - \gamma + \gamma \chi_1 (1 - (1 - \varepsilon) \mu^C)} \iff \chi_1 = \frac{(1 - \gamma) (1 - \mu^R)}{\gamma (1 - (1 - \varepsilon) \mu^C) \mu^R}.
\]

For \( \mu^{UC} \), we have
\[
\gamma \chi_1 (1 - (1 - \varepsilon) \mu^C) \mu^{R} = (1 - \gamma + \gamma \varepsilon \chi_2) \mu^{UC}.
\]

Using the above equation for \( \chi_1 \), we can solve for \( \chi_2 \) as
\[
\chi_2 = \frac{(1 - \gamma) \mu^D}{\gamma \varepsilon \mu^{UC}}.
\]

Note that Feasibility says that \( \chi_1 \) and \( \chi_2 \) must be less than 1.

We now consider the incentive constraints. The \( (C|C)_0 \) constraint is \( \gamma (1 - \varepsilon) \chi_1 (V^R - V^{UC}) > (1 - \gamma) g \), which is equivalent to
\[
V^R - V^{UC} > \frac{(1 - \gamma) g}{\gamma (1 - \varepsilon) \chi_1}.
\]

Note that
\[
V^R - V^{UC} = \frac{(1 - \gamma) \left( \mu^C - \mu^{UC} \right)}{1 - \gamma + \gamma \chi_1 (1 - (1 - \varepsilon) \mu^C)} = \mu^R \left( \mu^C - \mu^{UC} \right)
\]
and
\[
1 - \gamma (1 - \varepsilon \chi_2) = 1 - \gamma + \frac{(1 - \gamma) \mu^D}{\mu^{UC}} = (1 - \gamma) \left( 1 + \frac{\mu^D}{\mu^{UC}} \right) = (1 - \gamma) \frac{1 - \mu^R}{\mu^{UC}}.
\]

Therefore,
\[
\begin{align*}
V^{UC} &= \frac{(1 - \gamma) \left( \mu^C - l \mu^D \right) + \gamma \varepsilon \chi_2 (1 + g) \mu^{UC}}{1 - \gamma (1 - \varepsilon \chi_2)} = \frac{\mu^{UC}}{1 - \mu^R} \left( \mu^C - l \mu^D + \frac{\mu^D}{\mu^{UC}} (1 + g) \mu^{UC} \right) \\
&= \frac{\mu^{UC}}{1 - \mu^R} \left( 1 - \mu^D - l \mu^D + (1 + g) \mu^D \right) = \frac{\mu^{UC}}{1 - \mu^R} (1 - \left( l - g \right) \mu^D).
\end{align*}
\]
Thus, $(C|C)_0$ is equivalent to

$$
\mu^R \left( \mu^C - \frac{\mu^{UC}}{1 - \mu^R} (1 - (l - g) \mu^D) \right) > \frac{(1 - (1 - \delta) \mu^C) \mu^R g}{(1 - \delta) (1 - \mu^R)},
$$

which gives the corresponding constraint in Theorem 2.

Similarly, the $(D|D)_0$ constraint is $\gamma (1 - \varepsilon) \chi_1 (V^R - V^{UC}) < (1 - \gamma) l$, which is equivalent to the constraint given in Theorem 2 by the previous results.

Finally, the $(C|D)_1$ constraint is $\gamma (1 - \varepsilon) \chi_2 (V^{UC} - V^D) > (1 - \gamma) l$, which is equivalent to the constraint in Theorem 2 by the previous results.

We now consider the iterated limit. For fixed values of $\varepsilon, \mu^R, \mu^{UC}$, and $\mu^D$, Feasibility is satisfied for high enough $\gamma$. Thus, we can ignore Feasibility and simply ask when there exist $\mu^R, \mu^{UC}$, and $\mu^D$ that satisfy the $\varepsilon \to 0$ “limit” versions of the incentive constraints:

\begin{align*}
(C|C)_0 & : \quad \mu^R + (l - g) \mu^{UC} > g \\
(D|D)_0 & : \quad \mu^R + (l - g) \mu^{UC} < l \\
(C|D)_1 & : \quad \mu^R - g \mu^{UC} - l (1 - \mu^C) > 0.
\end{align*}

We show that if $g \geq 1$ and $l \leq g (1 + g)$ then these constraints cannot be satisfied for any values of $\mu^R, \mu^{UC}$, and $\mu^D$; while if $g < 1$ or $l > g (1 + g)$ then they can be satisfied for values of $\mu^R, \mu^{UC}$, and $\mu^D$ such that $\mu^D = 0$. This completes the proof.

Suppose $g \geq 1$ and $l \leq g (1 + g)$. Note that a necessary condition for $(C|D)_1$ is $\mu^R \geq \frac{g}{1 + g}$; otherwise, the left-hand side of $(C|D)_1$ must be negative. Now, if $\mu^R \geq \frac{g}{1 + g}$ and $g \geq 1$, for $(C|C)_0$ to hold it must be that

$$
\frac{g}{1 + g} + (l - g) \frac{1}{1 + g} > g \iff l > g (1 + g).
$$

Hence, if $l < g (1 + g)$ the constraints cannot be satisfied.

Now suppose $g < 1$ or $l > g (1 + g)$. If $\mu^C = 1$ then $\mu^R \geq \frac{g}{1 + g}$ is a sufficient
condition for \((D|D)_0\). We can therefore support an equilibrium with \(\mu^C = 1\) iff there exists \(\mu^R \geq \frac{g}{1+g}\) such that \(g < \mu^R + (l-g)\left(1-\mu^R\right) < l\), or equivalently

\[
2g - l < \mu^R (1 + g - l) < g.
\]

Consider three cases. First, if \(l = 1 + g\) then \(g < 1\), so \(2g - l < 0 < g\) and thus (30) is trivially satisfied.

Second, if \(l < 1 + g\), then (30) is equivalent to

\[
\frac{2g - l}{1 + g - l} < \mu^R < \frac{g}{1 + g - l}.
\]

In this case, note that \(\frac{g}{1+g} < \frac{g}{1+g-l}\), so there is a value of \(\mu^R\) satisfying the constraints iff \(\frac{2g-l}{1+g-l} < 1\), i.e. \(g < 1\). Thus, the constraints can be satisfied if \(l < 1 + g\) and \(g < 1\).

Third, if \(l > 1 + g\), then (30) is equivalent to

\[
\frac{l - 2g}{l - 1 - g} > \mu^R > -\frac{g}{l - 1 - g}.
\]

In this case, there is a value of \(\mu^R\) satisfying the constraints iff

\[
\frac{g}{1 + g} < \frac{l - 2g}{l - 1 - g} \iff l > g (1 + g).
\]

Thus, the constraints can be satisfied if \(l > \max\{g, 1\} (1 + g)\).

Putting this together, if \(g < 1\) then either \(l < 1 + g\) or \(l > \max\{g, 1\} (1 + g)\). In either case, the constraints can be satisfied, and they can also be satisfied if \(g \geq 1\) and \(l > g (1 + g)\).

\[\blacksquare\]

**OA.13 Higher-Order Information**

We analyze the efficiency properties of \textit{GrimK} when we no longer restrict the record-keeping system to only use first-order information. Here players are still reciprocators
for the first $K$ records, $0 \leq k \leq K - 1$, and defectors for all other records, $k \geq K$, but a player has record $k$ if the number of times she has played $D$ and her opponent has played $C$ is $k$, rather than if the number of times she has played $D$ in total is $k$. As we defined the function $\beta(\gamma, \varepsilon, \mu^C)$ for the analysis of $GrimK$ when records count $D$'s, so will it be useful here to define the function $\omega : (0, 1) \times (0, 1) \times [0, 1] \to (0, 1)$ given by

$$
\omega(\gamma, \varepsilon, \mu^C) = \frac{\gamma \varepsilon (1 - \varepsilon) \mu^C}{1 - \gamma(1 - \varepsilon(1 - \varepsilon)\mu^C)}.
$$

We first characterize the steady-state record shares. While $i_0 = 1 - \gamma$ remains true, now $i_{k+1} = \tau_k = \gamma \varepsilon (1 - \varepsilon) \mu^C \mu_k$ for all $0 \leq k < K - 1$, which is different than what Lemmas 4 and 5 would give. This is because a reciprocator’s record only increases when she plays $C$ and her opponent plays $D$. Equation 7 still applies and says that $\mu_0 = \frac{1 - \gamma}{1 - \gamma(1 - \varepsilon)(1 - \varepsilon)\mu^C} = 1 - \omega(\gamma, \varepsilon, \mu^C)$ and $\mu_{k+1} = \omega(\gamma, \varepsilon, \mu^C) \mu_k$ for $0 \leq k < K - 1$. Induction gives $\mu_k = \omega(\gamma, \varepsilon, \mu^C)^k (1 - \omega(\gamma, \varepsilon, \mu^C))$ for $0 \leq k \leq K - 1$. Thus,

$$
\mu^C = \sum_{k=0}^{K-1} \mu_k = \sum_{k=0}^{K-1} \omega(\gamma, \varepsilon, \mu^C)^k (1 - \omega(\gamma, \varepsilon, \mu^C)) = 1 - \omega(\gamma, \varepsilon, \mu^C)^K.
$$

We now compute the value functions. Since defectors receive a flow payoff of 0 in every period, $V_k = 0$ for all $k \geq K$. Just as before, a reciprocator receives a flow payoff of $\mu^C$. Moreover, before being matched in the current period, a reciprocator with record $k$ has a probability of $1 - \varepsilon (1 - \varepsilon) \mu^C$ of retaining her record of $k$ at the end of the period and a probability of $\varepsilon (1 - \varepsilon) \mu^C$ of her record increasing to $k + 1$. Thus,

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

for all $0 \leq k \leq K - 1$, which is equivalent to

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

$$
= (1 - \omega(\gamma, \varepsilon, \mu^C)) \mu^C + \omega(\gamma, \varepsilon, \mu^C) V_{k+1}.
$$
Recursively solving this gives

\[ V_k = (1 - \omega(\gamma, \varepsilon, \mu_C)K^{-k})\mu_C \]  \hspace{1cm} (32)

for \( 0 \leq k \leq K - 1 \).

Finally, the only incentive constraints we need worry about are the \((C|C)_k\) constraints, since a reciprocator’s record never increases when the strategy calls upon her to play \(D\). The \((C|C)_k\) constraints take the following form, which is slightly different than those in Lemma 3:

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

which is equivalent to

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

By the usual argument, the \((C|C)_0\) constraint implies all other \((C|C)_k\) constraints. Furthermore, Equation 32 shows that \((C|C)_0\) is equivalent to

\[ \frac{1 - \varepsilon(1 - \varepsilon)}{\varepsilon(1 - \varepsilon)} \omega(\gamma, \varepsilon, \mu_C)^K > g. \]

Combining this with the steady-state condition given in Equation 31 implies that \((C|C)_0\) is equivalent to

\[ \mu_C < 1 - \frac{\varepsilon(1 - \varepsilon)}{1 - \varepsilon(1 - \varepsilon)}g. \]  \hspace{1cm} (33)

Equations 31 and 33 together give the following characterization of \textit{GrimK} equilibria.

\textbf{Proposition 3.} There is a \textit{GrimK} equilibrium with total share of cooperators \(\mu_C\) if and only if the following conditions hold:

1. **Feasibility:** \(\mu_C = 1 - \omega(\gamma, \varepsilon, \mu_C)K\).

2. **Incentives:** \(\mu_C < 1 - \frac{\varepsilon(1 - \varepsilon)}{1 - \varepsilon(1 - \varepsilon)}g.\)
The following result shows that GrimK can always achieve limit efficiency in this setting, regardless of the values of \(g\) and \(l\).

**Theorem 3.** \(\lim_{\varepsilon \to 0} \lim_{\gamma \to 1} P_K^C(\gamma, \varepsilon) = 1\).

Theorem 3 follows from combining \(\lim_{\varepsilon \to 0} 1 - \varepsilon(1 - \varepsilon)/(1 - \varepsilon(1 - \varepsilon))g = 1\) with the following lemma, which is an analog of Lemma 9.

**Lemma 22.** Fix \(\varepsilon \in (0, 1)\). For all \(\Delta > 0\), there exists \(\tilde{\gamma} < 1\) such that, for all \(\gamma > \tilde{\gamma}\) 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 3 for some \(K\).

**Proof of Lemma 22.** We first state the properties of the function \(\omega(\gamma, \varepsilon, \mu^C)\) that we use in the proof.

1. For all \((\gamma, \varepsilon) \in (0, 1) \times (0, 1)\), \(\omega(\gamma, \varepsilon, \mu^C)\) is continuous and non-decreasing in \(\mu^C \in [0, 1]\).

2. For all \((\gamma, \varepsilon) \in (0, 1) \times (0, 1)\), \(\omega(\gamma, \varepsilon, \mu^C) > 0\) for all \(\mu^C > 0\).

3. For all \((\varepsilon, \mu^C) \in (0, 1) \times (0, 1]\), \(\lim_{\gamma \to 1} \omega(\gamma, \varepsilon, \mu^C) = 1\).

Note that, by the intermediate value theorem, for all \(\gamma \in (0, 1)\) and non-negative integers \(K\), there exists some \(\mu^C \in [0, 1]\) such that \(\mu^C = 1 - \omega(\gamma, \varepsilon, \mu^C)K\). Let \(\underline{\mu}^C(\gamma, \varepsilon, K)\) denote the smallest such \(\mu^C\). That is,

\[
\underline{\mu}^C(\gamma, \varepsilon, K) = \min\{\mu^C \in [0, 1] : \mu^C = 1 - \omega(\gamma, \varepsilon, \mu^C)K\}.
\]

Fix \(\Delta > 0\). There exists some \(0 < \bar{\gamma} < 1\) such that \(1 - \Delta/2 < \omega(\gamma, \varepsilon, \mu^C) < 1\) for all \(\gamma \in (\bar{\gamma}, 1)\) and \(\mu \in [\Delta/2, 1]\). For the remainder of the proof, we assume that \(\gamma \in (\bar{\gamma}, 1)\).

Since \(1 - \Delta/2 < \omega(\gamma, \varepsilon, \mu^C) < 1\) for all \(\mu \in [\Delta/2, 1]\), \(\underline{\mu}^C(\gamma, \varepsilon, K) < \Delta/2\). Moreover, because \(\lim_{K \to \infty} \overline{P}^C(\gamma, \varepsilon, K) = 1\), there exists some integer, \(K > 1\), such that \(\overline{P}^C(\gamma, \varepsilon, K) \geq 1 - \Delta/2\) and \(\overline{P}^C(\gamma, \varepsilon, K) < 1 - \Delta/2\) for all \(K < K\). Since \(1 - \Delta/2 < \omega(\gamma, \varepsilon, \mu^C) < 1\) for all \(\mu \in [\Delta/2, 1]\), it follows that both \(\underline{\mu}^C(\gamma, \varepsilon, K) < \Delta/2\).
and $\mu^C(\gamma, \varepsilon, K) < \mu^C(\gamma, \varepsilon, K + 1) < \mu^C(\gamma, \varepsilon, K) + \Delta$ hold for all $K \geq K$. These two conditions together imply that the subset $\{\mu^C(\gamma, \varepsilon, K)\}_{K \geq K}$ of $[0, 1]$ is of distance no more than $\Delta$ from any point $\mu^C \in [0, 1]$. \hfill \blacksquare