Gradualism and Rough Transition with Reference Dependence[*,†]

Takashi Onoda[‡]

Japan Bank for International Cooperation

TDB-CAREE, Graduate School of Economics, Hitotsubashi University

February 20, 2025

Abstract

I explain gradualism in cooperation through reference dependence, analyzing a dynamic game among reference-dependent, loss-averse players. The players choose cooperation levels, receive intrinsic payoffs, and update their reference points in a backward-looking manner. The subgame-perfect equilibrium that maximizes utility using Nash reversion exhibits gradualism. As cooperation begins, the players experience higher payoffs than their initial reference points, causing these points to rise, which makes the penalty for deviation more severe and facilitates further cooperation. Additionally, I examine transitional dynamics after a structural shock. Findings show that when the transition entails a loss, it creates an additional incentive to deviate, leading cooperation levels to undershoot before gradually converging to the new steady state.

Keywords: Gradualism, Starting small, Reference dependence, Loss aversion, Long-term relationship, Trade liberalization

JEL classification: C72, C73, D91

[*]I am deeply indebted to Michihiro Kandori and Ichiro Obara for their invaluable guidance and support. I also thank Rodrigo Adão, Jonathan Dingel, Taiji Furusawa, Gene Grossman, Robert Gulotty, Sota Ichiba, Fuhito Kojima, Elliot Lipnowski, Takeshi Murooka, Shunya Noda, Felix Tintelnot, Yuichi Yamamoto, Kai Hao Yang, and seminar participants at the CIRJE and UTMD workshop at the University of Tokyo, the DC conference at Osaka University, the Contract Theory Workshop, and the International Trade Working Group at the University of Chicago.

[†]An earlier version of this paper was circulated as “Gradual Development and Rough Transition of Cooperation with Reference Dependence.”

[‡]Email: onoda.takashi@gmail.com

1 Introduction

Gradualism is prevalent in cooperative relationships. Business relationships deepen gradually over time (Souchou 1987), trade credit usage increases with the duration of a relationship (McMillan and Woodruff 1999), loan sizes start small in microfinance (Morduch 1999), and countries do not eliminate tariffs immediately in a free trade agreement (Bond and Park 2002).

In this paper, I explain gradualism through reference dependence. Unlike standard models, which assume rational agents, real-world agents are reference-dependent and loss-averse. They evaluate economic returns relative to their reference points (Kahneman and Tversky 1979), with a strong aversion to losses, which means they perceive payoffs below their reference point as particularly undesirable. However, I show that the optimal level of cooperation for reference-dependent agents gradually develops over time, resembling gradualism or "starting small." Additionally, this paper illustrates how the developed cooperation responds to changes in the economic environment. Specifically, when players adjust to a lower steady-state cooperation level, they initially undershoot before stabilizing.

I investigate an infinite horizon dynamic game with complete information in which players have reference points. The players are symmetric, choosing their cooperation levels individually and simultaneously every period. Their flow utility consists of an intrinsic payoff and a loss utility. The intrinsic payoff is a function of all the players' levels of cooperation and exhibits prisoner's-dilemma-type properties; the first-best cooperation is infeasible in a one-shot game because deviation is lucrative, and a static Nash equilibrium yields low cooperation. The loss utility becomes non-zero and negative only if the intrinsic payoff is below the player's reference point, reflecting loss aversion. The reference point is backward-looking, as in DellaVigna et al. (2017) and Bowman et al. (1999). The reference point is updated every period as the weighted average of the previous value and the latest intrinsic payoff, following Bowman et al. (1999).

Throughout the paper, I consider subgame-perfect equilibria (SPEs) in which the players employ a Nash-reversion strategy to obtain analytical results. When a player deviates from a cooperative level of cooperation, the others impose a static Nash equilibrium that generates low intrinsic payoffs in the following periods as a penalty. An infinite number of SPEs exist, but I focus on the one that maximizes the players' present value of utility.

Reference dependence generates three forces that influence the development and transitions of cooperation. The first is cooperation loss aversion, the only force at work when players start cooperation with low initial reference points. Once they receive intrinsic payoffs higher than their reference points in return for their cooperation, their reference points rise. Consequently, the penalty would cause a loss, inflicting loss utilities in the deviation path. Thus, the experience of a high level of cooperation makes the players averse to losing it—an additional deterrent to deviation.

Cooperation develops gradually as the cooperation loss aversion strengthens over time. Experiencing high payoffs raises the players' reference points, generating the cooperation loss aversion. This aversion disincentives the players to deviate, enabling them to implement a higher level of cooperation in the next period. The resulting higher cooperation produces higher payoffs, elevating the reference points further. In turn, these elevated reference points widen the hypothetical losses after deviation, strengthening the cooperation loss aversion. Consequently, an even higher level of cooperation becomes possible in the following periods. Repeating this process results in a gradual development of cooperation, eventually converging to a steady state.

When the economic environment changes unfavorably, the developed cooperation faces two additional forces. When considering players whose cooperation has reached a steady state and who have reference points at that level, a structural shock that reduces intrinsic payoffs from a given level of cooperation implies a decline of the steady-state payoff. Thus, the players would incur losses from the new steady-state payoff. This situation generates two forces. First, the loss in the current period incentivizes the players to avoid the loss, augmenting the incentive to deviate. Second, if loss utilities exist in future periods, they reduce the value of future cooperative levels of cooperation, discouraging cooperation. These two effects counteract the cooperation loss aversion.

The net effect necessitates that players undergo cooperation levels lower than the new steady-state level. Thus, a reference point higher than the new steady state hinders their cooperation. Only after experiencing payoffs lower than the new steady state level and lowering their reference points can players reach the new steady state; in other words, those living in past glories are not trustworthy because they try to cling to them by betraying others; it is not until they become used to the new and less pleasant reality that they can be trustworthy. This analysis highlights that players' reputations or trustworthiness in this game are not only shaped by the history of actions, but also by the reference point, or equivalently the history of outcomes (intrinsic payoffs).

To the best of my knowledge, this paper is the first to demonstrate that reference dependence and loss aversion generate gradualism, or "starting small," in cooperation. Unlike my model, most existing works on gradualism in cooperation-building or relationship-building feature incomplete information. Previous research has analyzed situations in which players are uncertain about their partners' types (Sobel 1985; Watson 1999, 2002; Rauch and Watson 2003; Hua and Watson 2022; Furusawa and Kawakami 2008). Additionally, in some cases, players can match with other partners after a relationship ends (Ghosh and Ray 1996; Kranton 1996). [1] However, gradualism also arises in relationships where information is reasonably assumed to be complete—for example, business relationships among kin, which take years to cultivate (Souchou 1987). Research on gradualism in games with complete information focuses on other factors, such as the irreversibility of cooperation (Lockwood and Thomas 2002).

The results of this work apply also to gradualism in other fields, such as trade policy. The literature on trade policy analyzes trade liberalization with incomplete tariff elimination as an equilibrium result of a tariff-setting game between strategic governments in an infinite time horizon. The welfare function that the governments maximize falls into a class of intrinsic payoff functions described in my model.[2] While there have been various explanations of gradualism in trade liberalization (Staiger 1994; Devereux 1997; Furusawa and Lai 1999; Bond and Park 2002; Maggi and Rodriguez-Clare 2007; Zissimos 2007; Chan 2019), research on reference dependence as the driver in a strategic setting is lacking. Although reference dependence or loss aversion at an aggregate level is not as established a concept as

[1] Datta (1996) also studies gradualism in a model with random matching but with complete information.

[2] The welfare for a government to maximize is the sum of consumer surplus, producer surplus, and tariff revenues. When a government redistributes an increase in this welfare among its citizens, the welfare is the sum of the gains from tariff reductions for individual citizens. If citizens are reference-dependent and loss-averse, their individual loss utilities aggregate to a loss utility at the government level.

at the individual level, efforts have been made to theoretically and empirically investigate behavioral effects in trade policy (Bernardes 2003; Freund and Özden 2008; Tovar 2009).

The implication for transitional dynamics is a unique contribution. From collusion by firms to trade agreements to gang memberships, no cooperation is immune to changes in their economic environments, such as new whistleblower regulations, growing trade protectionism, or stricter law enforcement. However, analysis of the transitional dynamics is limited in the literature. A mature relationship has resolved the problem of incomplete information; therefore, models of cooperation development with incomplete information typically imply an instant transition from a steady state to a new one. My model provides new implications for transitional dynamics of mature relationships and cooperation.

The rest of this paper is structured as follows. Section 2 introduces the model. Section 3 examines Nash-reversion strategies in this model. Section 4 identifies the three effects from reference dependence. Section 5 solves for the SPE that maximizes utility with a Nash-reversion strategy, illustrating the gradual development of cooperation. Section 6 analyzes how the cooperation level transitions to a new steady state when the economic environment changes. Section 7 concludes this paper.

2 Setup

I study a dynamic game with complete information by N symmetric players. I denote the set of players by $\boldsymbol{N}$. There are infinite periods, which I denote by t. In each period, the players individually and simultaneously choose their cooperation levels. I denote player i's cooperation level in period t by $\alpha_{i,t} \in A$ where $A = [0, \bar{\alpha}]$ and the vector of cooperation levels in period t by $\boldsymbol{\alpha}_t = (\alpha_{1t}, \alpha_{2t}, \ldots, \alpha_{Nt})$.

Player i's flow utility consists of an intrinsic payoff $\pi_i(\boldsymbol{\alpha}_t)$ and a loss utility $\ell_i(\boldsymbol{\alpha}_t|r_{it})$. Given the reference point r_{it} of player i in period t, it is given by

$$u_i(\boldsymbol{\alpha}_t|r_{i,t}) = \pi_i(\boldsymbol{\alpha}_t) + \ell_i(\boldsymbol{\alpha}_t|r_{it}),$$

where

$$\ell_i(\boldsymbol{\alpha}_t|r_{it}) = \begin{cases} 0 & \text{if } \pi(\boldsymbol{\alpha}_t) \geq r_{i,t}, \\ -\eta\,[r_{it} - \pi_i(\boldsymbol{\alpha}_t)] & \text{if } \pi(\boldsymbol{\alpha}_t) < r_{i,t}, \end{cases}$$

and $0 < \eta \leq 1$. The intrinsic payoff is symmetric among the players in the sense that $\boldsymbol{\alpha}'_t = (\alpha'_{1t}, \alpha'_{2t}, \ldots, \alpha'_{Nt})$ with $\alpha_{it} = \alpha'_{jt}$, $\alpha_{jt} = \alpha'_{it}$, and $\alpha_{kt} = \alpha'_{kt}$ for else implies $\pi_i(\boldsymbol{\alpha}_t) = \pi_j(\boldsymbol{\alpha}'_t)$. The loss utility term $\ell_i(\boldsymbol{\alpha}_t|r_{it})$ captures the reference dependence and loss aversion. It measures losses in terms of intrinsic payoffs as in Kőszegi and Rabin (2006); when the intrinsic payoff $\pi_i(\boldsymbol{\alpha}_t)$ is below the reference point r_{it}, they perceive it as a loss, incurring the negative loss utility. The parameter $\eta > 0$ measures the degree of loss aversion, and $\eta = 0$ would eliminate reference dependence, making the model standard.

This functional form reflects simplifying assumptions. First, the flow utility u comprises only the loss utility, not both loss and gain utilities.[3] The lack of gain utilities prevents us from weighining the importance of gain-loss utilities relative to intrinsic payoff separately from loss utility relative to gain utility. However, it retains the property of loss aversion and simplifies the algebra. Moreover, in structural estimations of a reference-dependent job search model, DellaVigna et al. (2017) report that a model only with loss utility fits the data similarly to a model with both gain and loss utilities. Thus, using only the loss utility for tractability is a reasonable choice. Second, it abstracts from the diminishing sensitivity of the prospect theory (Kahneman and Tversky 1979); the loss utility is piecewise linear in loss ($r_{it} - \pi_i(\boldsymbol{\alpha}_t)$). This functional form enables us to study the main idea of this paper in a relatively simple model. The parameter η belongs to $(0, 1]$, which is roughly consistent with the observation that people evaluate an economic outcome to be twice as large when they perceive it as a loss compared with when they perceive it as a gain in terms of absolute value in riskless decisions (Kahneman et al. 1990, Tversky and Kahneman 1991).

I assume several properties of the intrinsic payoff $\pi_i(\boldsymbol{\alpha}_t)$ that can be micro-founded by either collusion in Cournot competition or trade liberalization as explained in the Online Appendix. $\pi_i(\boldsymbol{\alpha}_t)$ is continuous in α_j and twice differentiable with respect to α_j for all $j \in N$. In this paper, I focus on equilibria where players symmetrically choose the same level of cooperation. To simplify the notations, I define $\tilde{\pi}(\alpha)$ as Definition 1 and assume $\tilde{\pi}$ is

[3] Other works that employ this formulation includes Freund and Özden (2008).

increasing as Assumption 1.

Definition 1 (Symmetric Cooperation). $\tilde{\pi}(\alpha) = \pi_i(\boldsymbol{\alpha}_t)$ where $\alpha_{it} = \alpha$ for all $i \in N$.

Assumption 1. $\tilde{\pi}(\alpha)$ strictly increases in α.

I let $\bar{\pi} = \max_{\alpha \in A} \tilde{\pi}(\alpha) = \tilde{\pi}(\bar{\alpha})$ and refer to $\bar{\alpha}$ as the first-best cooperation. Assumption 1 ensures that the inverse function $\tilde{\pi}^{-1}(\pi)$ exists. Using this inverse function, I define the best-response intrinsic payoff b as a function of π.[4]

Definition 2 (Best-Response). $b(\pi) = \max_{\alpha_i \in A} \pi_i(\alpha_i, \tilde{\pi}^{-1}(\pi), \ldots, \tilde{\pi}^{-1}(\pi))$.

I assume two properties for the best-response intrinsic payoff as Assumption 2.

Assumption 2. b is convex, and there exists unique $\underline{\pi} \in [\tilde{\pi}(0), \bar{\pi})$ that solves $b(\underline{\pi}) = \underline{\pi}$.

Assumption 2 implies that $\underline{\pi}$ is the intrinsic payoff in a unique symmetric pure-strategy Nash equilibrium in a one-shot game without a loss utility ($\eta = 0$). I denote the cooperation level that produces $\underline{\pi}$ by α^N, referring to α^N as the Nash cooperation in the rest of the paper, although whether $\underline{\pi}$ is the intrinsic payoff in a Nash equilibrium with a loss utility ($\eta > 0$) is unclear at this point. Section 3 shows that $\underline{\pi}$ is indeed the intrinsic payoff in a SPE even with a loss utility ($\eta > 0$).

The reference point is backward-looking and adaptive, following Bowman et al. (1999) and DellaVigna et al. (2017). Specifically, it is given by

$$r_{i,t} = \begin{cases} \rho r_{i,t-1} + (1-\rho)\pi_i(\boldsymbol{\alpha}_{t-1}) & \text{if } t \geq 2, \\ r_1 & \text{if } t = 1, \end{cases} \tag{1}$$

where $0 < \rho < 1$. The initial reference point $r_1 \in \mathbb{R}$ takes an exogenously given common value among the players, which is public information. The parameter ρ measures the persistence of a reference point.

This backward-looking reference point is a crucial assumption in my model. If it were forward-looking, deviation from a given sequence of intrinsic payoffs would always produce

[4] $b(\pi)$ is the best response in terms of intrinsic payoff and in terms of flow utility.

the same size of losses in subsequent periods. Thus, the optimal level of cooperation would be invariant to the history of payoffs, becoming constant over periods. [5] In the structural estimations by DellaVigna et al. (2017), the backward-looking and adaptive reference point outperforms other types of reference points, including the forward-looking one (Kőszegi and Rabin 2006) and the status quo, which corresponds to $\rho = 0$ in my model, in terms of fitting job search patterns.[6][7] Also, Post et al. (2008) report the persistence of reference points, although they use a different functional form than mine. Thus, a backward-looking and adaptive reference point is a reasonable assumption.

The players correctly understand how their own reference points evolve. Given a path of intrinsic payoffs, they derive their future reference points according to the law of motion (1), evaluating future intrinsic payoffs against them. They discount future flow utilities by a discount factor $\delta \in (0, 1)$, maximizing the present value of their utility $\sum_{t=0}^{\infty} \delta^t u(\alpha_t | r_{i,t})$. The Online Appendix presents a study of the game in which players incorrectly believe their reference points remain fixed at the current level indefinitely. Gradualism and rough transitions still occur, demonstrating the robustness of the model's results.

3 Nash-Reversion Strategy

The analysis in this paper focuses on the pure-strategy SPEs with Nash-reversion strategies, which use a static Nash equilibrium as the penalty (Friedman 1971). Consequently, the range of cooperation levels feasible in these SPEs is generally limited compared with the case where the optimal penalty is applied (Abreu 1988). However, regardless of the specific form of the penalty, it imposes low payoffs, generating loss utilities. The magnitude of these loss utilities increases with the reference point, meaning that the core mechanism

[5] I predict that completely eliminating the effect from expected utilities is unnecessary. As long as the payoff history affects the reference point, gradualism should appear for the same reasoning.

[6] DellaVigna et al. (2017) use the average income over the N previous periods in their benchmark model instead of the AR(1) formation like mine for computational reasons. They report that the fits of these two reference points are similar.

[7] In addition to the empirical support, there is another reason why I do not allow ρ to be one. $\rho = 1$ disconnects a reference point before deviation and the loss utilities by a penalty. If $\rho = 1$, the reference point in the first penalty period is the intrinsic utility by the best deviation, which I show independent of the reference point in Section 3. Given this, the magnitude of the cooperation loss aversion would become independent of the reference point, and gradualism does not emerge.

driving gradualism remains the same across penalty structures. While the optimal penalty theoretically expands the range of feasible cooperation, its implementation is extraordinarily complex, making its equilibrium path highly intricate and difficult to solve.[8] Given that even a simple penalty already leads to a challenging equilibrium, as we will see in this paper, adopting a more tractable penalty structure is a reasonable choice. Owing to this choice, this study yields rich analytical results, as a penalty that is independent of reference points forms an SPE, as analyzed later in this section.

This section examines the Nash-reversion strategies and the best deviation when the others employ a Nash-reversion strategy. To do so, we must first determine where to revert. A natural candidate is a Nash strategy s^N in Definition 3.

Definition 3 (Nash Strategy). A Nash strategy s^N is a sequence of mapping σ_t^N that maps any history of cooperations $\{\boldsymbol{\alpha}_k\}_{k=1}^{t-1} \in A^{N(t-1)}$ and the initial reference point $r_1 \in \mathbb{R}$ into the Nash cooperation α^N. That is, $\forall t \in \mathbb{N}$, $\sigma_t^N : A^{N(t-1)} \times \mathbb{R} \mapsto \alpha^N$.

We need to examine if all the players employing s^N, $(s^N, s^N, \ldots s^N)$, indeed forms a SPE in this game. An SPE may not be formed for two reasons. First, the flow utility $u(\boldsymbol{\alpha}_t|r_{i,t})$ additionally contains the loss-aversion term $\ell(\boldsymbol{\alpha}_t|r_{i,t})$. Second, this model has a dynamic effect. An action in a given period affects the reference points in the following periods. Thus, the maximizer of the flow utility u possibly differs from that of the present value of utility $(\sum_{k=t}^{\infty} \delta^{k-t} u_{i,k})$, even when the other players choose the same actions every period. To see these effects clearly, I lay out the derivative of the utility with respect to $\pi_i(\boldsymbol{\alpha}_{t+k})$ below:

$$\delta^{t+k}\Big(1 + \mathbb{1}\{\pi_i(\boldsymbol{\alpha}_{t+k}) < r_{i,t+k}\} \cdot \eta - \sum_{m=1}^{\infty} \delta^m \mathbb{1}\{\pi_i(\boldsymbol{\alpha}_{t+k+m}) < r_{i,t+k+m}\} \cdot \eta \cdot \frac{\partial r_{i,t+k+m}}{\partial \pi_i(\boldsymbol{\alpha}_{t+k})}\Big).$$

This derivative shows the standard positive effect $(\delta^{t+k} \cdot 1)$ and additional positive and negative effects. First, the second term is the positive effect: a reduction of the current loss utility

[8] The optimal penalty generally depends on a reference point. If it follows a stick-and-carrot structure (Abreu 1988), the stick phase entails losses, reinforcing the penalty and creating an incentive to deviate from the punishment. Analogous to the loss-evading incentive discussed in Section 4, the magnitude of this incentive increases with the reference point in that period, but only when deviation from the punishment eliminates the loss. Additionally, the reference point, which is further influenced by the payoff in the stick phase, affects the value of the carrot phase, as the feasible set of cooperation levels depends on the reference point.

when the intrinsic payoff $\pi_i(\boldsymbol{\alpha}_{t+k})$ is in the loss region. Second, the third term summarizes the indirect negative effect: a higher intrinsic payoff today $\pi_i(\boldsymbol{\alpha}_{t+k})$ raises future reference points ($\partial r_{i,t+k+m}/\partial \pi_i(\boldsymbol{\alpha}_{t+k}) > 0$), widening future losses, if they exist. One can prove that the standard positive effect dominates the indirect negative effect, making the net effect positive. Thus, Lemma 1 follows.

Lemma 1. $(s^N, s^N, \ldots s^N)$ is a SPE.

Proof. See Appendix A. □

Given this result, I formally define a Nash-reversion strategy—a grim trigger strategy with the Nash penalty s^N—in Definition 4.

Definition 4. A Nash-reversion strategy with a cooperation path $\{\alpha_t^c\}_{t=1}^{\infty}$ for player i, $s(\{\alpha_t^c\}_{t=1}^{\infty})$, is a sequence of mapping $\{s_k(\{\alpha_t^c\}_{t=1}^{k})\}_{k=1}^{\infty}$ such that

$$s_k(\{\alpha_t^c\}_{t=1}^{k}) : A^{N(k-1)} \times \mathbb{R} \mapsto \begin{cases} \alpha_k^c & \text{if } \forall t \in \{1, 2, \ldots, k-1\},\ \forall \mathrm{j} \in \boldsymbol{N}, \alpha_{jt} = \alpha_t^c \ \text{ or if } k = 1 \quad \text{and} \\ \alpha^N & \text{otherwise.} \end{cases}$$

As the next step, I identify the best deviation. Given that the other players employ s^N from the next period, the best deviation maximizes the present value of utility. Like the analysis of the Nash strategy, the marginal utility with respect to the own cooperation level has, possibly, two additional effects: directly alleviating the current loss utility and indirectly aggravating the future loss utilities by raising future reference points. Nevertheless, the best deviation in my model becomes the same as the best deviation in the model without the reference dependence ($\eta = 0$) because the standard positive effect on the current intrinsic payoff dominates the indirect effect of aggravating the future loss utilities. Formally, Lemma 2 states this result.

Lemma 2. When the others play a Nash-reversion strategy $s(\{\alpha_t^c\}_{t=1}^{\infty})$, the best deviation in period t (α_t^b) is unique and the same as that without the reference dependence ($\eta = 0$). That is, for any $\boldsymbol{i} \in \boldsymbol{N}$, $\forall r_{i,t} \in \mathbb{R}$,

$$\underset{\alpha_{it}\in A}{\operatorname{argmax}}\left(\pi_i(\alpha_{it}, \boldsymbol{\alpha}^c_{-i,t}) + \ell(\alpha_{it}, \boldsymbol{\alpha}^c_{-i,t}|r_{i,t}) + \sum_{k=1}^{\infty}\delta^k\left[\tilde{\pi}(\alpha^N) + \ell(\boldsymbol{\alpha}^N|r_{i,t+k})\right]\right) = \underset{\alpha_{it}\in A}{\operatorname{argmax}}\,\pi_i(\alpha_{it}, \boldsymbol{\alpha}^c_{-i,t}),$$

where $\boldsymbol{\alpha}^c_{-i,t}$ is the vector of cooperation levels by all the players but $\boldsymbol{i}$ where $\alpha_{jt} = \alpha^c_t$ for all $j \neq \boldsymbol{i}$.

Proof. See Appendix B. □

Having validated the penalty and identified the best deviation, I analyze the effects of reference dependence in SPEs with Nash-reversion strategies in the next section.

4 Three Effects

This section identifies the three effects that reference dependence generates in SPEs with Nash-reverting strategies. In the rest of the paper, I focus on cooperation paths that take values between the Nash cooperation and the first-best cooperation, $[\alpha^N, \bar{\alpha}]$. [9] Given this interval, the focus on the symmetric equilibria with Nash-reversion strategies, and Lemmas 1 and 2, I simplify notations using π_t and $b(\pi_t)$ where $\pi_t = \tilde{\pi}(\alpha_t)$, as if the players directly choose π_t instead of $\boldsymbol{\alpha}_t$, and a player earns $b(\pi_t)$ by the best deviation. The notation for a Nash-reversion strategy changes accordingly from $s(\{\alpha^c_t\}_{t=1}^{\infty})$ to $s(\{\pi^c_t\}_{t=1}^{\infty})$. Since $\tilde{\pi}$ is increasing on $[\alpha^N, \bar{\alpha}]$, I sometimes refer to the level of π_t as that of cooperation and, in particular, $\bar{\pi}$ as the first-best cooperation. A cooperation level greater than α^N and an intrinsic payoff greater than $\underline{\pi}$ are called cooperative. I assume the initial reference point falls into the same interval as intrinsic payoffs as Assumption 3.

Assumption 3. $\underline{\pi} \leq r_1 \leq \bar{\pi}$.

In this game with the initial reference point r_1, $(s(\{\pi^c_t\}_{t=1}^{\infty}), s(\{\pi^c_t\}_{t=1}^{\infty}), \ldots, s(\{\pi^c_t\}_{t=1}^{\infty}))$ forms a SPE if and only if, for all $t \in \mathbb{N}$, every player weakly prefers following $s(\{\pi^c_t\}_{t=1}^{\infty})$ to the best deviation. I call a cooperation path $\{\pi^c_t\}_{t=1}^{\infty}$ feasible if and only if $(s(\{\pi^c_t\}_{t=1}^{\infty}), s(\{\pi^c_t\}_{t=1}^{\infty}), \ldots, s(\{\pi^c_t\}_{t=1}^{\infty}))$ forms a SPE. Lemma 3 states the condition for feasibility.

[9] This focus does not forbid a best deviation from taking a value lower than α^N.

Lemma 3 (Feasibility). $\{\pi_t^c\}_{t=1}^{\infty}$ is feasible if and only if,

$$\begin{aligned}
\forall t \in \mathbb{N}, \quad & b(\pi_t^c) - \pi_t^c + \mathbb{1}\{r_t > \pi_t^c\} \cdot \eta \left[\min\{r_t, b(\pi_t^c)\} - \pi_t^c\right] \\
& \leq \sum_{k=1}^{\infty} \delta^k \Big[\pi_{t+k}^c - \mathbb{1}\{r_{t+k} > \pi_{t+k}^c\} \cdot \eta \left[r_{t+k} - \pi_{t+k}^c\right] \\
& \qquad - \left\{\underline{\pi} - \mathbb{1}\{r_{t+k}^d > \underline{\pi}\} \cdot \eta \left[r_{t+k}^d - \underline{\pi}\right]\right\}\Big], \qquad (2) \\
s.t. \forall k \in \mathbb{N}, \; & r_{t+k} = \rho r_{t+k-1} + (1-\rho)\pi_{t+k}^c, \\
\forall k \in \mathbb{N}, \; & r_{t+k}^d = \rho r_{t+k-1}^d + (1-\rho)\underline{\pi} \quad \text{where } r_t^d = r_t.
\end{aligned}$$

Proof. It follows from Lemmas 1 and 2. □

This condition inherits, from a standard model, a tension between the short-term gain and the long-term loss by deviation. The left-hand side of the inequality expresses the short-term gain from deviation in period t. It consists of the improvement of the current intrinsic payoff ($b(\pi_t^c) - \pi_t^c$) and the reduction of the loss utility, if any. The right-hand side is the difference in the present value of utilities from the following periods between the equilibrium and deviation paths, which I refer to as a long-term loss. The value of future utilities from the cooperative levels of cooperation is measured with reference points in the equilibrium path $\{r_{t+k}\}_{k=1}^{\infty}$, and that of deviation is measured with reference points in the deviation path $\{r_{t+k}^d\}_{k=1}^{\infty}$.

To analyze the standard effects and the effects from reference dependence separately, I suppose that reference dependence does not exist ($\eta = 0$) and that the cooperation path is constant ($\pi_t^c = \pi^c$). Inequality (2) becomes

$$b(\pi^c) - \pi^c \leq \sum_{k=1}^{\infty} \delta^k \left[\pi^c - \underline{\pi}\right].$$

On the one hand, the short-term gain is convex and possibly becomes very large when the cooperative intrinsic payoff π^c is sufficiently high. On the other hand, the long-term loss is linear in π^c. Thus, rising π^c makes the short-term gain overwhelm the long-term loss at some point, subject to parameter values, making high cooperation levels infeasible.

Reference dependence adds three effects to this tension. To see them clearly, I express the future reference points r_{t+k} and r^d_{t+k} as functions of the current reference point r_t and future intrinsic payoffs. Manipulating inequality (2) yields the following inequality, showing the three effects as:

$$
\begin{aligned}
&b(\pi^c_t) - \pi^c_t + \mathbb{1}\{r_t > \pi^c_t\} \cdot \underbrace{\eta \left[\min\{r_t, b(\pi^c_t)\} - \pi^c_t\right]}_{\text{Loss-evading incentive}} \\
&\leq \sum_{k=1}^{\infty} \delta^k \pi^c_{t+k} - \underbrace{\sum_{k=1}^{\infty} \mathbb{1}\{r_{t+k} > \pi^c_{t+k}\} \cdot \delta^k \eta \left[\rho^k r_t + \sum_{\ell=0}^{k-1} \rho^{k-1-\ell}(1-\rho)\pi^c_{t+\ell} - \pi^c_{t+k}\right]}_{\text{Future cooperation losses}} \\
&\quad - \sum_{k=1}^{\infty} \delta^k \underline{\pi} + \underbrace{\sum_{k=1}^{\infty} \delta^k \eta \left[\rho^k r_t + \rho^{k-1}(1-\rho) b(\pi^c_t) - \rho^{k-1}\underline{\pi}\right]}_{\text{Cooperation loss aversion}}. \qquad (3)
\end{aligned}
$$

First, the cooperation loss aversion adds to the long-term loss, and a higher reference point r_t strengthens this effect. The last summation term on the right-hand side is non-negative and increases with r_t, capturing this effect. After deviation, a player incurs a loss utility every period. A higher reference point in the current period has a persistent effect on future reference points as $\rho > 0$, aggravating these loss utilities. Intuitively, the more players are used to high payoffs, the more they hesitate to deviate due to future losses.

The second effect is the loss-evading incentive; loss from π^c_t incentivizes players to deviate, captured by the third term on the left-hand side. Given a cooperation level π^c_t, the magnitude of this effect weakly increases with r_t. The larger a loss is, the more a player can potentially gain by deviating. However, a reference point higher than $b(\pi^c_t)$ creates too big a loss to eliminate, making the magnitude of this effect unresponsive to a further increase.

Third, the equilibrium path can entail losses in the future ($r_{t+k} > \pi^c_{t+k}$), labeled as future cooperation losses, which discourage cooperation in the current period. The second summation term on the right-hand side captures this effect, reducing the long-term loss by deviation. Given a cooperation path $\{\pi^c_t\}^{\infty}_{t=1}$, the size of this effect increases with r_t; a higher reference point implies more significant future losses because of its persistence.

Whether a higher reference point r_t relaxes or tightens this inequality condition depends on whether the equilibrium path entails losses. Without a loss in any period, only the cooperation loss aversion is operative; consequently, a higher r_t relaxes the condition. In contrast, if there is a loss every period ($r_{t+k} > \pi^c_{t+k}$ for all k), all three effects exist, and each of them is strengthened with a higher r_t. Eq. (3) shows that the contributions of r_t to the cooperation loss aversion and the future cooperation losses are both $\sum_{k=1}^{\infty} \delta^k \eta \rho^k r_t$, canceling each other out. Thus, the only remaining effect is the strengthening loss-evading incentive, tightening the condition. This case-dependent net effect becomes crucial in the optimized path, which I study in Section 5.

I exclude an uninteresting case in which the first-best cooperation $\bar{\pi}$ is feasible from period 1 regardless of the initial reference point. Imposing Assumption 4 ensures it.

Assumption 4. The first-best cooperation every period $\{\bar{\pi}\}_{t=1}^{\infty}$ is not feasible when the reference point is at the Nash cooperation level ($r_1 = \alpha^N$) as

$$b(\bar{\pi}) - \bar{\pi} > \frac{\delta}{1-\delta}\left[\bar{\pi} - \underline{\pi}\right] + \frac{\delta\eta}{1-\delta\rho}\left[\rho\underline{\pi} + (1-\rho)b(\bar{\pi}) - \underline{\pi}\right].$$

The less patient (low δ) and the less loss-averse (low η) the players are, and the less persistent a reference point is (low ρ), the more likely Assumption 4 holds. A smaller discount factor lowers the present value of the long-term loss. Less persistence of a reference point and less loss-aversion weaken the cooperation loss aversion. Another factor that can support Assumption 4 is a lucrative best deviation. The higher $b(\bar{\pi})$ is relative to $\bar{\pi}$, the more likely the inequality holds. Given Assumption 4, the first-best cooperation $\bar{\pi}$ is not feasible for some r_1, allowing us to study the development path of cooperation in the next section.

5 Gradual Development

In this section, I investigate the optimal SPE, defined as one that maximizes the present value of utility among those feasible by a Nash-reversion strategy, showing that cooperation optimally develops gradually. The path of π_t in that equilibrium is the solution to the following problem:

$$
\begin{aligned}
v^*(r_1) &= \max_{\{\pi_t\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \delta^{t-1} \left[\pi_t - \mathbb{1}\{r_t > \pi_t\} \cdot \eta(r_t - \pi_t)\right], \\
s.t. \forall t \in \mathbb{N},\ & \pi_t \in \Gamma(r_t, v^*), \\
\forall t \in \mathbb{N},\ & r_{t+1} = \rho r_t + (1-\rho)\pi_t,
\end{aligned} \tag{4}
$$

where the feasible set $\Gamma(r_t, v^*)$ is defined as

$$
\begin{aligned}
\Gamma(r_t, v^*) = \Bigg\{ \pi_t \in [\underline{\pi}, \bar{\pi}] : & \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(\pi_t) - \mathbb{1}\{r_t > b(\pi_t)\} \cdot \eta\left[r_t - b(\pi_t)\right] \\
& + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right] \underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} r_t \le v^*(r_t) \Bigg\}.
\end{aligned}
$$

The first constraint ensures the feasibility following Lemma 3; the inequality compares the value of the deviation path on the left-hand side and that of the equilibrium path on the right-hand side. To characterize the solution, I set up the corresponding functional equation to apply dynamic programming. It takes r_t as the state variable and π_t as the control variable. I obtain Lemma 4.

Lemma 4. The function v^* satisfies the following functional equation.

$$
v^*(r) = \max_{y \in \Gamma(r, v^*)} \left[y - \mathbb{1}\{r > y\} \cdot \eta(r - y) + \delta v^*(\rho r + [1-\rho] y)\right]. \tag{5}
$$

Proof. See Appendix C. □

The functional equation (5) differs from a standard dynamic programming problem in that the correspondence $\Gamma(r, v^*)$ has the function v^* as an argument. Nevertheless, Lemma 4 states that the function v^* satisfies eq. (5), enabling us to characterize the value function v^* and the policy function $g^*(r) = \text{argmax}_{y \in \Gamma(r,v^*)}\, y - \mathbb{1}\{r > y\} \cdot \eta(r - y) + \delta v^*(\rho r + [1-\rho] y)$.

The properties of the value function v^* and the policy function g^* depend on whether the first-best cooperation $\bar{\pi}$ is sustainable. I define sustainability in Definition 5.

Definition 5 (Sustainability). A cooperation (or reference-point) level π is sustainable if the constant cooperation path at that level ($\{\pi_t^c\}_{t=1}^{\infty}$ where $\pi_t^c = \pi$) is feasible with a Nash-reversion strategy when $r_1 = \pi$. Otherwise, they are unsustainable.

The constant cooperation path at the initial reference point level does not change the reference point over periods, generating no loss. Thus, the condition for π to be sustainable follows from Lemma 3 as

$$b(\pi) - \pi \leq \frac{\delta}{1-\delta}\left[\pi - \underline{\pi}\right] + \frac{\delta\eta}{1-\delta\rho}\left[\rho\pi + (1-\rho)b(\pi) - \underline{\pi}\right]. \tag{6}$$

Given this definition, I partition all the cases into two.

Case 1. The first-best cooperation is sustainable.

Case 2. The first-best cooperation is unsustainable.

Similar to the condition of Assumption 4, given $\pi = \bar{\pi}$, the right-hand side of the inequality (6) increases in δ, η, and ρ; the more patient (higher δ) and the more loss-averse (higher η) the players are, the more likely it falls into Case 1. Also, the more persistent a reference point is (higher ρ), the more likely it is Case 1. These increase the magnitude of the long-term loss. On the other hand, a more lucrative best deviation (higher $b(\bar{\pi})$) makes Case 2 more likely by raising the short-term gain. In the rest of this section, I first provide analytical results for Case 1. Subsequently, I provide analytical results for Case 2. Finally, I provide numerical examples of Case 2, obtaining additional insights.

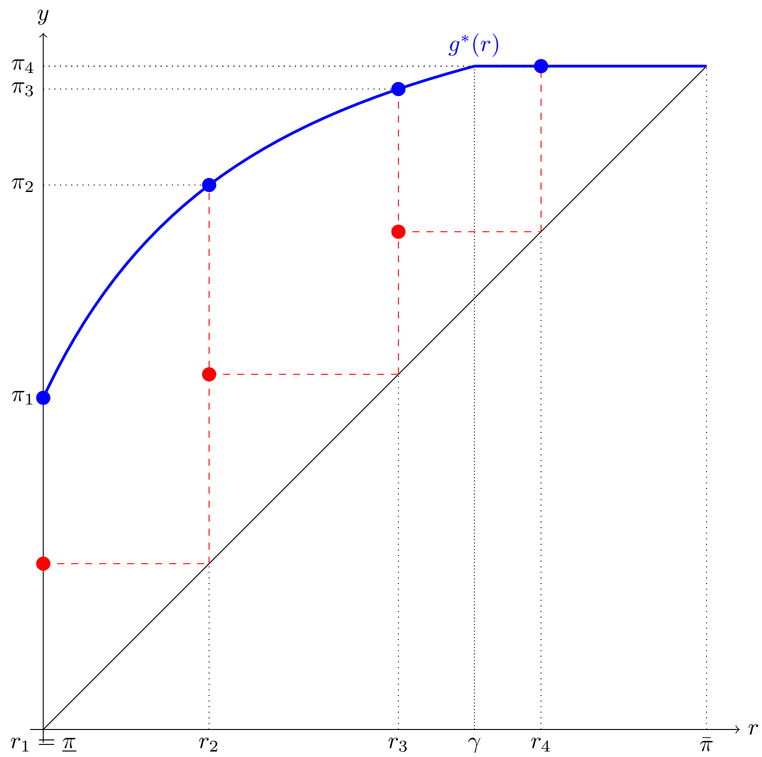
5.1 Case 1

For Case 1, Lemma 5 summarizes the main properties of the policy function g^*.

Lemma 5 (Case 1). The policy function g^* is continuous on $[\underline{\pi}, \bar{\pi}]$ and strictly increasing and concave on $[\underline{\pi}, \gamma]$. Additionally, $r < g^*(r) < \bar{\pi}$ on $[\underline{\pi}, \gamma)$, and $g^*(r) = \bar{\pi}$ on $[\gamma, \bar{\pi}]$. Consequently, $\bar{\pi}$ is the unique steady state.

Proof. See Appendix D. □

Figure 1
Gradual Development Path in Case 1



Notes: The solid blue line depicts the policy function g^* according to the properties in Proposition 5 for Case 1. The blue dots locate $\{\pi_1, \pi_2, \pi_3, \pi_4\}$ and the red dots locate $\{r_2, r_3, r_4\}$ of the optimized cooperation path with $r_1 = \underline{\pi}$.

Figure 1 illustrates the optimized cooperation path for players with the initial reference point at $\underline{\pi}$ according to these properties. The policy function $g^*(r_t)$ locates an intrinsic payoffs π_t (blue dots). Subsequently, π_t determines the reference point in the next period r_{t+1} (red dots) as the weighted average of π_t and r_t. From the beginning, the optimized cooperation $g^*(r_t)$ is higher than r_t, raising the reference point. Consequently, the next optimized cooperation becomes even higher. This way, the cooperation level gradually and monotonically rises and reaches the first-best level $\bar{\pi}$ when the reference point reaches or exceeds γ. Thus, I obtain Proposition 1.

Proposition 1 (Cooperation Path in Case 1). (i) In a game of players with initial reference points lower than γ, they monotonically increase the level of cooperation over time, with the level converging to the first-best level. (ii) In a game of players with initial reference points not lower than γ, they keep implementing the first-best cooperation level from period 1.

Proof. It follows from Lemma 5. ◻

The driver of this monotonically increasing cooperation is the monotonically strengthening cooperation loss aversion ($\delta\eta\rho r_t/(1-\delta\rho)$ in $\Gamma(r_t, v^*)$). Experiencing a higher utility raises the reference point and aggravates the loss utilities in the deviation path. Consequently, the aversion strengthens, enabling the players to shift up the cooperation level. In turn, this even higher cooperation level elevates the reference point, reinforcing the aversion further. The repetition of this process continues until the cooperation level reaches the first-best level. This cooperation path never generates a loss, ensuring the other two effects remain absent.

5.2 Case 2

In Case 2, reference point r can be unsustainable, allowing the other two effects to operate. Lemma 6 states the properties of g^*, describing how the path differs depending on the sustainability of r.

Lemma 6 (Case 2). The policy function g^* is continuous, strictly increasing, and strictly concave on $[\underline{\pi}, \bar{z}]$ where $\bar{z}$ is the highest sustainable cooperation level that is implicitly defined by

$$b(\bar{z}) - \bar{z} = \frac{\delta}{1-\delta}\left[\bar{z} - \underline{\pi}\right] + \frac{\delta\eta}{1-\delta\rho}\left[\rho\bar{z} + (1-\rho)b(\bar{z}) - \underline{\pi}\right]. \tag{7}$$

Additionally, $r < g^*(r) < \bar{z}$ for $r < \bar{z}$, $g^*(\bar{z}) = \bar{z}$, and $g^*(r) < \bar{z}$ for $r > \bar{z}$. Consequently, $\bar{z}$ is the unique steady state.

Proof. See Appendix E. ◻

Lemma 6 immediately implies Proposition 2.

Proposition 2 (Cooperation Path in Case 2). (i) In a game of players with initial reference points lower than $\bar{z}$, the players monotonically increase the level of cooperation over time, with the level converging to $\bar{z}$. (ii) In a game of players with initial reference points higher than $\bar{z}$, they start cooperation below $\bar{z}$, with the level eventually reaching $\bar{z}$.

Proof. It follows from Lemma 6. ◻

When players' initial reference point r_1 is not as high as $\bar{z}$—the case we expect for a new relationship—the development of cooperation exhibits gradualism. The properties of the policy function on $[\underline{\pi}, \bar{z}]$ are the same as those on $[\underline{\pi}, \gamma]$ in Case 1. Heuristically, this result is due to the non-existence of a loss from any reference point $r_1 \in [\underline{\pi}, \bar{z}]$. Thus, only the cooperation loss aversion is operative, driving gradually developing cooperation similarly to Case 1.

An excessively high reference point hinders cooperation. In Case 2, a high $b(r)$ prevents the players from implementing a cooperation level above $\bar{z}$, making a reference point higher than $\bar{z}$ unsustainable. Given such a reference point, the players must follow an equilibrium path with losses. The future cooperation losses cancel out the cooperation loss aversion in incentivizing cooperation, as discussed in Section 4. Additionally, the loss-evading incentive—the only remaining effect—operates, preventing the players from maintaining the cooperation level even at $\bar{z}$.

Whether a higher reference point in the region above $\bar{z}$ lowers the cooperation level is generally unclear because the magnitude of the loss-evading incentive can be unresponsive to r, as discussed in Section 4. Nevertheless, Assumption 5 enables us to obtain the monotonicity as Proposition 3.

Assumption 5. The highest sustainable $\bar{z}$ is sufficiently close to $\bar{\pi}$.

Proposition 3 (Case 2 with Additional Assumption). Given Assumption 5, g^* is strictly decreasing on $[\bar{z}, \bar{\pi}]$. Thus, the higher the players' initial reference points are, the lower the cooperation level is in the first period, as long as their initial reference points are above $\bar{z}$.

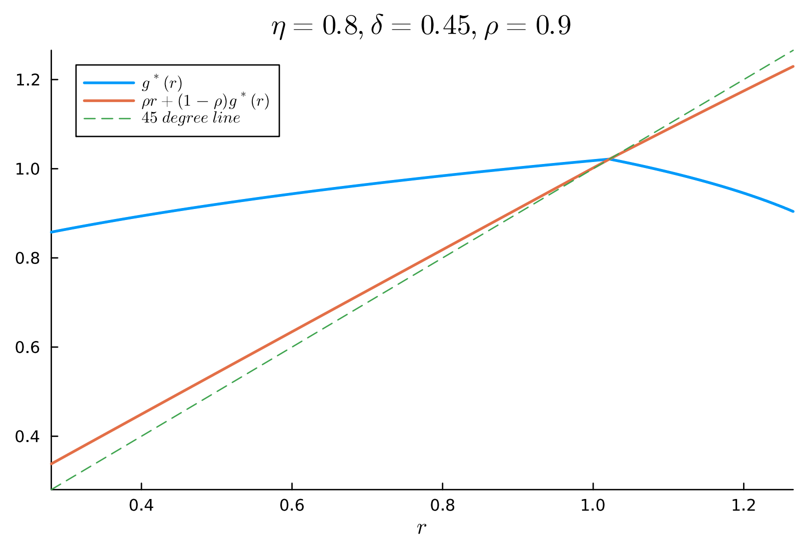
Proof. See Appendix E. □

The decreasing g^* reflects the strengthening loss-evading incentive. Assumption 5 ensures that the best deviation payoff always exceeds a reference point ($b(g(r)) > r$ for all r), guaranteeing the strictly increasing loss-evading incentive. The next subsection provides numerical examples, consistent with this result.

5.3 Numerical Examples of Case 2

This subsection provides numeric examples of Case 2 to obtain more sense of the optimized paths. Figure 2 shows Example 1: an example with high ρ and not very low δ. The policy function $g^*(r)$ (solid blue) is upward-sloping until it intersects the 45-degree (dashed green) line and then becomes downward-sloping. The intersection corresponds to the highest sustainable cooperation $\bar{z}$, and the properties of g^* are exactly as stated in Lemma 6 and Proposition 3. The other solid line (orange) locates the reference point in the next period ($\rho r + (1-\rho)g^*(r)$), which increases even after the intersection. In this example, when the initial reference point is unsustainable ($r > \bar{z}$) along the equilibrium path, the reference point gradually declines, the cooperation gradually rises, and both converge to the intersection from the right of the graph. Throughout this converging path, the players incur loss utilities.

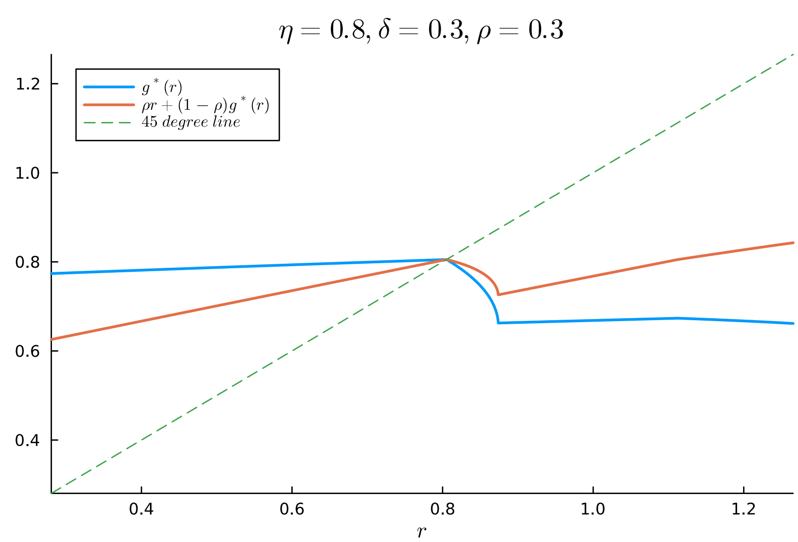
Figure 2
Case 2 - Example 1: Policy Function and Reference Point in Next Period



Notes: The intrinsic payoff $\pi(\alpha)$ is derived from Cournot competition by $N = 15$ firms with common marginal cost $c = 1$ with the inverse demand function $10 - \sum x_i$ where x_i is the quantity supplied by firm i.

An unsustainable reference point inhibits cooperation more severely when the players are less patient and when a reference point is less persistent. Figure 3 shows Example 2 of Case 2 in which δ and ρ are smaller than in Figure 2. There are four differences from Figure 2. First, the optimized cooperation g^* slope in the sustainable region ($\underline{\pi} \leq r \leq \bar{z}$) flattens, demonstrating that the cooperation loss aversion becomes weaker with less persistent losses in the deviation path and more discounting of future losses. Second, the highest sustainable cooperation $\bar{z}$ becomes smaller, caused by the smaller standard long-term loss with more discounting and the weakening cooperation loss aversion. Third, the downward slope in the unsustainable region ($\bar{z} < r$) close to $\bar{z}$ becomes steeper. This steeper slope indicates that the magnitude of the loss-evading incentive becomes relatively greater than in Example 1. A smaller discount factor and a less persistent reference point both reduce the long-term losses, whereas the standard short-term gain or the loss-evading incentive is unaffected. Consequently, a given increase in the loss-evading incentive tightens the constraint in $\Gamma(r, v^*)$ relatively more, generating a greater drop in cooperation. Finally, there are kinks in $g^*(r)$ and, subsequently, $\rho r + (1-\rho)g^*(r)$ around $r = 0.95$. Where the reference point is higher than this point, the best deviation cannot eliminate a loss in the current period ($r > b(g^*(r))$. A higher reference point in this region does not generate an additional incentive for deviation. This example works as a base for the analysis of transitional dynamics in the next section.

Figure 3
Case 2 - Example 2: Policy Function and Reference Point in Next Period



Notes: The intrinsic payoff $\pi(\alpha)$ is derived from Cournot competition by $N = 15$ firms with common marginal cost $c = 1$ with the inverse demand function $10 - \sum x_i$ where x_i is the quantity supplied by firm i.

6 Rough Transition

In this section, I study the model's implications for transitions of cooperation levels when the economic environment changes. To this end, I consider Case 2 for richer implications and assume that a reference point is at the steady state $\bar{z}$, corresponding to a mature relationship.

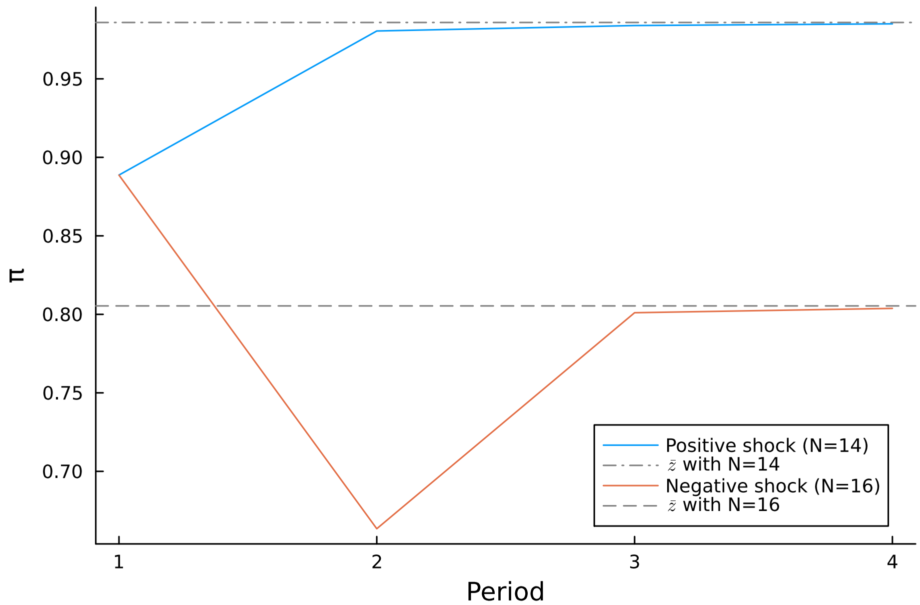
Reference dependence generates asymmetric transitional dynamics when $\bar{z}$ shifts. The steady state $\bar{z}$ moves when any of function b and parameters $(\underline{\pi}, \bar{\pi}, \delta, \eta, \rho)$ changes, which occurs for various reasons; for example, an entry or exit in a Cournot competition changes b, $\underline{\pi}$, and $\bar{\pi}$.[10] In the following analysis, I consider a Cournot-competition example in which the incumbents are tacitly colluding and achieving the highest sustainable profit $\bar{z}$ below the monopolist level $(\bar{\pi})$.

Cooperation monotonically transitions to the new steady state when the environment favorably changes. Suppose an incumbent exits the market at the end of period t. Then, it increases the profits of the remaining firms from period $t + 1$, working as a positive shock for them. Consequently, $\bar{z}$ rises, becoming higher than the reference point r_{t+1}, which is unaffected.[11] The result of $g(r) > r$ for $r < \bar{z}$ implies higher profits for the remaining firms in period $t+1$. Subsequently, the profits monotonically rise to $\bar{z}$. Figure 4 shows this path by a solid blue line. It uses the same parameters except for the number of firms as in Example 2 in Section 5, which uses a functional form of Cournot competition among 15 firms. In Figure 4, the number of firms decreases from 15 to 14 in period 2. This figure illustrates a smooth transition for firms, who do not suffer losses after the shock.

[10] A more straightforward example of the shift of $\bar{z}$ is a permanent intrinsic payoff shock. That is, $\pi^{new}_{t+k}(\boldsymbol{\alpha}) = \pi^{old}_{t+k}(\boldsymbol{\alpha}) + \epsilon$ for all $\boldsymbol{\alpha} \in A^N$ and all $k \in \mathbb{N}$ for a shock at the beginning of period $t+1$. This shock shifts everything but the reference point to the same extent. The insights in this section applies also to this simple case.

[11] Additionally, $\bar{\pi}$ and $\underline{\pi}$ rise, and $b(\pi)$ changes.

Figure 4
Transitional Dynamics from $\bar{z}$ with $N = 15$ in Cournot competition



Notes: These lines trace the level of intrinsic payoffs with $\eta = 0.8$, $\delta = 0.3$, and $\rho = 0.3$ and a shock to the number of firms at the end of period 1 in collusion under Cournot competition. Firms have a common marginal cost $c = 1$ with the inverse demand function $10 - \sum x_i$ where x_i is the quantity supplied by firm i. In period 1, $r_1 = \pi_1 = \bar{z}$ for $N = 15$. In period 2 and later, π_t moves according to the policy functions for $N = 14$ (blue) and $N = 16$ (orange) with $r_2 = \pi_1$.

In contrast, a negative shock causes an excessive decline in profits in the short term, traced by the solid orange line in Figure 4 where the number of firms increases to 16. When a new firm enters the market and joins the tacit collusion, individual firms' profits fall. Thus, the entry has a negative effect, pushing down the steady state $\bar{z}$.[12] The result of $g(r) < \bar{z}$ for $r > \bar{z}$ implies a profit in period $t + 1$ (period 2 in Figure 4) lower than the new steady state level $\bar{z}$, let alone the previous steady-state level. This overshoot reflects the loss-evading incentive as discussed in Section 5. Firms are accustomed to a high level of profit. Even after this high profit level becomes infeasible in a new environment, they try to cling to it by deviating from the collusion. This incentive makes their cooperation difficult, lowering the profit. After the drop in period $t + 1$, the profit converges to the new steady state level

[12] I assume the entrant has the same reference point as the incumbents when it enters.

after firms eventually adapt to a lower profit level.

These dynamics are similar to those of Green and Porter (1984) in that a sharp drop in the payoff level does not imply a collapse of collusion or cooperative relationships. Green and Porter (1984) study a model with unobserved demand shocks in a stationary environment, showing that prices and profits can sharply decline from and return to the monopolistic levels even when collusion operates. In contrast to their work, the return to a higher price—the new steady state—is gradual in my model; thus, reference-dependent firms "restart small." As a similar exercise, I analyze how cooperation levels respond to unexpected one-off payoff shocks in the Online Appendix. The shock does not affect the steady state, making the analysis more comparable to Green and Porter (1984). Following the one-off shock, reference points shift away from the steady state, preventing players from implementing the steady-state level in the next period; thus, even a positive shock can reduce the cooperation level. The cooperation returns to the steady-state level only gradually, unlike in Green and Porter (1984).

7 Conclusion

This paper provides a new explanation for the gradual development of cooperation. In existing models of gradualism with incomplete information, the presence of myopic players forces cooperation to start at a low level. In contrast, reference dependence raises the initial level of cooperation and facilitates its growth over time. However, reference points also create an additional incentive to deviate—the loss-evading incentive—suppressing cooperation when players must scale back their cooperation levels.

A limitation of this analysis is that it examines SPEs using Nash-reversion strategies to derive the implications of reference dependence in a tractable way. Given the complexity of the model, this approach is reasonable while still capturing the core mechanism at play. Nevertheless, investigating the optimal penalty in the presence of reference dependence may yield further insights.

References

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. Econometrica: Journal of the Econometric Society, 383–396.

Bernardes, L. G. (2003). Reference-dependent preferences and the speed of economic liberalization. The Journal of Socio-Economics 32(5), 521–548.

Bond, E. W. and J.-H. Park (2002). Gradualism in trade agreements with asymmetric countries. The Review of Economic Studies 69(2), 379–406.

Bowman, D., D. Minehart, and M. Rabin (1999). Loss aversion in a consumption–savings model. Journal of Economic Behavior & Organization 38(2), 155–178.

Chan, J. M. (2019). Gradualism in the gatt: Strategic tariff bargaining and forward manipulation. Review of International Economics 27(1), 220–239.

Datta, S. (1996). Building trust. Sticerd - theoretical economics paper series, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

DellaVigna, S., A. Lindner, B. Reizer, and J. F. Schmieder (2017). Reference-dependent job search: Evidence from hungary. The Quarterly Journal of Economics 132(4), 1969–2018.

Devereux, M. B. (1997). Growth, specialization, and trade liberalization. International Economic Review, 565–585.

Freund, C. and Ç. Özden (2008). Trade policy and loss aversion. American Economic Review 98(4), 1675–1691.

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. The Review of Economic Studies 38(1), 1–12.

Furusawa, T. and T. Kawakami (2008). Gradual cooperation in the existence of outside options. Journal of Economic Behavior & Organization 68(2), 378–389.

Furusawa, T. and E. L.-C. Lai (1999). Adjustment costs and gradual trade liberalization. Journal of International Economics 49(2), 333–361.

Ghosh, P. and D. Ray (1996). Cooperation in community interaction without information flows. The Review of Economic Studies 63(3), 491–519.

Green, E. J. and R. H. Porter (1984). Noncooperative collusion under imperfect price information. Econometrica: Journal of the Econometric Society, 87–100.

Hua, X. and J. Watson (2022). Starting small in project choice: A discrete-time setting with a continuum of types. Journal of Economic Theory 204, 105490.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1990). Experimental tests of the endowment effect and the coase theorem. Journal of Political Economy 98(6), 1325–1348.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. Econometrica 47(2), 263–292.

Kőszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. The Quarterly Journal of Economics 121(4), 1133–1165.

Kranton, R. E. (1996). The formation of cooperative relationships. The Journal of Law, Economics, and Organization 12(1), 214–233.

Lockwood, B. and J. P. Thomas (2002). Gradualism and irreversibility. The Review of Economic Studies 69(2), 339–356.

Maggi, G. and A. Rodriguez-Clare (2007). A political-economy theory of trade agreements. American Economic Review 97(4), 1374–1406.

McMillan, J. and C. Woodruff (1999). Interfirm relationships and informal credit in vietnam. The Quarterly Journal of Economics 114(4), 1285–1320.

Morduch, J. (1999). The microfinance promise. Journal of economic literature 37(4), 1569–1614.

Post, T., M. J. Van den Assem, G. Baltussen, and R. H. Thaler (2008). Deal or no deal? decision making under risk in a large-payoff game show. American Economic Review 98(1), 38–71.

Rauch, J. E. and J. Watson (2003). Starting small in an unfamiliar environment. International Journal of industrial organization 21(7), 1021–1042.

Sobel, J. (1985). A theory of credibility. The Review of Economic Studies 52(4), 557–573.

Souchou, Y. (1987). The fetish of relationships: Chinese business transactions in singapore. Sojourn: Journal of Social Issues in Southeast Asia 2(1), 89–111.

Staiger, R. W. (1994). A theory of gradual trade liberalization.

Tovar, P. (2009). The effects of loss aversion on trade policy: Theory and evidence. Journal of International Economics 78(1), 154–167.

Tversky, A. and D. Kahneman (1991). Loss aversion in riskless choice: A reference-dependent model. The Quarterly Journal of Economics 106(4), 1039–1061.

Watson, J. (1999). Starting small and renegotiation. Journal of Eßconomic Theory 85(1), 52–90.

Watson, J. (2002). Starting small and commitment. Games and Economic Behavior 38(1), 176–199.

Zissimos, B. (2007). The gatt and gradualism. Journal of International Economics 71(2), 410–433.

Appendix

A Proof of Lemma 1

I consider a subgame starting from period t. The utility for player i is given by $U_{i,t}(\{\boldsymbol{\alpha}_{t+k}\}_{k=0}^{\infty}, r_{i,t})$. The derivative of the utility with respect to $\alpha_{i,t+m}$ is:

$$\frac{\partial U_{i,t}}{\partial \alpha_{i,t+m}} = \delta^m \frac{\partial \pi_i(\alpha_{i,t+m}, \boldsymbol{\alpha}_{-i,t+m})}{\partial \alpha_{i,t+m}} \Phi,$$

where

$$\Phi = 1 + \mathbb{1}\{\pi(\alpha_{i,t+m}, \boldsymbol{\alpha}_{-i,t+m}) < r_{i,t+m}\}\eta - \sum_{n=1}^{\infty} \delta^n \mathbb{1}\{\pi(\alpha_{i,t+m+n}, \boldsymbol{\alpha}_{-i,t+m}) < r_{i,t+m+n}\} \cdot \eta \cdot \frac{\partial r_{i,t+m+n}}{\partial \pi(\alpha_{i,t+m}, \boldsymbol{\alpha}_{-i,t+m})}.$$

Φ is positive as

$$\Phi \geq 1 - \sum_{n=1}^{\infty} \delta^n \eta \cdot \frac{\partial r_{i,t+m+n}}{\partial \pi_i(\alpha_{i,t+m}, \boldsymbol{\alpha}_{-i,t+m})} = 1 - \eta \sum_{n=1}^{\infty} \delta^n \cdot (1-\rho)\rho^{n-1} = 1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho} > 0,$$

where the last inequality follows from $\eta \in (0,1]$ and $\delta, \rho \in (0,1)$. Thus, $\partial U_{i,t}/\partial \alpha_{i,t+m}$ and $\partial \pi_i(\alpha_{i,t+m}, \boldsymbol{\alpha}_{-i,t+m})/\partial \alpha_i$ have the same sign, implying $\text{argmax}_{\alpha_{i,t+m} \in A} U_{i,t}(\{\alpha_{i,t+k}, \boldsymbol{\alpha}_{-i,t+k}\}_{k=1}^{\infty}, r_{i,t}) = \text{argmax}_{\alpha_{i,t+m} \in A} \pi_i(\alpha_{i,t+m}, \boldsymbol{\alpha}_{-i,t+m})$ for all $m \in \mathbb{N}$. Therefore, $(s^N, s^N, \ldots, s^N)$ is a Nash equilibrium in this subgame. This completes the proof.

B Proof of Lemma 2

Suppose player i deviates in period t when the others play $s(\{\alpha_{t+k}^c\}_{k=0}^{\infty}, r_{i,t})$. Then, the derivative of utility with respect to $\alpha_{i,t}$ is $\partial \pi_i(\alpha_{i,t}, \alpha_t^c)/\partial \alpha_{i,t})\Phi'$ where $\Phi' = 1 + \mathbb{1}\{\pi_i(\alpha_{i,t}, \boldsymbol{\alpha}_{-i,t}^c) < r_t\} \cdot \eta - \sum_{k=1}^{\infty} \delta^k \cdot \mathbb{1}\{\tilde{\pi}(\alpha^N) < r_{i,t+k}\} \cdot \eta \cdot (\partial r_{i,t+k}/\partial \pi_i(\boldsymbol{\alpha}_t))$ evaluated at $\boldsymbol{\alpha}_t = (\alpha_{i,t}, \boldsymbol{\alpha}_{-i,t}^c)$. I manipulate this Φ' as:

$$\Phi' \geq 1 - \eta \sum_{k=1}^{\infty} \delta^k \frac{\partial r_{i,t+k}}{\partial \pi_i(\boldsymbol{\alpha}_t)}_{|\alpha_t = (\alpha_{i,t}, \alpha_{-i,t}^c)} = 1 - \eta \sum_{k=1}^{\infty} \delta^k \cdot (1-\rho)\rho^{k-1} = 1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho} > 0,$$

where the last inequality follows from $\eta \in (0,1]$ and $\delta, \rho \in (0,1)$. Thus, the sign of $(\partial \pi_i(\alpha_{i,t}, \alpha_t^c)/\partial \alpha_{i,t})\Phi'$ equals to that of $\partial \pi_i(\alpha_{i,t}, \alpha_t^c)/\partial \alpha_{i,t}$, implying Lemma 2.

C Proof of Lemma 4

I show the function v^* satisfies (5). v^* is bounded as: $\underline{\pi}/(1-\delta) \leq v^* \leq \bar{\pi}/(1-\delta)$. Since $\pi_t = \underline{\pi}$ is always feasible, $\Gamma(r_t, v^*)$ is nonempty for all r_t. Let $\Psi(r_t) \equiv \Big\{ \{\pi_k\}_{k=t}^{\infty} \in [\underline{\pi}, \bar{\pi}]^{\infty} : \pi_k \in \Gamma(r_k, v^*),\ s.t.\ r_{k+1} = \rho r_k + (1-\rho)\pi_k \Big\}$ be the set of feasible cooperation paths from r_t. Given $\rho r_t + (1-\rho)\pi_t^o \in [\underline{\pi}, \bar{\pi}]$, there exists an optimized path $\{\pi_k\}_{k=t+1}^{\infty}$ such that $v^*(\rho r_t + [1-\rho]\pi_t^o) = F(\{\pi_k\}_{k=t+1}^{\infty})$ where $F(\{\pi_k\}_{k=t+1}^{\infty}) = \sum_{k=t+1}^{\infty} \delta^{k-t} \Big[\pi_k + \mathbb{1}\{r_k > \pi_k\} \cdot \eta(r_k - \pi_k) \Big]$. It follows that, given $\pi_t, r_t \in [\underline{\pi}, \bar{\pi}]$,

$$\begin{aligned} &\pi_t^o - \mathbb{1}\{r_t > \pi_t^o\} \cdot \eta(r_t - \pi_t^o) + \delta v^*(\rho r_t + [1-\rho]\pi_t^o) \\ &= \pi_t^o - \mathbb{1}\{r_t > \pi_t^o\} \cdot \eta(r_t - \pi_t^o) + \delta \max_{\{\pi_k\}_{k=t+1}^{\infty} \in \Psi([1-\rho] r_t + \rho \pi_t^o)} F(\{\pi_k\}_{k=t+1}^{\infty}) \\ &\leq \max_{\{\pi_k\}_{k=t}^{\infty} \in \Psi(r_t)} \{\pi_t^o - \mathbb{1}\{r_t > \pi_t^o\} \cdot \eta(r_t - \pi_t^o) + \delta F(\{\pi_k\}_{k=t+1}^{\infty})\} = v^*(r_t) \end{aligned} \tag{8}$$

where I use the fact that $\{\pi_t^o, \{\pi_k\}_{k=t+1}^{\infty}\} \in \Psi(r_t)$ when $\{\pi_k\}_{k=t+1}^{\infty} \in \Psi([1-\rho] r_t + \rho \pi_t^o)$. At the same time, given $r_t \in [\underline{\pi}, \bar{\pi}]$,

$$\begin{aligned} v^*(r_t) &= \max_{\{\pi_k\}_{k=t}^{\infty} \in \Psi(r_t)} F(\{\pi_k\}_{k=t}^{\infty}) = \pi_t^o - \mathbb{1}\{r_t > \pi_t^o\} \cdot \eta(r_t - \pi_t^o) + \delta F(\{\pi_k^o\}_{k=t+1}^{\infty}) \\ &\leq \pi_t^o - \mathbb{1}\{r_t > \pi_t^o\} \cdot \eta(r_t - \pi_t^o) + \delta v^*(\rho r_t + [1-\rho]\pi_t^o) \end{aligned} \tag{9}$$

where $\{\pi_t^o, \{\pi_k^o\}_{k=t+1}^{\infty}\} = \text{argmax}_{\{\pi_k\}_{k=t}^{\infty} \in \Psi(r_t)} F(\{\pi_k\}_{k=t}^{\infty})$. Thus, it follows from inequalities (8) and (9) that

$$v^*(r_t) = \pi_t^o - \mathbb{1}\{r_t > \pi_t^o\} \cdot \eta(r_t - \pi_t^o) + \delta v^*(\rho r_t + [1-\rho]\pi_t^o)$$

v^* satisfies the functional equation (5).

D Proof of Lemma 5

I first obtain Lemma 7 that enables us to exclude the possibility of $v^*(\underline{\pi}) = \underline{\pi}/(1-\delta)$.

Lemma 7. There exists $\pi' > \underline{\pi}$ such that $v^*(\underline{\pi}) \geq \pi'/(1-\delta)$.

Proof. Consider a constant cooperation path $\{\pi\}_{t=1}^{\infty}$ in problem 4. Given $r_1 = \underline{\pi}$, the following inequality is sufficient for $v^*(\underline{\pi}) \geq \pi/(1-\pi)$.

$$\forall t, \ \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(\pi) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} r_t \leq \frac{\pi}{1-\delta}, \tag{10}$$

where the indicator function is absent because $\pi \geq r_t$ for all t. Subsequently, inequality (10) for $t = 1$ becomes

$$f(\pi) \equiv \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(\pi) - \frac{\pi}{1-\delta} + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right]\underline{\pi} \leq 0.$$

$b(\underline{\pi}) = \underline{\pi}$ implies $f(\underline{\pi}) = 0$. Taking derivative of $f(\pi)$ yields

$$\begin{aligned} f'(\pi) &= \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b'(\pi) - \frac{1}{1-\delta} \\ &= \left(1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right)\left(\sum_{j\neq i} \frac{\partial \pi_i(\alpha_i, \boldsymbol{\alpha}^c{}_{-i})}{\partial \alpha_{j,t}}_{|\alpha_i = \operatorname{argmax} \pi_i(\alpha_i, \boldsymbol{\alpha}^c{}_{-i})}\right) \frac{d\alpha_t}{d\tilde{\pi}(\alpha_t)}_{|\alpha_t = \alpha^c} - \frac{1}{1-\delta}, \end{aligned}$$

where the second equality follows from the first order condition for $\alpha_i = \operatorname{argmax} \pi_i(\alpha_i, \boldsymbol{\alpha}^c{}_{-i})$. Evaluating this derivative at $\pi = \underline{\pi}$ yields

$$\begin{aligned} f'(\underline{\pi}) &= \left(1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right)\left(\sum_{j\neq i} \frac{\partial \pi_i(\boldsymbol{\alpha}^N)}{\partial \alpha_{j,t}}\right) \frac{d\alpha_t}{d\tilde{\pi}(\alpha_t)}_{|\alpha_t = \alpha^c} - \frac{1}{1-\delta} \\ &= \left(1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right)\left(\sum_{j\neq i} \frac{\partial \pi_i(\boldsymbol{\alpha}^N)}{\partial \alpha_{j,t}}\right)\left(\sum_{j\neq i} \frac{\partial \pi_i(\boldsymbol{\alpha}^N)}{\partial \alpha_{j,t}}\right)^{-1} - \frac{1}{1-\delta} \\ &= \frac{1 - \delta\rho - \delta\eta + \delta\eta\rho}{1-\delta\rho} - \frac{1}{1-\delta} < 0, \end{aligned}$$

where the second equality follows from $\boldsymbol{\alpha}^N = \operatorname{argmax} \pi_i(\alpha_i, \boldsymbol{\alpha}^N{}_{-i})$. $f(\underline{\pi}) = 0$ and $f'(\underline{\pi}) < 0$ imply that there exists $\epsilon > 0$ such that $f(\underline{\pi} + \epsilon) < 0$. Thus, $\{\underline{\pi} + \epsilon\}_{t=1}^{\infty}$ satisfy inequalities (10) for $t = 1$ and for $t \geq 2$ because $r_t \geq r_1$. Let $\pi' = \underline{\pi} + \epsilon$. Then, $v^*(\underline{\pi}) \geq \pi'/(1-\delta)$. This completes the proof of Lemma 7. □

Using π' in Lemma 7, I define the interval X and the set of functions $C(X)$ as

$$\begin{aligned}
X &: [\underline{\pi}, \bar{\pi}] \\
C(X) &: \text{the set of bounded, continuous, and weakly increasing} \\
&\quad \text{functions } f : X \to R \text{ with the sup norm that are weakly concave on } X \\
&\quad s.t. \begin{cases} \pi'/(1-\delta) \leq f(x) \leq \bar{\pi}/(1-\delta) & x \leq \gamma, \\ f(x) = \bar{\pi}/(1-\delta) & x \geq \gamma, \end{cases}
\end{aligned}$$

where γ is implicitly defined by the following equation:

$$\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(\bar{\pi}) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho}\gamma = \frac{\bar{\pi}}{1-\delta}. \tag{11}$$

Assumption 4 and the definition of Case 1 imply $\underline{\pi} < \gamma < \bar{\pi}$. On $C(X)$, I define the operator T by

$$Tf(x) = \max_{y \in \Gamma(x;f)} \left[y - \mathbb{1}\{x > y\} \cdot \eta(x-y) + \delta f(\rho x + [1-\rho]y)\right],$$

where

$$\begin{aligned}
\Gamma(x; f) = \Bigg\{ x \in [\underline{\pi}, \bar{\pi}] : &\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(y) + \mathbb{1}\{x > y\} \cdot \eta\left[\min\{x, b(y)\} - y\right] \\
&+ \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} x \leq y + \delta f(\rho x + [1-\rho]y) \Bigg\}.
\end{aligned}$$

Given function f, I obtain properties of Tf. The policy function, $g(x; f)$, is defined by

$$g(x; f) = \operatorname*{argmax}_{y \in \Gamma(x;f)} \; y - \mathbb{1}\{x > y\} \cdot \eta(x-y) + \delta f(\rho x + [1-\rho]y).$$

Since f is a weakly increasing function, $g(x; f) = \max \Gamma(x; f)$. By manipulating the condition of $\Gamma(x; f)$, let $h(x, y; f)$ be

$$\begin{aligned}
h(x, y; f) = &\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(y) + \mathbb{1}\{x > y\} \cdot \eta\left[\min\{x, b(y)\} - y\right] \\
&+ \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} x - y - \delta f(\rho x + [1-\rho]y).
\end{aligned}$$

y belongs to $\Gamma(x; f)$ if and only if $h(x, y; f) \leq 0$.

I eliminate the indicator function by showing $g(x; f) \geq x$. I define $\tilde{h}(x; f)$ by

$$\tilde{h}(x; f) = h(x, x; f) = \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(x) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right] \underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} x - x - \delta f(x).$$

$\tilde{h}(x; f)$ is negative at $x = \underline{\pi}$ as:

$$\begin{aligned}
\tilde{h}(\underline{\pi}; f) &= \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(\underline{\pi}) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right] \underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} \underline{\pi} - \underline{\pi} - \delta f(\underline{\pi}) \\
&\leq \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] \underline{\pi} + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right] \underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} \underline{\pi} - \underline{\pi} - \frac{\delta\pi'}{1-\delta} \\
&= -\frac{\delta}{1-\delta}(\pi' - \pi) < 0,
\end{aligned}$$

where the inequality follows from $b(\underline{\pi}) = \underline{\pi}$ and $f(\underline{\pi}) \geq \pi'/(1-\delta)$. Also, $\tilde{h}(x; f)$ is non-positive at $\bar{\pi}$ by construction of Case 1. The strict convexity of $b(\cdot)$ and the weak concavity of $f(\cdot)$ imply $\tilde{h}(x; f)$ is strictly convex in x. These imply that, for all $x \in X$, $\tilde{h}(x, f) \leq 0$. In turn, this implies that $g(x; f) \geq x$ for all $x \in X$, eliminating the indicator function.

$h(x, g(x); f) = 0$ on $[\underline{\pi}, \gamma]$ I now show that the equality of $h(x, y; f) \leq 0$ holds with $y = g(x; f) = \max \Gamma(x; f)$ for all $x \leq \gamma$. From the result above, given $x < \gamma$, $h(x, x; f) = \tilde{h}(x; f) < 0$. On the other hand,

$$\begin{aligned}
h(x, \bar{\pi}; f) &= \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(\bar{\pi}) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right] \underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} x - \bar{\pi} - \delta f(\rho x + [1-\rho]\bar{\pi}) \\
&= \frac{\delta\eta\rho}{1-\delta\rho}(\gamma - \bar{\pi}) + \delta f(\rho\gamma + [1-\rho]\bar{\pi}) - \delta f(\rho x + [1-\rho]\bar{\pi}) > 0,
\end{aligned}$$

where the second inequality follows from eq. (11) and f is weakly increasing. Since $b(\cdot)$ and $-f(\cdot)$ are convex and continuous functions, given x, $h(x, y; f)$ is convex and continuous in y. This convexity, continuity, $h(x, x; f) < 0$ and $h(x, \bar{\pi}; f) > 0$ imply $h(x, g(x; f); f) = 0$.

Tf and g are strictly increasing on $[\underline{\pi}, \gamma]$ $Tf(x)$ strictly increases in x on $[\underline{\pi}, \gamma]$ because $g(x)$ strictly increases on $[\underline{\pi}, \gamma]$, following from $g(x) = \max \Gamma(x; f)$.

Tf and g are strictly concave on $[\underline{\pi}, \gamma]$ $\quad Tf(x)$ is strictly concave on $(\underline{\pi}, \gamma)$ because $g(x)$ is strictly concave on $(\underline{\pi}, \gamma)$, as below. For $x_1, x_2 \in [\underline{\pi}, \gamma]$ such that $x_1 \neq x_2$, $g(x_1)$ and $g(x_2)$ satisfy the following equations, respectively.

$$\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(x_1)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} x_1 = g(x_1) + \delta f(\rho x_1 + [1-\rho] g(x_1)),$$

$$\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(x_2)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} x_2 = g(x_2) + \delta f(\rho x_2 + [1-\rho] g(x_2)). \tag{12}$$

Combining these two equations with a weight $\theta \in (0, 1)$ yields

$$\begin{aligned} &\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] \{[1-\theta] b(g(x_1)) + \theta b(g(x_2))\} + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} \\ &\quad - \frac{\delta\eta\rho}{1-\delta\rho} ([1-\theta] x_1 + \theta x_2) \\ &= (1-\theta) g(x_1) + \theta g(x_2) + (1-\theta)\,\delta f\,(\rho x_1 + [1-\rho] g(x_1)) + \theta\delta f\,(\rho x_2 + [1-\rho] g(x_2)). \end{aligned}$$

Since $g(x)$ strictly increases on $[\underline{\pi}, \gamma]$, $g(x_1) \neq g(x_2)$. It follows from the convexity of $b(\cdot)$ and the concavity of $f(\cdot)$ that:

$$\begin{aligned} &\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] \{b([1-\theta] g(x_1) + \theta g(x_2))\} + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} ([1-\theta] x_1 + \theta x_2) \\ &< (1-\theta) g(x_1) + \theta g(x_2) + \delta f\,(\rho\,\{[1-\theta] x_1 + \theta x_2\} + (1-\rho)\,\{[1-\theta] g(x_1) + \theta g(x_2)\}). \end{aligned} \tag{13}$$

Since $[1-\theta] x_1 + \theta x_2 \in (\underline{\pi}, \gamma)$, $g([1-\theta] x_1 + \theta x_2) \in (\underline{\pi}, \gamma)$ satisfies the following equality.

$$\begin{aligned} &\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] \{b(g([1-\theta] x_1 + \theta x_2))\} + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} ([1-\theta] x_1 + \theta x_2) \\ &\qquad = g([1-\theta] x_1 + \theta x_2) + \delta f\,(\rho\,\{[1-\theta] x_1 + \theta x_2\} + (1-\rho) g([1-\theta] x_1 + \theta x_2)). \end{aligned} \tag{14}$$

Eq. (13) and (14) imply $(1-\theta) g(x_1) + \theta g(x_2) < g([1-\theta] x_1 + \theta x_2)$. That is, $g(x)$ is strictly concave on $(\underline{\pi}, \gamma)$. Consequently, $Tf(x)$ is strictly concave on $(\underline{\pi}, \gamma)$.

Properties on $[\gamma, \bar{\pi}]$ For the properties on $[\gamma, \bar{\pi}]$, $h(x, \bar{\pi}; f) \leq 0$ for all $x \geq \gamma$, following from eq. (11), and $\bar{\pi} \in \Gamma(x; f)$. This immediately implies $g(x; f) = \bar{\pi}$ and $Tf(x) = \bar{\pi}/(1-\delta)$. These also imply that $g(x; f)$ and $Tf(x)$ are weakly concave and weakly increasing on X.

Continuity The weak concavities of $g(x)$ and $Tf(x)$ imply their continuities on $(\underline{\pi}, \bar{\pi})$. They are also continuous at $\bar{\pi}$, following from the results on $[\gamma, \bar{\pi}]$. As for the continuity at $\underline{\pi}$, I prove that of $g(x)$ first. I take a sequence, $\{x_t\} \subset [\underline{\pi}, \gamma]$, that converges to $\underline{\pi}$ as $t \to \infty$. Then, for all x_t,

$$g(x_t) = \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(x_t)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} x_t - \delta f(\rho\underline{\pi} + [1-\rho] g(x_t)).$$

The continuities of $b(\cdot)$ and $f(\cdot)$ imply

$$\lim_{x_t \to \underline{\pi}} g(x_t) = \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(\lim_{x_t \to \underline{\pi}} g(x_t)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} \\ - \frac{\delta\eta\rho}{1-\delta\rho}\underline{\pi} - \delta f(\rho\underline{\pi} + [1-\rho] \lim_{x_t \to \underline{\pi}} g(x_t)).$$

This equality implies $\lim_{x_t \to \underline{\pi}} g(x_t) = g(\underline{\pi})$.

Pointwise Convergence The results above imply $T : C(X) \to C'(X)$ where $C'(x)$ is defined as:

$$\begin{aligned} C'(X) : &\text{ the set of bounded, continuous, weakly increasing, and weakly concave} \\ &\text{ functions } f : X \to R \text{ with the sup norm that are} \\ &\text{ strictly increasing and strictly concave on } [\underline{\pi}, \gamma] \\ &\ s.t. \begin{cases} \underline{\pi}'/(1-\delta) \leq f(x) \leq \bar{\pi}/(1-\delta) & x \leq \gamma, \\ f(x) = \bar{\pi}/(1-\delta) & x \geq \gamma, \end{cases} \end{aligned}$$

which is a subset of $C(X)$. Using this operator T, I define a sequence of functions produced by T by

$$\begin{cases} f_1(x) & = \frac{\bar{\pi}}{1-\delta} \quad \text{for all} \quad x \in X, \\ f_{n+1} & = Tf_n \quad \text{for} \quad n \in \mathbb{N}. \end{cases}$$

f_1 is in $C(X)$. It follows that

$$\begin{cases} f_2(x) < f_1(x) & x < \gamma \\ f_2(x) = f_1(x) & x \geq \gamma \end{cases}$$

For $x < \gamma$, given $f(x) < f'(x)$, $g(x; f) = \max \Gamma(x; f) < \max \Gamma(x; f') = g(x, f')$, implying $Tf(x) < Tf'(x)$. By this monotonicity of T, $f_3 = Tf_2 < Tf_1 = f_2$. Repeating this operation yields

$$\begin{cases} f_n(x) < \cdots < f_1(x) & x < \gamma, \\ f_n(x) = \cdots = f_1(x) & x \geq \gamma. \end{cases}$$

Since $\{f_n\}_{n \in \mathbb{N}}$ is bounded, there exists $f^*(x)$ such that $f_n(x) \to f^*(x)$ as $n \to \infty$.

Uniform Convergence   I show the convergence is uniform. For all $n + 1$, and $y < x \in X$,

$$|f_{n+1}(x) - f_{n+1}(y)| = |Tf_n(x) - Tf_n(y)| = \begin{cases} 0 & x, y \geq \gamma, \\ \frac{\bar{\pi}}{1-\delta} - Tf_n(y) & y < \gamma \leq x, \\ Tf_n(x) - Tf_n(y) & y < x < \gamma. \end{cases}$$

When $y < x < \gamma$,

$$\begin{aligned}|f_{n+1}(x) - f_{n+1}(y)| =& \Bigg| \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(x; f_n)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho}x \\ & - \left\{\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(y; f_n)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho}y\right\} \Bigg| \\ =& |x-y| \cdot \left| \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] \frac{b(g(x; f_n)) - b(g(y; f_n))}{x-y} - \frac{\delta\eta\rho}{1-\delta\rho} \right| \\ <& |x-y| \cdot \max\left\{ \left| \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b'(\bar{\pi}) \max_{x,y} \frac{g(x)-g(y)}{x-y} - \frac{\delta\eta\rho}{1-\delta\rho} \right|, \right. \\ & \left. \left| \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b'(\underline{\pi}) \min_{x,y} \frac{g(x)-g(y)}{x-y} - \frac{\delta\eta\rho}{1-\delta\rho} \right| \right\}. \end{aligned} \tag{15}$$

where the last inequality follows from the convexity of $b(\cdot)$. I show $[g(x) - g(y)]/(x - y)$ on the right-hand side (RHS) is bounded. $g(x)$ strictly increasing implies $h(x_1, g(x_2); f) - h(x_1, g(x_1); f) > 0$ for $x_1 < x_2 < \gamma$. Thus,

$$\begin{aligned} \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(x_1)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho}x_1 &= g(x_1) + \delta f(\rho x_1 + [1-\rho]g(x_1)). \\ \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(x_2)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho}x_1 &> g(x_2) + \delta f(\rho x_1 + [1-\rho]g(x_2)). \end{aligned}$$

It follows

$$\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] \frac{b(g(x_2)) - b(g(x_1))}{g(x_2) - g(x_1)} - 1 - \delta\frac{f(\rho x_1 + [1-\rho]g(x_2)) - f(\rho x_1 + [1-\rho]g(x_1))}{g(x_2) - g(x_1)} > 0. \tag{16}$$

Subsequently, $h(x_2, g(x_2); f) - h(x_1, g(x_1); f) = 0$ implies

$$\begin{aligned}
&\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] \frac{b(g(x_2)) - b(g(x_1))}{g(x_2) - g(x_1)} \frac{g(x_2) - g(x_1)}{x_2 - x_1} - \frac{\delta\eta\rho}{1-\delta\rho} - \frac{g(x_2) - g(x_1)}{x_2 - x_1} \\
&- \delta \frac{f(\rho x_2 + [1-\rho]g(x_2)) - f(\rho x_1 + [1-\rho]g(x_2))}{x_2 - x_1} \\
&- \delta \frac{f(\rho x_1 + [1-\rho]g(x_2)) - f(\rho x_1 + [1-\rho]g(x_1))}{g(x_2) - g(x_1)} \frac{g(x_2) - g(x_1)}{x_2 - x_1} = 0 \\
\iff &\left\{ \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] \frac{b(g(x_2)) - b(g(x_1))}{g(x_2) - g(x_1)} - 1 \right. \\
&\left. - \delta \frac{f(\rho x_1 + [1-\rho]g(x_2)) - f(\rho x_1 + [1-\rho]g(x_1))}{g(x_2) - g(x_1)} \right\} \frac{g(x_2) - g(x_1)}{x_2 - x_1} \\
&- \delta \frac{f(\rho x_2 + [1-\rho]g(x_2)) - f(\rho x_1 + [1-\rho]g(x_2))}{x_2 - x_1} = \frac{\delta\eta\rho}{1-\delta\rho}.
\end{aligned}$$

The combination of $\rho x_1 + [1-\rho]g(x_2) > \underline{\pi}$ and the concavity of f implies that the second term on the left-hand side of the last equality is bounded. The RHS is also bounded. Consequently, given inequality (16), $[g(x_2) - g(x_1)]/[x_2 - x_1]$ must be bounded. Consequently, there exists some $K_{n1} < \infty$ such that $|f_{n+1}(x) - f_{n+1}(y)| < K_{n1}|x - y|$ for all $x, y < \gamma$ because of inequality (15).

When $y < \gamma \leq x$,

$$\begin{aligned}
|f_{n+1}(x) - f_{n+1}(y)| = &\left|(x - y)\frac{\gamma - y}{x - y}(T f_n(\gamma) - T f_n(x))\right| \\
\leq &|x - y| \left|\frac{\gamma - y}{x - y}\right| \\
&\cdot \max\left\{ \left| \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b'(\bar{\pi}) \max_{x,y} \frac{g(x) - g(y)}{x - y} - \frac{\delta\eta\rho}{1-\delta\rho} \right|, \right. \\
&\left. \left| \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b'(\underline{\pi}) \min_{x,y} \frac{g(x) - g(y)}{x - y} - \frac{\delta\eta\rho}{1-\delta\rho} \right| \right\}.
\end{aligned}$$

Similar to the previous case, there exists some $K_{n2} < \infty$ such that $|f_{n+1}(x) - f_{n+1}(y)| < K_{n2}|x - y|$ for all $y < \gamma$ and $x \geq \gamma$. These results imply each f_n is Lipschitz continuous with $\max\{K_{11}, K_{12}, \ldots K_{n1}, K_{n2} \ldots\}$ as the common Lipschitz constant. In turn, it implies $\{f_n\}_{n\in\mathbb{N}}$ is equicontinuous. Additionally, $\{f_n\}_{n\in\mathbb{N}}$ is bounded, and X is closed and bounded. Thus, the convergence of $\{f_n\}_{n\in\mathbb{N}}$ is uniform by the Ascoli–Arzelà theorem. Finally, $(C(X), \|\cdot\|_\infty)$ is a compact set, implying $f^* \in C(X)$.

f^* satisfies the functional equation   Next, I prove that f^* satisfies the functional equation by showing that f_n converges to Tf^*. It follows from the uniform convergence that for any $\varepsilon > 0$, there exists N such that $\|f_n - f^*\| < \varepsilon$ for all $n > N$. Additionally, $f^* \leq f_n$, implying $f_n < f^* + \varepsilon$ on X for $n > N$. By the monotonicity, $Tf_n < T(f^* + \varepsilon)$. Subsequently,

$$\begin{aligned}
\|Tf_n - Tf^*\| &= \sup_{x\in X} |Tf_n(x) - Tf^*(x)| \\
&< \sup_{x\in X} |Tf^*(x+\varepsilon) - Tf^*(x)| \quad (n > N) \\
&= \sup_{x\in X}\ l(x, f^*, \varepsilon),
\end{aligned} \tag{17}$$

where

$$\begin{aligned}
l(x, f^*, \varepsilon) =& g(x; f^* + \varepsilon) + \delta f^*(\rho x + [1-\rho]g(x; f^* + \varepsilon)) + \delta\varepsilon \\
& - \{g(x; f^*) + \delta f^*(\rho x + [1-\rho]g(x; f^*))\}.
\end{aligned}$$

First, I show that $l(x, f^*, \varepsilon) \leq l(\underline{\pi}, f^*, \varepsilon)$. I let $\Delta g(x) = g(x; f^* + \varepsilon) - g(x; f^*)$ and $\Delta f^*(x) = f^*(\rho x + [1-\rho]g(x; f^* + \varepsilon)) - f^*(\rho x + [1-\rho]g(x; f^*))$. Then, $l(x, f^*, \varepsilon)$ increases with $\Delta g(x)$ and $\Delta f^*(x)$. To show that $x = \underline{\pi}$ maximizes $l(x, f^*, \varepsilon)$, I show that, given ε, it maximizes both $\Delta g(x)$ and $\Delta f^*(x)$. The following two equations hold for $x \leq \gamma - \epsilon$.

$$\begin{aligned}
&\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(x; f^* + \varepsilon)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho}x \\
&= g(x' f^* + \varepsilon) + \delta f^*(\rho x + [1-\rho]g(x; f^* + \varepsilon)) + \delta\varepsilon \\
&\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(g(x; f^*)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho}x \\
&= g(x; f^*) + \delta f^*(\rho x + [1-\rho]g(x; f^*))
\end{aligned}$$

Combining these two equations yields

$$\begin{aligned}
Y(x, \Delta g(x)) \equiv& \left[1 - \frac{h\delta\rho}{1-\delta(1-\rho)}\right]\{b(g(x; f^*) + \Delta g(x)) - b\,(g(x; f^*))\} \\
& - \Delta g(x) - \delta\{f^*(\rho x + [1-\rho]g(x; f^*) + \rho\Delta g(x)) - f^*\,(\rho x + [1-\rho]g(x; f^*))\} = \delta\varepsilon
\end{aligned}$$

I let $\bar{x} = \text{argmax}_{x \in \{x \in X | Y(x, \Delta g(x) = \delta\varepsilon\}} \Delta g(x)$. Suppose $\bar{x} > \underline{\pi}$. Then, the convexity of b, the weak concavity of f^*, and $g(x)$ increasing in x implies $\delta\varepsilon = Y(\bar{x}, \Delta g(\bar{x})) > Y(\underline{\pi}, \Delta g(\bar{x}))$. $Y(\underline{\pi}, \Delta g(x))$ is convex in $\Delta g(x)$, and $Y(\underline{\pi}, 0) = 0$. Because of the convexity, there exits $\Delta g > \Delta g(\bar{x})$ such that $Y(\underline{\pi}, \Delta g) = \delta\varepsilon$. This contradicts $\bar{x} = \text{argmax}_{x \in \{x \in X | Y(x, \Delta g(x) = \delta\varepsilon\}} \Delta g(x)$. Thus, $\bar{x} = \underline{\pi}$. As for $\Delta f^*(x)$,

$$\Delta f^*(x) \leq f^*(\rho\underline{\pi} + [1-\rho]g(\underline{\pi}; f^*) + \rho\Delta g(\underline{\pi})) - f^*(\rho\underline{\pi} + [1-\rho]g(\underline{\pi}; f^*)) = \Delta f^*(\underline{\pi}),$$

following the weak concavity of f^*, $g(x; f^*)$ weakly increasing in x, and $\max \Delta g(x) = \Delta g(\underline{\pi})$. These results imply

$$\begin{aligned} \|Tf_n - Tf^*\| < \max_{x \in X} l(x, f^*, \varepsilon) &\leq l(\underline{\pi}, f^*, \varepsilon) \\ &= g(\underline{\pi}; f^* + \varepsilon) + \delta f^*(\rho\underline{\pi} + [1-\rho]g(\underline{\pi}; f^* + \varepsilon)) + \delta\varepsilon \\ &\quad - \{g(\underline{\pi}; f^*) + \delta f^*(\rho x + [1-\rho]g(\underline{\pi}; f^*))\}. \end{aligned} \tag{18}$$

Next, I show that $g(\underline{\pi}, f)$ is continuous in f at $f = f^*$. At $x = \underline{\pi}$, the following condition holds for small ε,

$$\begin{aligned} &\left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho}\right] b(\tilde{g}(\varepsilon)) + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho}\right]\underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho}\underline{\pi} \\ &= \tilde{g}(\varepsilon) + \delta f^*(\rho\underline{\pi} + [1-\rho]\tilde{g}(\varepsilon)) + \delta\varepsilon, \end{aligned} \tag{19}$$

where $\tilde{g}(\varepsilon)$ denotes $g(\underline{\pi}; f^* + \varepsilon)$. I rewrite this equation as $Z(\tilde{g}(\varepsilon)) = \varepsilon$ where $Z(x) = (1/\delta) \cdot \{[1 - \delta\eta(1-\rho)/(1-\delta\rho)]b(x) + \{\delta/(1-\delta) + \delta\eta/[1-\delta\rho]\}\underline{\pi} - \delta\eta\rho/(1-\delta\rho)\underline{\pi} - x - \delta f(\rho\underline{\pi} + [1-\rho]x)\}$. $Z(\cdot)$ is a continuous and convex function, and $Z(\tilde{g}(0)) = 0$, implying $\tilde{g}(\varepsilon)$ is strictly increasing in ε and continuous in ε.

Finally, I can prove that Tf_n converges to Tf^*. Since $\tilde{g}$ and f^* are both continuous, $\tilde{g}(\varepsilon) + \delta f^*([1-\rho]\underline{\pi} + \rho q(\varepsilon)) + \delta\varepsilon$ is continuous in ε. Subsequently, for any $\omega > 0$, there exists

$\zeta > 0$ such that for all $\varepsilon < \zeta$,

$$\|Tf_n - Tf^*\| < \tilde{g}(\varepsilon) + \delta f^*(\rho\underline{\pi} + [1-\rho]\tilde{g}(\varepsilon)) + \delta\varepsilon$$
$$- \{\tilde{g}(0) + \delta f^*(\rho\underline{\pi} + [1-\rho]\tilde{g}(0)) + \delta \cdot 0\} < \omega$$

Thus, Tf_n uniformly converges to Tf^*, implying f_n converging to Tf^*. Since f_n converges also to f^*, $f^* = Tf^*$. Since f^* is the solution to the functional equation whose value is bounded, $v^* = f^*$. Additionally, since $v^* = f^* \in C(X)$, $v^* = Tv^* \in C'(X)$. The policy function $g^*(x) = g(x; v^*)$ has the same properties as $g(x; f)$ for $f \in C(X)$. This completes the proof of Lemma 5.

E Proof of Lemma 6

I prove Lemma 6 in three steps, each of which follows a procedure similar to that in the proof of Lemma 5. First, I derive the properties of the value function $w_1^*(r)$ and the policy function $g_{w1}^*(r)$ of a modified problem that does not have the loss-evading incentive or the future cooperation losses: both $w_1^*(r)$ and $g_{w1}^*(r)$ are strictly increasing and strictly concave on X. Second, I set up another modified problem in which only the loss-evading incentive is absent. I derive the properties of the value function $w_2^*(r)$ and the policy function $g_{w2}^*(r)$, using the results of the first modified problem: $w_2^*(r) = w_1^*(r)$ and $g_{w2}^*(r) = g_{w1}^*(r)$ on $[\underline{\pi}, \bar{z}]$; $w_2^*(r) = \bar{z}/(1-\delta) - \eta(r - \bar{z})/(1 - \delta\rho)$ and $g_{w2}^*(r) = \bar{z}$. Finally, using the results from the second modified problem, I derive the properties of the value function and the policy function of the original problem (5) as in Lemma 6. See the Online Appendix for further details.

F Proof of Proposition 3

I define X and $C(X)$ as:

$$
\begin{aligned}
X :& [\underline{\pi}, \bar{\pi}] \\
C(X) :& \text{the set of bounded and continuous functions } f : X \to R \\
& \text{with the sup norm that are weakly concave on } [\bar{z}, \bar{\pi}]. \\
& s.t.\ \forall x \in [\underline{\pi}, \bar{z}],\ f(x) = w_1^*(x), \\
& \quad \forall x \in [\bar{z}, \bar{\pi}],\ f(x) \leq \frac{\bar{z}}{1-\delta} - \frac{\eta}{1-\delta\rho}(x - \bar{z}), \\
& \quad \forall x, \forall y \in [\bar{z}, \bar{\pi}]\ s.t.\ x < y,\ \frac{f(y) - f(x)}{y - x} \geq -\frac{1+\eta}{\delta(1-\rho)},
\end{aligned}
$$

where w_1^* is that of Appendix E. On $C(X)$, I define the operator T by

$$
Tf(x) = \max_{y \in \Gamma(x,f)} \left[y - \mathbb{1}\{x > y\} \cdot \eta(x - y) + \delta f(\rho x + [1-\rho]y) \right],
$$

where

$$
\begin{aligned}
\Gamma(x, f) = \Bigg\{ x \in [\underline{\pi}, \bar{\pi}] : & \left[1 - \frac{\delta\eta(1-\rho)}{1-\delta\rho} \right] b(y) + \mathbb{1}\{x > y\} \cdot \eta \left[\min\{x, b(y)\} - y \right] \\
& + \left[\frac{\delta}{1-\delta} + \frac{\delta\eta}{1-\delta\rho} \right] \underline{\pi} - \frac{\delta\eta\rho}{1-\delta\rho} x \leq y + \delta f(\rho x + [1-\rho]y) \Bigg\}.
\end{aligned}
$$

Following steps similar to those in the proof of Lemma 5 completes the proof. See the Online Appendix for further details.