Cartel Formation under Manager Loss Aversion

Douglas C. Turner

December 26, 2025

Abstract

Cartels often emerge after a deterioration in market conditions, such as a decline in demand, an increase in costs, or the entry of a new competitor. However, in conventional theoretical models of collusion, such changes either reduce or do not impact incentives to collude. I show how deteriorations in market conditions can facilitate collusion and cause the formation of new cartels when, contrary to standard models, colluding managers are averse to losses. Additionally, a deterioration in market conditions can increase the maximum possible collusive payoff and permit colluding managers to set higher prices than would otherwise be sustainable.

Keywords: collusion, loss aversion, prospect theory, managerial incentives

JEL Codes: D91, L41, L21

Department of Economics, University of Florida, PO Box 117140, Gainesville, FL 32611 USA (douglasturner@ufl.edu). I thank Yonggyun Kim, Iwan Bos and David Sappington for helpful comments. I thank Daniel Garcia, Randy Koliha, and Isaiah Sarria for excellent research assistance.

1 Introduction

Cartels often form after a deterioration in market conditions, such as a reduction in demand, an increase in marginal cost or the entry of a new competitor.[1] This empirical fact stands in contrast to the predictions of conventional models of collusion. These models predict that the entry of a new competitor reduces incentives to collude. Additionally, permanent reductions in demand or increases in marginal cost either do not impact the sustainability of collusion or reduce managers' incentives to collude. Crucially, standard models of collusion assume that colluding managers are loss neutral (i.e., they do not perceive losses as more severe than an equivalent gain). In practice, managers may be averse to losses (Kahneman and Tversky, 1979). I present a model of collusion among loss averse managers and find that, consistent with empirical evidence, deteriorations in the profitability of a market can cause the formation of cartels[2] in previously competitive markets, increase the payoff managers earn from collusion, and cause higher prices.

To understand how a deterioration in market conditions can cause the formation of a cartel, consider a competitive industry in which each firm employs a loss averse manager. Additionally, suppose a manager's utility depends on the amount of profit the manager earns for his/her firm. Over time, managers become accustomed to competitive profit/utility levels. Thus, when market conditions unexpectedly deteriorate in a way that reduces the amount of profit that can be earned from competitive play, managers perceive continued competition as a loss in utility. This effect causes managers to turn to collusion for two reasons. First, collusion, which raises profits above Nash equilibrium levels, avoids a painful loss in utility. Second, a deterioration in market conditions relaxes the incentive compatibility constraints necessary for successful collusion. Because competitive profit levels (after the deterioration in market state) are perceived as a loss, the manager experiences a painful loss in utility when the cartel dissolves and competition resumes. As a result, managers have relatively weak incentives to cheat on the cartel because doing so would cause a breakdown in collusion, a return to competition, and significant losses in utility. This effect does not occur after improvements in market conditions (e.g., an increase in demand or a decline in marginal cost), because managers perceive post-change Nash profit levels as a gain, not a loss.

In addition to promoting cartel formation, deteriorations in market conditions can, due to the considerations outlined above, also increase the collusive payoff (i.e., the discounted present value of utility from collusion) and the gain in utility from collusion (i.e., the difference between collusive and competitive utility levels). Additionally, by enhancing the stability of collusion, a deterioration in market state can allow a cartel to set higher collusive prices than would otherwise be incentive compatible/sustainable, particularly

[1] See Aston and Pressey (2012), Herold and Paha (2018), and the empirical findings presented in Section 2.

[2] I refer to a group of colluding managers as a cartel throughout the analysis for ease of exposition. However, the model does not require collusion to be explicit/illegal.

in early periods of collusion.

Shalev (1998) and Bernard (2011) both study repeated games with loss averse agents. However, neither study analyzes the impact of deteriorations in market conditions on incentives to collude. Rotemberg and Saloner (1986) and Green and Porter (1984) study the impact of temporary profit shocks on firms' ability to collude. Motivated by empirical evidence suggesting that many cartels form following a permanent (or long lasting) change in market conditions rather than a transitory shock, I analyze the impact of permanent changes in market state rather than temporary profit shocks.

Most closely related to the current study is a series of insightful articles by Spagnolo (1999; 2000; 2005) which explore the impact of managerial incentives and compensation structures on managers' ability to collude. Spagnolo (2005) demonstrates that collusion is easier to sustain when managers have a preference for a smooth stream of income. Preferences for smooth income could arise from, among other factors, a psychological aversion to losses or a tendency to evaluate managerial performance relative to past levels. Thus, the fact that loss aversion can facilitate collusion was first established, more generally, in Proposition 1 of Spagnolo (2005). The present study focuses not on the potential for loss aversion to facilitate collusion, but on how a deterioration in market conditions can facilitate collusion when managers are loss averse.[3] Additionally, unlike the present study, Spagnolo (2005)'s model is not tailored specifically to the analysis of loss aversion and therefore does not explicitly model certain aspects of loss aversion such as reference point dynamics.

Section 2 presents empirical evidence that cartels often form following a deterioration in market conditions. Section 3 introduces a stylized model intended to illustrate the key mechanisms as simply as possible. Section 4 introduces the general model. Cartel formation is analyzed in Section 5. The value of collusion and gain from collusion are studied in Section 6. Section 7 presents results related to cartel pricing, and Section 8 concludes. The appendix and technical appendix contain proofs, as well as additional derivations, discussion, and simulation results.

2 Empirical Evidence

European Commission (EC) cartel decision documents typically contain detailed information regarding the relevant industry, the cartel's practices and structure, and the origins of the infringement. To identify which changes in market conditions lead to cartel formation, I review European Commission cartel decisions and record the stated causes of each cartel's formation. Specifically, I search for changes in the market environment that are alleged to have caused producers to engage in price fixing. I restrict attention to cases

[3] Spagnolo (2005) also analyzes collusive pricing under demand shocks/fluctuations in the spirit of Rotemberg and Saloner (1986).

with an available prohibition decision (in english) originally published between 1998 and 2024, which results in 92 cartel cases. See Technical Appendix B for additional details regarding data collection.

Table 1 presents results. Panel A of Table 1 reports that 45 of the 92 decisions reviewed did not mention a clear cause of cartel formation. I record any causes of cartel formation indicated within the decisions. Thus, there are multiple recorded causes for certain cartels.[4] Panel B of Table 1 presents results pertaining to improvements in market conditions. An improvement in market conditions of some kind (e.g., a cost reduction, demand increase or the exit of a competitor) is cited as a factor causing the cartel's formation in 5 cases, constituting 11% of all cases for which at least one cause is available. Panel C of Table 1 presents results pertaining to deteriorations in market conditions. Five different types of deteriorations in market conditions were cited within the decisions including an increase in cost, a decrease in demand, the entry of a new competitor, an increase in buyer power, and an increase in import competition. The most common deterioration in market conditions was a cost increase which was cited as a factor contributing to the cartel's formation in 12 cases (26% of all cases with at least one cause). The second most common deterioration in market conditions was the entry of a new competitor which was cited as a factor contributing to the cartel's formation in 8 cases (17% of all cases with at least one cause). In total, at least one deterioration in market conditions was cited as a cause in 27 cases (57% of all cases with at least one cause). Panel D presents results pertaining to other causes of cartel formation including overcapacity (7 cases), a regulatory policy change (3 cases), a price war (7 cases), and low prices (9 cases).

These observations suggest that deteriorations in market state are significantly more likely to be cited as the cause of a cartel's formation than an improvement in market conditions. This finding is consistent with prior literature (Herold and Paha, 2018; Grout and Sonderegger, 2005, 2007; Aston and Pressey, 2012; Levenstein and Suslow, 2015) which has previously identified various deteriorations in market state as potential causes of cartel formation. Most closely related to the analysis above is Herold and Paha (2018) who also review EC cartel decisions, augmenting their review with information from external sources, and record events potentially contributing to each cartel's formation. They find that a reduction in demand or the entry of a new competitor often precedes a cartel's formation. For example, 16 of the 41 cartels included in their sample were formed following a reduction demand while only 8 of 41 were formed after an increase in demand. Aston and Pressey (2012) also review EU cartel cases and find similar results. Specifically, they find that market conditions declined or were declining prior to cartel formation in 24 cases while conditions were improving in only 12 cases. Reflecting this pattern, the European Commission frequently notes within

[4]Cartel members often argue that their industry was in crisis at the time of the cartel's formation in order to request a reduction in the fine imposed on them by the European Commission. In 30% of cases, either the infringing firms or the EC contended that the industry was in a state of crisis when the cartel initially formed.

its cartel decisions that "[a]s a general rule, cartels come into being when a sector encounters problems."[5]

The empirical patterns above suggest that deteriorations in market conditions, such as a reduction in demand, an increase in marginal cost or the entry of a new competitor, may cause the formation of a cartel. However, in conventional theoretical models of collusion, these changes typically hinder or have no impact on the sustainability of collusion. In standard models of collusion, the entry of a new competitor reduces the sustainability of collusion (i.e., the critical discount factor is increasing in the number of firms (Ivaldi et al., 2003)). An increase in marginal cost or a reduction in demand typically does not impact the sustainability of collusion (Gallice, 2010; Klein and Schinkel, 2019). Thus, standard models of collusion typically cannot explain the empirical patterns established above (i.e., that cartels tend to form after deteriorations in market conditions). In the following sections, I present a model which illustrates how accounting for colluding managers' aversion to losses can explain this puzzle.

[5] See, for example, COMP/E-1/38.069 – Copper Plumbing Tubes, 9/3/2004, Commission Decision (¶742).

Table 1: Causes of Cartel Formation from EC Decisions

	Num Cases	% of Cases	% of Cases w/ Causes
Panel A: Data Collection Statistics			
Total Num. Cases	92	100%	
No Cause Cited	45	49%	
At Least One Cause Cited	47	51%	
One Cause Cited	35	38%	
More than One Cause Cited	12	13%	
Crisis Mentioned	28	30%	
Panel B: Improvements in Market Cond.			
Cost Dec.	1	1%	2%
Demand Inc.	3	3%	6%
Exit	1	1%	2%
Improvement Cited	5	5%	11%
Panel C: Deteriorations in Market Cond.			
Cost Inc.	12	13%	26%
Demand Dec.	6	7%	13%
Entry	8	9%	17%
Inc. Buyer Power	5	5%	11%
Inc. Import Comp.	2	2%	4%
Deterioration Cited	27	29%	57%
Panel D: Other Causes			
Overcapacity	7	8%	15%
Regulatory Policy Change	3	3%	6%
Price War	7	8%	15%
Low Prices	9	10%	19%

Notes: This table presents results from a review of European Commission cartel decisions from 1998 to 2024.

3 A Simple Illustrative Model

This section introduces a Cournot duopoly model with loss-averse managers to provide a simple illustration of how a deterioration in market conditions can strengthen incentives to collude. Two firms interact in an infinitely repeated quantity setting game. Let q_i denote the quantity chosen by the manager of firm $i \in \{1, 2\}$, and let total output be $Q = q_1 + q_2$. Marginal costs are normalized to zero. Prior to the initial period, the market experiences an unexpected, permanent decline in demand. Before the change, inverse demand was $P_b(Q) = A_b - Q$. After the decline, inverse demand becomes $P_a(Q) = A_a - Q$ where $A_a < A_b$. "b" denotes before and "a" denotes after. Managers have a common discount factor $\delta \in (0, 1)$. Before the decline, managers played the Nash equilibrium and earned per-period profits of $\frac{A_b^2}{9}$. Because managers have been accustomed to this level of profit, they treat it as a benchmark against which they evaluate outcomes: any profit level above $\frac{A_b^2}{9}$ is experienced as a gain, and any profit level below it as a loss. This benchmark is common to all managers, remains fixed over time, and is referred to as a reference point.

Manager utility is

$$u(\pi) = \begin{cases} \pi - l(r - \pi) & \text{if } \pi < r \\ \pi & \text{if } \pi \geq r \end{cases}$$

where r is the reference point and $l \geq 0$ is the degree of loss aversion. When $l = 0$, managers are loss neutral. For simplicity, suppose managers collude by choosing the joint profit maximizing quantity and enforcing cooperation through grim-trigger strategies.[6] Grim trigger strategies involve all managers producing the collusive output level in all periods. If any manager deviates and produces any quantity other than the collusive production level, all managers produce Nash equilibrium output in all future periods. Thus, defections from collusion are punished with an infinite reversion to Nash competition.

Collusion is sustainable (i.e., no manager wishes to deviate) only when managers are sufficiently patient. Formally, managers' discount factors must exceed a threshold level known as a critical discount factor. The critical discount factor before (after) the decline in demand is denoted $\tilde{\delta}_b$ ($\tilde{\delta}_a$). The lower the critical discount factor, the wider the range of discount factors for which collusion is sustainable. Before the decline in demand, collusion is sustainable if

$$\frac{1}{1-\delta} \underbrace{\left[\frac{A_b^2}{8}\right]}_{\text{Collusive Utility}} \geq \underbrace{\frac{9A_b^2}{64}}_{\text{Defection Utility}} + \frac{\delta}{1-\delta} \underbrace{\left[\frac{A_b^2}{9}\right]}_{\text{Punishment Utility}}. \tag{1}$$

[6] To ensure that collusive profits remain above the reference point, assume the decline in demand is moderate: $A_a > \frac{\sqrt{8}}{3} A_b$. If the decline is too large, even the monopoly profit level after the decline is below the reference point, and the effect detailed below is diminished. See discussion surrounding Assumption 4 in Section 4 for details.

Solving (1) for δ yields the critical discount factor $\tilde{\delta}_b = \frac{9}{17}$. After the decline in demand, collusion is sustainable if

$$\underbrace{\frac{1}{1-\delta}\left[\frac{A_a^2}{8}\right]}_{\text{Collusive Utility}} \geq \underbrace{\frac{9A_a^2}{64}}_{\text{Defection Utility}} + \underbrace{\frac{\delta}{1-\delta}\left[\frac{A_a^2}{9} - l\left(r - \frac{A_a^2}{9}\right)\right]}_{\text{Punishment Utility}}. \tag{2}$$

Because managers' reference points are fixed at the level of pre-decline competitive profits, the reduction in demand causes competitive profits to fall below r. Thus, a reversion to Nash competition is perceived as a loss after the change (see the punishment utility in (2)). Solving (1) for δ yields the critical discount factor

$$\tilde{\delta}_a = \frac{9}{17 + 64\,l\left(\frac{A_b^2}{A_a^2} - 1\right)}.$$

When $l = 0$ (as in conventional models of collusion), $\tilde{\delta}_a = \tilde{\delta}_b$ and a decline in demand has no impact on the sustainability of collusion. When $l > 0$, $\tilde{\delta}_a < \tilde{\delta}_b$ (recall that $\frac{A_b^2}{A_a^2} > 1$ as $A_b > A_a$). Therefore, a decline in demand reduces the critical discount factor and thereby enhances incentives to collude.

To understand this result, suppose a cartel forms after the decline in demand and a manager subsequently considers cheating on the collusive agreement. Cheating would trigger a breakdown of the cartel and a return to competitive play. But because post-decline competitive profits now lie below the manager's reference point, that punishment phase is perceived as a loss. The prospect of incurring this painful loss in utility makes cheating far less attractive, thereby stabilizing collusion. The next section presents a more general model in which reference points can evolve dynamically with experienced utilities, firms may choose prices, quantities or another strategic variable, and profit and utility functions need not take the simple forms used above. This broader framework yields additional results and establishes the mechanism illustrated here in greater generality.

4 A General Model

Consider a market consisting of N symmetric firms interacting in each of infinitely many periods indexed by $t = 1, 2, 3\ldots$ Each firm employs a manager responsible for overseeing its operations in the market. Specifically, in every period t, the manager chooses a strategic variable x_t (e.g., price or output level) on behalf of the firm. For ease of exposition, I refer to the strategic variable x as the firm's price throughout the ensuing analysis, recognizing that x could represent any variable chosen by managers that impacts a firm's profit. Similarly, I refer to a sequence $\{x_t\}_{t=1}^{\infty}$ as a price path. $x_t \in \Omega$ for all $t \in \{1, 2\ldots\}$ where $\Omega \subset \mathbb{R}$ is a compact set. Managers have a common discount factor $\delta \in (0, 1)$ and seek to maximize the discounted present value of their utility.

Managers are averse to losses. Loss aversion arises when an individual is more sensitive to losses than to gains of an equivalent magnitude (Barberis, 2013). Managers may display a distaste for losses due to a psychological aversion to losses (termed psychological loss aversion) or due to the negative career and wage consequences of a drop in performance below expected or target levels (termed target-based loss aversion). First, consider psychological loss aversion. Psychological loss aversion is an intrinsic cognitive aversion to utility levels below a pre-determined reference point. A substantial experimental and empirical literature has found evidence of psychological loss aversion.[7]

While manager loss aversion may be the result of an inherent, psychological distaste for losses, managers may also display an aversion to losses in profit or utility due to more practical considerations. Specifically, managers may be averse to losses if there are negative career or compensational consequences of performance below a predetermined target level (Sullivan and Kida, 1995; Crum, Laughhunn and Payne, 1981), termed target-based loss aversion. The utility functions employed in this study are consistent with both psychological and target-based loss aversion.[8]

Whether a particular utility level is perceived as a loss or a gain depends on a manager's reference point. Utility levels above (below) the reference point are perceived as gains (losses) by managers. When a manager earns a profit of π, their utility is $u(\pi; r, l)$ where r denotes the reference point and l denotes the degree of loss aversion. Larger values of l are associated with a stronger distaste for losses. $l = 0$ corresponds to a loss neutral manager who does not display an aversion to losses.

A manager's utility consists of two components: base utility and a separate term capturing managers' aversion to losses. A manager's base utility is denoted $u(\pi)$ where π is the profit a manager generates for the firm (through the manager's choice of x) in a given period. Following Shalev (2000) and Tversky and Kahneman (1991), manager utility satisfies the following assumption.

Assumption 1. *$u(\pi; r, l) \equiv u(\pi) - lL(r - u(\pi))$ where*

i) $u(\pi)$ is continuous for all $\pi \in \mathbb{R}$,

ii) $u(\pi)$ is strictly increasing for all $\pi \in \mathbb{R}$,

iii) $l \geq 0$, and

[7] For reviews, see Novemsky and Kahneman (2005), Zank (2010), and Tversky and Kahneman (1991). While loss aversion is often documented experimentally in laboratory studies involving college students; more experienced market participants, such as managers, are not immune to loss aversion. Prior literature has found that experienced corporate managers, in particular, exhibit loss aversion (e.g., executives in manufacturing companies (Sullivan and Kida, 1995), major league baseball managers (Pedace and Smith, 2013), professional futures and options pit traders (Haigh and List, 2005), managers in investment banks (Willman et al., 2002)). Additionally, loss aversion is important in competitive environments (Gill and Prowse, 2012) and strategic games (Feltovich, 2011).

[8] In Technical Appendix D.5, I develop a two models capturing target-based loss aversion: a model wherein a bonus is awarded if a manager meets a pre-defined performance target and a model wherein a manager may be terminated if their performance fails to meet a target. These models illustrate how assessing manager performance relative to a pre-defined target can generate a loss averse utility function satisfying Assumption 1. Thus, the utility functions in the main text may represent either psychological or target-based loss aversion.

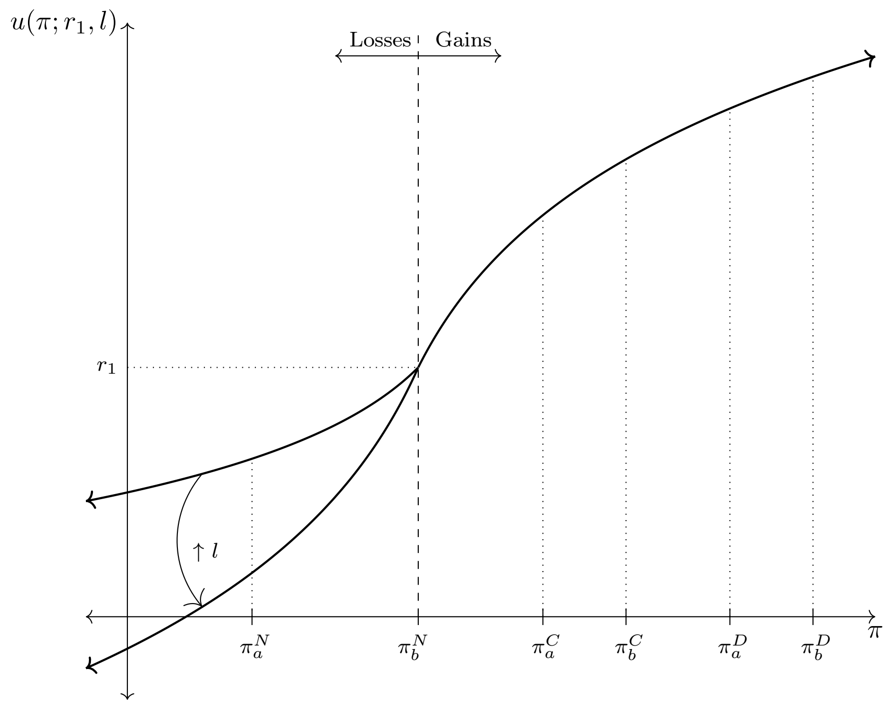


Figure 1: A loss averse utility function where π_i^C denotes collusive profits, π_i^D denotes defection profits, and π_i^N denotes Nash profit under regime $i \in \{a, b\}$.

iv) $L(x) : \mathbb{R} \to [0, \infty)$ *is continuous, strictly increasing for* $x > 0$*, and satisfies* $L(x) = 0$ *for* $x \leq 0$*.*

Assumption 1 permits a wide range of utility specifications including utility functions with a kink at the reference point and S-shaped utility functions.[9] For example, $u(\pi) = \pi$ and $L(x) = x$ result in linear loss aversion, also known as a kinked linear utility function (Shalev, 2000; Maggi, 2004).

Reference points divide utility levels into gains and losses. Reference points are defined differently under psychological and target-based loss aversion. First, consider psychological loss aversion. Reference points under psychological loss aversion are based on past utility levels. Intuitively, a manager who repeatedly experiences a particular level of utility over a significant period of time is likely to become accustomed to this utility level and perceive any greater (lower) utility levels as a gain (loss). Under target-based loss aversion, reference points are determined exogenously by performance targets (Sullivan and Kida, 1995; Crum, Laughhunn and Payne, 1981). These targets may be set explicitly by higher level managers (or shareholders) in the terms of a manager's compensation contract.

Reference points, under both types of loss aversion, may evolve over time in response to experienced utility. Following Bowman, Minehart and Rabin (1999), Ryder Jr and Heal (1973), and Karlsson, Loewenstein and Seppi (2009), a manager's reference point in period t is $r_t = m(r_{t-1}, u_{t-1}) = \alpha r_{t-1} + (1 - \alpha) u_{t-1}$ where

[9]See Technical Appendix D.2 for details.

r_{t-1} is the reference point in the previous period, u_{t-1} is utility in the previous period, and $\alpha \in [0,1]$ is a parameter governing the rate of reference point adjustment.

Managers' incentives to collude depend not only on how reference points evolve over time but also on managers' expectations regarding how their own reference point will adjust in the future. I consider two approaches to modeling managers' expectations regarding the evolution of their reference points: naive and sophisticated managers. Naive managers evaluate future utilities relative to their current reference point. In contrast, sophisticated managers fully anticipate future changes in their reference point when making decisions. The main results of this study hold, under appropriate assumptions, for both naive and sophisticated managers. In the main text, I restrict attention to the case of sophisticated managers. Naive managers are analyzed in Technical Appendix D.1.

As illustrated in Section 2, cartels often form after the profitability of a market declines. Formally, a parameter in firms' profit functions has changed in a way that negatively impacts profitability. In the ensuing analysis, this parameter is referred to as the market state. The market environment after the deterioration in market state is referred to as regime a. The market environment absent the deterioration in market state is referred to as regime b.

Let π_i^N denote Nash equilibrium profit under regime $i \in \{a, b\}$. The following assumption characterizes Nash equilibrium play under both regimes.

Assumption 2. *For $i \in \{a, b\}$, there exists a unique symmetric Nash equilibrium wherein all managers play $x_i^N \in \Omega$ and earn a profit of π_i^N.*

If a cartel forms, managers collude by setting a common price in each period t. Let $\pi_i(x)$ denote per-firm profit during collusion under regime i when all firms charge a common price x. $\pi_i^D(x)$ denotes the profit a manager earns if they choose to defect from collusion when the collusive level of the choice variable is x under regime i. $\pi_i(x)$ and $\pi_i^D(x)$ satisfy the following restrictions.

Assumption 3. *For $i \in \{a, b\}$, $\pi_i(x)$ and $\pi_i^D(x)$ satisfy the following assumptions:*

i) $\pi_i(x)$ and $\pi_i^D(x)$ are continuous in x for all $x \in \Omega$,

ii) $x_i^M \equiv argmax_{x\in\Omega}\pi_i(x)$ exists and is unique, and

iii) $\pi_i^D(x) \geq \pi_i(x)$ for all $x \in \Omega$.

Assumption 3(ii) ensures that a unique monopoly price x_i^M exists under both regimes. Additionally, let $\pi_i^M \equiv \pi_i(x_i^M)$ denote monopoly profit. $\pi_i(x)$ includes any fixed costs of collusion (Thomadsen and Rhee, 2007; Colombo, 2013; Klein and Schinkel, 2019).[10] Assumption 3(iii) ensures that defection does not reduce

[10]Fixed costs of collusion include any moral dis-utilities from participating in an illegal activity, fixed costs of monitoring rivals (e.g., payments made to a third party tasked with monitoring compliance with the collusive scheme), fixed costs involved

a manager's profit. A broad variety of common oligopoly models satisfy Assumptions 2 and 3, including homogenous product Cournot competition, differentiated product Bertrand competition, and differentiated product Cournot competition.

The timing of the game is as follows. The market state unexpectedly and permanently deteriorates immediately prior to the initial period. In the beginning of the initial period ($t = 1$), managers decide whether to form or not form a cartel. If a cartel does not form, managers engage in Nash competition in all future periods. If a cartel forms, then managers jointly set a common collusive price in each period. The remainder of the initial period ($t = 1$) and each subsequent period consists of three phases. In the first phase, managers determine x_t either collusively or competitively (depending on whether a cartel has formed or not). In the second phase, managers receive profits/utilities. In the third phase, managers' reference points are updated, in accordance with the utilities experienced in phase 2.

Let r_t represent the common reference point of managers in period t. The following assumption governs managers' initial reference point r_1 and characterizes the market state after the change (i.e., regime a).

Assumption 4. *i)* $r_1 = u(\pi_b^N)$,

ii) $r_1 < u(\pi_a^M)$, *and*

iii) $\pi_a^N < \pi_b^N$.

Assumption 4(i) states that the initial reference point is the utility level from Nash competition under regime b. Assumption 4(i) is consistent with a setting where Nash competition has prevailed in the past and, as a result, managers have become accustomed to the Nash equilibrium utility level. Assumption 4(i) is made primarily for concreteness and ease of exposition in the main text. In Technical Appendix C, I show that this assumption can be relaxed to permit initial reference points above and below $u(\pi_b^N)$.

In light of Assumption 4(i), Assumption 4(ii) requires that the deterioration in market conditions is sufficiently moderate that monopoly profits after the change exceed Nash equilibrium profits prior to the change. In the results to follow, managers have an incentive to form a cartel after a deterioration in market state in order to avoid losses in utility. If the change in market state is sufficiently severe that losses cannot be avoided by colluding (even if such collusion generates the maximum possible profits), then this effect will not occur. Assumption 4(iii) restricts attention to changes in market state that reduce the profitability of Nash competition, such as those discussed in Section 2.

If a cartel forms, collusion is sustained through grim trigger strategies. Specifically, collusion continues if all firms charge the agreed upon price. If any firm does not charge the agreed price, then collusion breaks down and all managers set prices competitively.

in concealing collusive activities (including managerial effort) and communicating with other managers involved in the cartel, and costs of buying out potential entrants (Ganslandt, Persson and Vasconcelos, 2012).

Assumption 5. *Reference points do not update (i.e., $m(r, u) = r$) in response to utilities experienced during the defection period.*

Assumption 5 reflects two considerations. First, the structure of grim trigger strategies implies that relatively large profit/utility levels earned during defection are transient as punishment phases immediately follow defection. Moreover, managers recognize that defection will result in elevated profits/utilities for a single period and therefore may not adjust to utility levels experienced during defection as they would to utility levels they believe may continue in future periods. As Chen and Rao (2002) write, "[i]f people are aware that a second event is going to undo the first, the reference point will likely not shift after the first event." Second, the adjustment of reference points to utilities in the defection period can result in unrealistic outcomes. Specifically, managers may be hesitant to defect as they would anticipate becoming acclimatized to relatively high utilities during defection and, as a result, perceive the punishment phase as a particularly large loss in utility. Declining the possibility of obtaining large profits due to a fear of becoming acclimatized to the resulting increase in utility may be unrealistic.

$V_i^N(r)$ denotes the discounted present value of utility from repeated Nash equilibrium play in all periods (hereafter, the competitive payoff) when the current reference point is r under regime $i \in \{a, b\}$. Thus,

$$V_i^N(r) = u(\pi_i^N; r) + \delta V_i^N(m(r, u(\pi_i^N))).$$

If a cartel does not form, all managers earn payoff $V_i^N \equiv V_i^N(r_1)$ where r_1 is the initial reference point. For collusion to occur successfully, collusive prices must be such that no manager wishes to defect or cheat on collusion in any period. Suppose $\{x_t\}_{t=1}^{\infty}$ denotes a collusive price path and let $\{r_t\}_{t=1}^{\infty}$ denote the corresponding sequence of reference points (i.e., $r_t = m(r_{t-1}, u(\pi_i(x_{t-1})))$ for $t > 1$). No manager wishes to defect in period T if

$$\sum_{t=T}^{\infty} \delta^{t-T} u(\pi_i(x_t); r_t) \geq u(\pi_i^D(x_T); r_T) + \delta V_i^N(r_T). \tag{3}$$

Inequality (3) is hereafter referred to as the incentive compatibility constraint (ICC) in period T.

During collusion, managers set prices to maximize their discounted present value of utility subject to the constraint that no manager wishes to defect in the present period or any future period (the ICC is satisfied for all T). V_i^C denotes the discounted present value of utility from collusion in the initial period under regime i. Thus, if a cartel forms, each manager earns a payoff

$$V_i^C = \max_{\{x_t\}_{t=1}^{\infty} \in \Psi_i} \sum_{t=1}^{\infty} \delta^{t-1} u(\pi_i(x_t); r_t) \tag{4}$$

and

$$\Psi_i = \left\{ \{x_t\}_{t=1}^{\infty} : x_t \in \Omega \text{ and } \sum_{t=T}^{\infty} \delta^{t-T} u(\pi_i(x_t); r_t) \geq u(\pi_i^D(x_T); r_T) + \delta V_i^N(r_T) \quad \text{for all } T \in \{1, 2 \ldots\} \right\} \tag{5}$$

where, in each of the inequalities, r_t denotes the reference point in period t consistent with the price path $\{x_t\}_{t=1}^{\infty}$. Ψ_i is the set of price paths $\{x_t\}_{t=1}^{\infty}$ that are incentive compatible in all periods. Let $V_i^C = -\infty$ when $\Psi_i = \emptyset$ (i.e., when no price path satisfies the ICCs).

A cartel forms if the discounted present value of utility from collusion (hereafter, the "collusive payoff") is greater than the discounted present value of utility from Nash competition (i.e., $V_i^C > V_i^N$). This condition involves two requirements. First, there must exist a price path that satisfies the inequalities in Equation (5) (i.e., $\Psi_i \neq \emptyset$). Second, given an incentive compatible price path exists, the optimal price path must result in a collusive payoff which is greater than the competitive payoff.[11] If $V_i^C < V_i^N$, then a cartel does not form and managers set Nash equilibrium prices in all periods and earn a payoff of V_i^N. A solution to (4) is referred to as an optimal price path. Additionally, let V_i^M denote the discounted present value of utility, under regime i, if the cartel charges the monopoly price in all periods (hereafter, the monopoly payoff).

The following lemma establishes that an optimal price path exists.

Lemma 1. *If $\Psi_i \neq \emptyset$, then a solution to (4) exists under regime $i \in \{a, b\}$.*

5 Cartel Formation

In this section, I analyze the impact of a deterioration in market conditions on loss averse managers' incentives to form a cartel. The following condition is assumed to hold throughout this section.

Condition 1. $V_b^C(l) \leq V_b^N$

Condition 1 ensures that a cartel does not form under regime b (i.e., a cartel does not form absent a change in market conditions). This condition holds when, for example, the discount factor is sufficiently low or fixed costs of collusion are sufficiently large.[12]

Proposition 1. *Suppose Condition 1 holds. There exists an $\bar{l}$ such that $V_a^C(l) > V_a^N(l)$ when $l \geq \bar{l}$.*

Proposition 1 implies that, when Condition 1 holds and managers are sufficiently loss averse, a deterioration in market conditions will cause a cartel to form in an otherwise competitive market.

[11] In Technical Appendix D.3, I show that the existence of an incentive compatible price path (i.e., $\Psi \neq \emptyset$) does not necessarily imply that the collusive payoff exceeds the payoff from Nash competition.

[12] Note that $V_b^C(l)$, and thus Condition 1, does not depend on l when reference points are constant over time (i.e., $m(r, u) = r$). See Figure 2.

To understand this result, recall that, in the initial period when managers decide whether to form or not form a cartel, managers are accustomed to competitive utility levels (i.e., $r_1 = u(\pi_b^N)$ by Assumption 4(i)). When the market state unexpectedly deteriorates in a way that reduces the profit that can be earned through competitive play (Assumption 4(iii)), managers' perception of continued competition changes. Specifically, continued competition is now perceived as a loss because the utilities that can be earned through competitive play are less than managers' reference points. When managers are particularly averse to losses, the incentive to avoid this painful loss in utility causes managers to form a cartel which, by Assumption 4(ii), can generate utility levels that exceed the reference point and are therefore perceived as a gain, not a loss.

Cartel activity becomes increasingly attractive to managers after a deterioration in the market environment for another, more subtle, reason. Under grim trigger strategies, defection from the cartel is punished through a reversion to competitive play in all subsequent periods. However, as discussed previously, managers perceive competitive utility levels as a loss after the deterioration in market conditions. As a result, managers have relatively weak incentives to cheat on the cartel and endure relatively large losses in utility during the subsequent punishment phase. This effect relaxes the incentive compatibility constraints, stabilizes collusion, and enables managers to successfully collude after a deterioration in market state.

Figure 1 illustrates managers' incentives to collude in the initial period. Specifically, Figure 1 depicts a classical S-shaped utility function indicating competitive, collusive and defection profits under both regimes. Figure 1 illustrates how a deterioration in market conditions reduces competitive profit levels and, as a result, causes continued competition to be perceived as a loss. Crucially, collusive profits represent a gain. Thus, managers can avoid painful losses in utility from continued competition by turning to collusion.

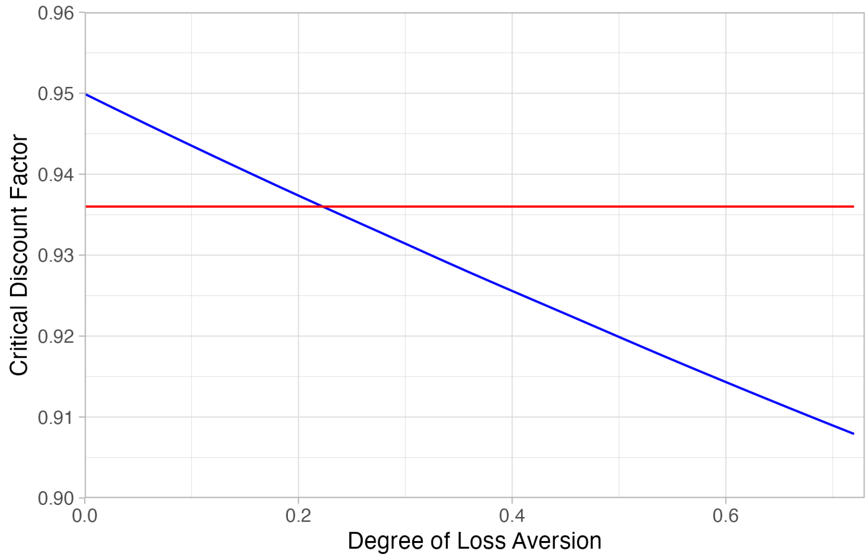


Figure 2: Critical Discount Factor by Degree of Loss Aversion Before (Red) and After (Blue) the Entry of a New Firm.

Notes: This figure depicts the critical discount factor as a function of the degree of loss aversion. See Technical Appendix G.1 for additional details regarding the simulations employed to generate this figure. Parameters: $a = 100$, $b = 2$, $e = 1$, $c = 0$, $\alpha = 1$ and $F = 125$. The blue curve denotes the critical discount factor after entry (i.e., $N = 6$) and the red curve denotes the critical discount factor absent entry (i.e., $N = 5$).

Figure 2 depicts the critical discount factor, the smallest discount factor for which a cartel forms, before (red) and after (blue) the entry of a new firm. See Technical Appendix G.1 for additional details regarding the simulations conducted to generate this figure and other figures presented in the main text.[13] First, observe that the entry of a new firm, as expected, increases the critical discount factor when managers are loss neutral (i.e., $l = 0$). This finding reflects the standard result that the sustainability of collusion in a repeated game is declining in the number of firms. However, when managers are sufficiently loss averse, the entry of a new firm can, counterintuitively, reduce the critical discount factor and enhance incentives to collude. Thus, when managers are loss averse, there exists a range of discount factors under which the entry of a new firm causes the formation of a cartel in a previously competitive market.

[13] All figures in the main text reflect outcomes from a setting involving N firms choosing prices and selling symmetrically differentiated products where the representative consumer has a utility function of (Singh and Vives, 1984)

$$U(q_1, \dots q_N) = a \sum_{i=1}^{N} q_i - \left(\frac{1}{2}\right) \left(b \sum_{i=1}^{N} q_i^2 + e \sum_{i=1}^{N} \sum_{j \neq i} q_i q_j \right)$$

where $a > 0$, $b > e > 0$ and q_i is the quantity of firm i's product consumed. Managers have a loss function of $L(x) = x$ and linear base utility $u(x) = x$. Additionally, there is a fixed cost of collusion $F \geq 0$.

6 Collusive Payoff and Gain from Collusion

Section 5 establishes that a deterioration in market state may cause a cartel to form in an otherwise competitive market. However, Proposition 1 does not speak to the magnitude of the payoff managers obtain from forming a cartel. In the model analyzed in this study, a cartel forms whenever the collusive payoff exceeds the competitive payoff (regardless of the size of the collusive payoff or the magnitude of the difference between collusive and competitive payoffs). In practice, cartels seem more likely to form when the potential gain from collusive activity is large. This may reflect two distinct considerations. First, there may be a set-up cost (not explicitly incorporated in the current model) of initiating a collusive agreement that is paid once in the beginning of the initial period when a cartel first forms.[14] Second, recall that the infinite repetition of the Nash equilibrium is always an equilibrium of the dynamic game. Managers may be more likely to coordinate on a collusive equilibrium than the competitive equilibrium when the collusive payoff is relatively large. Motivated by these considerations, I next explore how a deterioration in market conditions impacts both the collusive payoff ($V_i^C(l)$) and the gain from collusion ($V_i^C(l) - V_i^N(l)$).

Condition 2. i) $V_b^C(l) > V_b^N(l)$, and

ii) $V_a^M > V_b^C(l)$.

Condition 2(i), the reverse of Condition 1, restricts attention to markets wherein a cartel forms absent a change in market state. This condition focuses the analysis on the impact of a change in market state on the gain/value from collusion rather than incentives to form cartels (which was analyzed in the previous section). Condition 2(ii) ensures that the collusive payoff under regime b is less than the monopoly payoff under regime a. If managers can obtain a collusive payoff, absent a change in market state, which exceeds the maximum possible collusive payoff (i.e., the monopoly payoff) after a deterioration in market state, then a deterioration in market state will never increase the value of collusion. In Technical Appendix D.4, I provide a lower-level sufficient condition that ensures Condition 2(ii) holds. Specifically, I show that Condition 2(ii) is satisfied when 1) the discount factor is sufficiently low and 2) the magnitude of the change in market state is moderate.

Proposition 2. *Suppose Condition 2 holds. There exists an $\bar{l}$ such that the following hold when $l \geq \bar{l}$:*

i) $V_a^C(l) > V_b^C(l)$, *and*

ii) $V_a^C(l) - V_a^N(l) > V_b^C(l) - V_b^N$.

[14] This cost may reflect a initial costs of communication/coordination necessary to establish the cartel, or other start-up costs involved in successfully forming and organizing a cartel. To illustrate, suppose the collusive payoff equals the competitive payoff by an infinitesimal margin. In this case, it seems unlikely that managers' will expend the effort and time necessary to setup a cartel for a negligible gain.

Proposition 2 states that a deterioration in market conditions enhances the collusive payoff (part i) and increases the potential gain from collusion (part ii) when managers are sufficiently loss averse. Note that this result occurs despite the fact that a deterioration in market state, such as a reduction in demand or an increase in input cost, typically reduces collusive payoffs when managers are loss neutral.

Proposition 2(i) is driven by the fact, discussed in the previous section, that a deterioration in market state stabilizes collusion and reduces incentives to defect when managers are loss averse. Proposition 2(ii) is driven by the effects outlined in the previous section, but also by a distinct, albeit closely related, consideration. At the beginning of the initial period, when managers decide whether to form a cartel, the competitive payoff is smaller after the deterioration in market state (i.e., $V_a^N(l) < V_b^N$). This is the case not only due to the fact that the deterioration in market state reduces Nash profits (Assumption 4(iii)), but also due to the fact that managers perceive the competitive payoff as a loss.

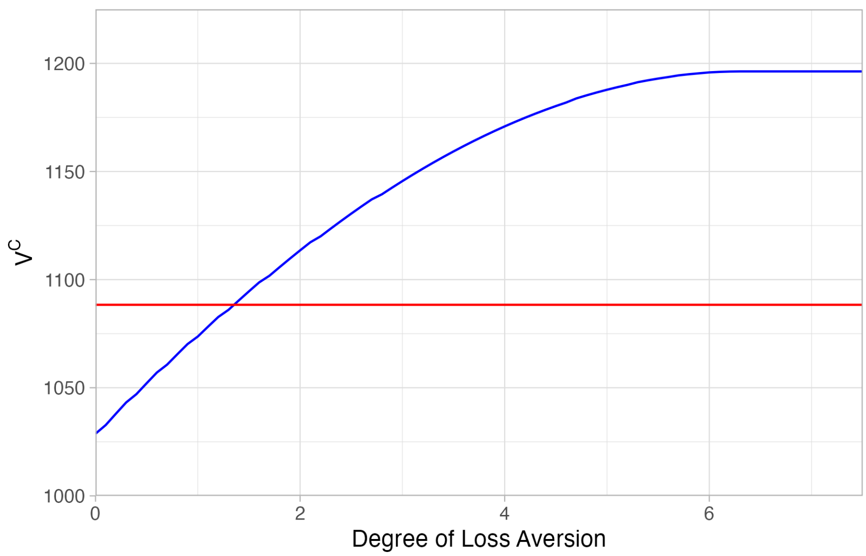


Figure 3: Collusive Payoff by Degree of Loss Aversion Before (Red) and After (Blue) a 25% Increase in Marginal Cost.

Notes: This figure depicts the collusive payoff as a function of the degree of loss aversion. Parameters: $a = 100$, $b = 2$, $e = 1$, $N = 7$, $F = 0$, $\alpha = 1$ and $\delta = .8$. Absent the deterioration in market conditions, all firms have a marginal cost of $c = 10$. The blue curve denotes the collusive payoff V_a^C after a 25% increase in marginal cost (i.e., $c = 12.5$) and the red curve denotes the collusive payoff V_b^C absent an increase in marginal cost (i.e., $c = 10$).

Figure 3 depicts the collusive payoff after an increase in marginal cost (in blue) and absent an increase in marginal cost (in red) as a function of the degree of loss aversion.[15] As expected, an increase in marginal cost reduces the collusive payoff when managers are relatively loss neutral (i.e., when l is sufficiently close to 0). When managers are sufficiently loss averse, the increase in marginal cost enhances the collusive payoff.

[15] Additional simulation results are presented in Technical Appendix G.4.

Note that when l is particularly large, managers earn the monopoly payoff after the deterioration in market conditions (i.e., $V_a^C(l) = V_a^M$).

7 Pricing Results

In this section, I show that a deterioration in market state can result in higher prices, particularly during early periods of collusion. To characterize optimal cartel prices, it is necessary to place a number of additional restrictions on the model. These restrictions are contained in the following assumption which is assumed to hold for the remainder of this section.

Assumption 6. *i)* $\Omega = [c, d] \subset \mathbb{R}$ *where* $c < d$,

ii) for $i \in \{a, b\}$, $\pi_i(x)$ *is strictly increasing in* x *for* $x < x_i^M$ *and strictly decreasing in* x *for* $x > x_i^M$,

iii) for $i \in \{a, b\}$, $x_i^N < x_i^M$, *and*

iv) for $i \in \{a, b\}$, $u(\pi_i^D(x)) - u(\pi_i(x))$ *is strictly increasing in* x *for all* $x \in \Omega$ *such that* $x \geq x_i^N$.

Assumption 6(ii) ensures that collusive profits are strictly quasi-concave. This assumption prevents the existence of multiple optimal price paths, simplifying the analysis. Assumption 6(iii) ensures that the monopoly value of the choice variable x exceeds the Nash value of the choice variable. This assumption is made without loss of generality because if $x^N > x^M$, then the choice variable can instead be taken to be $\tilde{x} = -x$.[16] Assumption 6(iv) mirrors Assumption C3 in Harrington (2004) and ensures that collusion involving higher prices is more difficult to sustain than collusion involving lower prices.

The main result of this section applies when the following condition holds.

Condition 3. i) $\delta < \frac{u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M))}{u(\pi_b^M) - u(\pi_b^N) + u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M))}$, and

ii) $x_b^N < x_a^M$.

Condition 3(i) holds if managers' discount factors are sufficiently low. When managers are sufficiently patient that incentive compatibility constraints do not bind prior to the deterioration in market state, managers can set the monopoly price in all periods. In this case, a deterioration in market state, unless the deterioration increases the monopoly price, cannot result in increased collusive prices. Condition 3(ii) ensures the monopoly price after a deterioration in market state exceeds the Nash equilibrium price absent a deterioration in market state. This condition holds if the magnitude of the deterioration is sufficiently moderate. If the change in market state is more drastic, then monopoly prices after the change in market

[16] Recall that the choice variable x is referred to as a price, for expositional ease, throughout the analysis. However, the model does not require x to represent prices. For example, x could instead represent output levels q. In this case, the choice variable is $x = -q$ and the results of this section establish that a deterioration in market state can cause lower output levels.

state may be less than competitive prices absent the change in market state. In this case, a deterioration in market state always causes lower prices.

Let $\{x_{b,t}\}_{t=1}^{\infty}$ and $\{x_{a,t}\}_{t=1}^{\infty}$ denote optimal price paths under regime b and a, respectively.

Proposition 3. *Suppose Condition 3 holds. There exists an $\bar{l}$ such that if $l \geq \bar{l}$, then there exists a $T \in \mathbb{N} \cup \{\infty\}$ such that $x_{b,t} < x_{a,t}$ for all $t \leq T$.*

Proposition 3 establishes that a deterioration in market state can increase collusive prices in early time periods, when managers are sufficiently loss averse. To understand this result, suppose a cartel would form both with and without the deterioration in market state. The pricing dynamics in Proposition 3 reflect two countervailing effects. First, due to conventional considerations, deteriorations in market state tend to reduce collusive prices. This effect is termed the *standard effect* and, all else equal, causes collusive prices to decline following a deterioration in market state.[17]

Second, when managers are loss averse, a reversion to Nash equilibrium play is perceived as a loss. The size of this loss depends on the difference between the current reference point r_t and the Nash equilibrium utility level. When this difference is large, managers perceive the punishment phase as a significant loss in utility and, as a result, are hesitant to defect from collusion. This effect enhances the stability of collusion (i.e., relaxes the incentive compatibility constraints) and permits the cartel to set higher prices. This effect is termed the *stability effect* and is largest when $r_t - u(\pi^N)$ is large (i.e., when the difference between the reference point and the competitive utility level is large).

When managers are loss neutral, the stability effect does not occur and, as a result, a deterioration in market conditions reduces collusive prices due to the standard effect. However, when managers are loss averse, the stability effect can overpower the standard effect. In early periods of collusion, the reference point is $r_1 = u(\pi_b^N)$ or, depending on how reference points adjust over time, slightly above $r_1 = u(\pi_b^N)$. Absent a deterioration in market state, a reversion to Nash competition results in a utility level of $u(\pi_b^N)$ which is not perceived as a loss (or, at a minimum, not as a significant loss). However, after a deterioration in market state, a reversion to Nash competition results in a utility of $u(\pi_a^N) < r_1$ which is perceived as a pronounced loss. Thus, in early periods of collusion, the stability effect is zero (or small) absent a deterioration in market state, but is strong if market conditions have deteriorated. This results in higher prices in early periods of collusion under regime a than under regime b.

Cartel prices in later periods of collusion are more difficult to characterize. As collusion continues, reference points adjust upwards and, in later periods of collusion, reference points exceed Nash equilibrium

[17]For certain deteriorations in market state, such as an increase in marginal cost, the standard effect may lead to higher prices after the deterioration in market state. In these cases, the deterioration in market state could cause higher prices even absent the presence of loss aversion.

utility levels under both regimes. Thus, the stability effect is strong in both regimes and, as a result, prices absent a deterioration in market state can exceed prices after a deterioration in market state.

Figure 4 plots the optimal price path before (red) and after (blue) a reduction in demand (specifically, the demand intercept). When managers are loss neutral (Panel A of Figure 4), the stability effect does not occur and, as a result, a reduction in demand reduces cartel prices in all periods due to the standard effect. However, when managers are loss averse (Panel B of Figure 4), collusive prices under regime a (blue) exceed prices under regime b (red) in early periods of collusion due to the relatively pronounced stability effect under regime a. In later periods of collusion, the stability effect is present under both regimes and prices are instead driven by the standard effect, resulting in higher prices absent a reduction in demand (regime b).

Panel A: $l = 0$ Panel B: $l = 2$

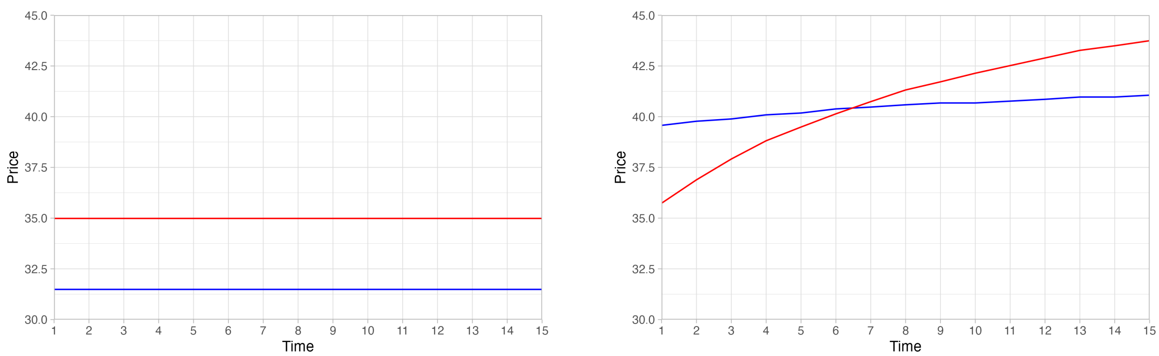


Figure 4: Optimal Price Paths Before (Red) and After (Blue) a 10% Reduction in Demand by Degree of Loss Aversion.

Notes: This figure depicts optimal price paths before (red) and after (blue) a 10% reduction in a for loss neutral managers (Panel A) and managers with a degree of loss aversion of 2 (Panel B). Parameters: $b = 2$, $e = 1$, $c = 0$, $N = 3$, $\delta = .25$, $F = 0$ and $\alpha = .9$. Prior to the deterioration in market conditions, $a = 100$. The blue curve depicts the optimal price path after the deterioration in market conditions (i.e., $a = 90$) and the red curve depicts the optimal price path prior to the deterioration in market conditions (i.e., $a = 100$). The Nash equilibrium price before (after) the change is 25 (22.5). The monopoly price before (after) the change is 50 (45).

While the exact length of time for which $x_{b,t} < x_{a,t}$ depends on a variety of factors (especially, the rate of reference point adjustment α), the following proposition demonstrates that this length of time may be substantial.

Proposition 4. *Suppose Condition 3 holds and $m(r, u) = r$. There exists an $\bar{l}$ such that if $l \geq \bar{l}$, then $x_{b,t} < x_{a,t}$ for all t.*

Proposition 4 implies that, when reference points are constant, collusive prices after the deterioration in market state exceed prices absent a deterioration in market state in all periods. To illustrate Proposition 4, first consider regime b. If reference points remain constant at $r_1 = u(\pi_b^N)$, the punishment for defection (Nash competition) is never perceived as a loss in utility. Thus, the stability effect is zero in all periods under regime b. However, after the market state deteriorates, the punishment for defection is perceived as a loss in all future periods. Thus, the stability effect occurs under regime a and, when managers are sufficiently loss averse, overpowers the standard effect, resulting in higher prices in all periods.

8 Conclusion

Empirical evidence suggests that deteriorations in market conditions, such as a reduction in demand, an increase in marginal cost or the entry of a competitor, often precede the formation of a cartel. However, conventional theoretical models, which assume managers are loss neutral, do not imply that deteriorations of this kind enhance the sustainability of collusion. The preceding analysis establishes that, when colluding managers are instead averse to losses, deteriorations in market conditions can cause the formation of a cartel. Loss averse managers perceive continued competition as a painful loss in utility after a deterioration in market conditions. To avoid a loss, managers turn to collusion. Due to similar considerations, I find that deteriorations in market conditions can also enhance the payoff managers receive from collusion, increase the gain in utility from collusion, and result in higher collusive prices. These effects do not arise after improvements in market conditions, because managers then perceive competitive profits as gains rather than losses.

A number of model extensions are presented in the technical appendix. Section E extends the model to include an antitrust authority that detects and penalizes cartels. The case of naive managers who do not anticipate changes in their reference points is explored in Section D.1. Gradual deteriorations in market conditions that occur over a number of periods are explored in Section F. The robustness of results to alternative initial reference points is established in Section C. Section G presents additional simulation results.

Future research might consider two variations on the foregoing analysis. First, the model considers collusion by means of grim trigger strategies rather than alternative, more sophisticated punishment strategies. However, the effects outlined in this study seem likely to occur in any model wherein firms punish infringing firms with Nash or below Nash profit levels (e.g., stick and carrot punishments (Abreu, 1986)). This is the case as these profit levels will be perceived by colluding managers as a loss, causing the effects described above to occur. Second, managers/firms are assumed to be symmetric throughout the analysis. The key mechanisms should nonetheless extend to asymmetric settings, provided that worsening market conditions lead each manager to view competitive outcomes as losses.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process During the preparation of this work the author used ChatGPT in order to check grammar and improve the flow of the manuscript. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

References

Abreu, Dilip. 1986. "Extremal equilibria of oligopolistic supergames." *Journal of Economic Theory*, 39(1): 191–225.

Aston, John, and Andrew Pressey. 2012. "Who manages cartels? the role of sales and marketing managers within international cartels: Evidence from the european union 1990-2009." *ESRC Centre for Competition Policy, University of East Anglia, CCP Working Paper*, 12–1.

Barberis, Nicholas C. 2013. "Thirty years of prospect theory in economics: A review and assessment." *Journal of Economic Perspectives*, 27(1): 173–196.

Barkan, Rachel, and Jerome R Busemeyer. 1999. "Changing plans: dynamic inconsistency and the effect of experience on the reference point." *Psychonomic Bulletin & Review*, 6(4): 547–554.

Bernard, Mark. 2011. "A folk theorem for endogenous reference points." *Economics Letters*, 112(3): 223–225.

Bos, Iwan, Stephen Davies, Joseph E Harrington Jr, and Peter L Ormosi. 2018. "Does enforcement deter cartels? A tale of two tails." *International Journal of Industrial Organization*, 59: 372–405.

Bowman, David, Deborah Minehart, and Matthew Rabin. 1999. "Loss aversion in a consumption–savings model." *Journal of Economic Behavior & Organization*, 38(2): 155–178.

Chen, Haipeng, and Akshay R Rao. 2002. "Close encounters of two kinds: False alarms and dashed hopes." *Marketing Science*, 21(2): 178–196.

Colombo, Stefano. 2013. "Product differentiation and collusion sustainability when collusion is costly." *Marketing Science*, 32(4): 669–674.

Crum, Roy L, Dan J Laughhunn, and John W Payne. 1981. "Risk-seeking behavior and its implications for financial models." *Financial Management*, 20–27.

Ely, Kirsten M. 1991. "Interindustry differences in the relation between compensation and firm performance variables." *Journal of Accounting Research*, 29(1): 37–58.

Feltovich, Nick. 2011. "The Effect of Subtracting a Constant from all Payoffs in a Hawk-Dove Game: Experimental Evidence of Loss Aversion in Strategic Behavior." *Southern Economic Journal*, 77(4): 814–826.

Frederick, Shane, and George Loewenstein. 1999. "Well-Being the Foundations of Hedonic Psychology." , ed. N. Schwarz D. Kahneman ED, Chapter Hedonic adaptation, 302–329.

Gallice, Andrea. 2010. "The neglected effects of demand characteristics on the sustainability of collusion." *Research in Economics*, 64(4): 240–246.

Ganslandt, Mattias, Lars Persson, and Helder Vasconcelos. 2012. "Endogenous mergers and collusion in asymmetric market structures." *Economica*, 79(316): 766–791.

Gill, David, and Victoria Prowse. 2012. "A structural analysis of disappointment aversion in a real effort competition." *American Economic Review*, 102(1): 469–503.

Green, Edward J, and Robert H Porter. 1984. "Noncooperative collusion under imperfect price information." *Econometrica*, 87–100.

Grout, Paul A, and SMIA Sonderegger. 2005. "Predicting Cartels (OFT 773)."

Grout, Paul A, and SMIA Sonderegger. 2007. "Structural approaches to cartel detection." In *European Competition Law Annual: 2006. Enforcement of Prohibition of Cartels*. 83–103. Hart Publishing.

Haigh, Michael S, and John A List. 2005. "Do professional traders exhibit myopic loss aversion? An experimental analysis." *Journal of Finance*, 60(1): 523–534.

Harrington, Jr., Joseph E. 2004. "Cartel pricing dynamics in the presence of an antitrust authority." *RAND Journal of Economics*, 651–673.

Harrington, Jr., Joseph E. 2006. "How do cartels operate?" *Foundations and Trends in Microeconomics*, 2(1): 1–105.

Healy, Paul M. 1985. "The effect of bonus schemes on accounting decisions." *Journal of Accounting and Economics*, 7(1): 85–107.

Helson, Harry. 1964. "Adaptation-level theory: An experimental and systematic approach to behavior."

Herold, Daniel, and Johannes Paha. 2018. "Cartels as defensive devices: evidence from decisions of the European Commission 2001–2010." *Review of Law & Economics*, 14(1): 20160035.

Ivaldi, Marc, Bruno Jullien, Patrick Rey, Paul Seabright, and Jean Tirole. 2003. "The economics of tacit collusion." *IDEI Working Paper*.

Ivaldi, Marc, Bruno Jullien, Patrick Rey, Paul Seabright, and Jean Tirole. 2007. "The economics of tacit collusion: implications for merger control." *Contributions to Economic Analysis*, 282: 217–239.

Jepson, Christopher, George Loewenstein, and Peter Ubel. 2001. "Actual versus estimated differences in quality of life before and after renal transplant."

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: An analysis of decisions under risk." *Econometrica*, 47(2): 278.

Kahneman, Daniel, and Jackie Snell. 1992. "Predicting a changing taste: Do people know what they will like?" *Journal of Behavioral Decision Making*, 5(3): 187–200.

Karlsson, Niklas, George Loewenstein, and Duane Seppi. 2009. "The ostrich effect: Selective attention to information." *Journal of Risk and Uncertainty*, 38: 95–115.

Klein, Timo, and Maarten Pieter Schinkel. 2019. "Cartel Stability by a Margin." *Amsterdam Law School Research Paper.*

Levenstein, Margaret C, and Valerie Y Suslow. 2015. "Cartels and collusion-Empirical evidence." In *Oxford Handbook on International Antitrust Economics.* , ed. Roger D. Blair and D. Daniel Sokol, Chapter 18, 442–63. Oxford University Press.

Loewenstein, George, Ted O'Donoghue, and Matthew Rabin. 2003. "Projection bias in predicting future utility." *Quarterly Journal of Economics*, 1209–1248.

Maggi, Mario Alessandro. 2004. "A characterization of s-shaped utility functions displaying loss aversion." Quaderni di Dipartimento-EPMQ.

Merchant, Kenneth A, and Jean-Francois Manzoni. 1989. "The achievability of budget targets in profit centers: A field study." *Readings in Accounting for Management Control*, 496–520.

Novemsky, Nathan, and Daniel Kahneman. 2005. "The boundaries of loss aversion." *Journal of Marketing Research*, 42(2): 119–128.

Pedace, Roberto, and Janet Kiholm Smith. 2013. "Loss aversion and managerial decisions: Evidence from major league baseball." *Economic Inquiry*, 51(2): 1475–1488.

Rotemberg, Julio J, and Garth Saloner. 1986. "A supergame-theoretic model of price wars during booms." *The American Economic Review*, 76(3): 390–407.

Ryder Jr, Harl E, and Geoffrey M Heal. 1973. "Optimal growth with intertemporally dependent preferences." *The Review of Economic Studies*, 40(1): 1–31.

Shalev, Jonathan. 1998. "Loss Aversion in Repeated Games." UCL-Université Catholique de Louvain. Université Catholique de Louvain. Center for Operations Research and Econometrics [CORE].

Shalev, Jonathan. 2000. "Loss aversion equilibrium." *International Journal of Game Theory*, 29: 269–287.

Singh, Nirvikar, and Xavier Vives. 1984. "Price and quantity competition in a differentiated duopoly." *RAND Journal of Economics*, 15(4): 546–554.

Spagnolo, Giancarlo. 1999. "On Interdependent Supergames: Multimarket Contact, Concavity, and Collusion." *Journal of Economic Theory*, 89(1): 127–139.

Spagnolo, Giancarlo. 2000. "Stock-related compensation and product-market competition." *RAND Journal of Economics*, 22–42.

Spagnolo, Giancarlo. 2005. "Managerial incentives and collusive behavior." *European Economic Review*, 49(6): 1501–1523.

Sullivan, Kathryn, and Thomas Kida. 1995. "The effect of multiple reference points and prior gains and losses on managers risky decision making." *Organizational Behavior and Human Decision Processes*, 64(1): 76–83.

Thomadsen, Raphael, and Ki-Eun Rhee. 2007. "Costly collusion in differentiated industries." *Marketing Science*, 26(5): 660–665.

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

Willman, Paul, Mark Fenton-O'Creevy, Nigel Nicholson, and Emma Soane. 2002. "Traders, managers and loss aversion in investment banking: a field study." *Accounting, Organizations and Society*, 27(1-2): 85–98.

Wilson, Timothy D, and Daniel T Gilbert. 2005. "Affective forecasting: Knowing what to want." *Current Directions in Psychological Science*, 14(3): 131–134.

Zank, Horst. 2010. "On probabilities and loss aversion." *Theory and Decision*, 68: 243–261.

A Proofs

Proof of Lemma 1. The proof is identical for both regimes. For expositional ease, I have suppressed the subscript $i \in \{a, b\}$ in the remainder of the proof. First, I show that the payoff function $F\left(\{x_t\}_{t=1}^{\infty}\right) = \sum_{t=1}^{\infty} \delta^{t-1} u(\pi(x_t); r_t)$ is a continuous function $F : \Omega^{\infty} \to \mathbb{R}$ where $\mathbb{R}$ is endowed with the standard topology. Ω^{∞} is endowed with the product topology which implies the projection $p_T\left(\{x_t\}_{t=1}^{\infty}\right) = x_T$ is continuous for all T.

First, I show that $u(\pi(x_T); r_T)$ is continuous in $\{x_t\}_{t=1}^{\infty}$ on the product space Ω^{∞} for all T. Let $u_T = u\left(\pi\left(p_T\left(\{x_t\}_{t=1}^{\infty}\right)\right)\right)$ denote base utility in period T as a function of $\{x_t\}_{t=1}^{\infty}$. u_T is continuous in $\{x_t\}_{t=1}^{\infty}$ by the continuity of p_T, the continuity of π (Assumption 3(i)), and the continuity of u (Assumption 1(i)). Next, I show that r_T is continuous in $\{x_t\}_{t=1}^{\infty}$ for all T. The proof follows by induction. $r_2 = m(r_1, u_1)$ is continuous in $\{x_t\}_{t=1}^{\infty}$ by the continuity of m and the continuity of u_1 in $\{x_t\}_{t=1}^{\infty}$. Suppose r_T is continuous in $\{x_t\}_{t=1}^{\infty}$. Then $r_{T+1} = m(r_T, u_T)$ is continuous in $\{x_t\}_{t=1}^{\infty}$ by the continuity of m, the continuity of r_T, the continuity of u_T, and the fact that compositions of continuous functions are continuous. Finally, $u(\pi(x_T); r_T) = u_T - lL\left(r_T - u_T\right)$ is continuous in $\{x_t\}_{t=1}^{\infty}$ by the continuity of L (Assumption 1(iv)), the continuity of u_T in $\{x_t\}_{t=1}^{\infty}$, the continuity of r_T in $\{x_t\}_{t=1}^{\infty}$, and the fact that differences and compositions of continuous functions are continuous.

As sums of continuous functions are continuous,

$$F_n\left(\{x_t\}_{t=1}^{\infty}\right) = \sum_{t=1}^{n} \delta^{t-1} u(\pi(x_t); r_t)$$

is continuous for all n. $\left\{F_n\left(\{x_t\}_{t=1}^{\infty}\right)\right\}_{n=1}^{\infty}$ is a sequence of continuous functions that converges uniformly to $F\left(\{x_t\}_{t=1}^{\infty}\right)$. To see that $\left\{F_n\left(\{x_t\}_{t=1}^{\infty}\right)\right\}_{n=1}^{\infty}$ converges uniformly, note that $\pi(x)$ is a continuous function (Assumption 3(i)) on a compact set. Thus, $\pi(x)$ is bounded above and below. Let $\bar{\pi}$ denote the upper bound and let $\underline{\pi}$ denote the lower bound. $u(\pi(x); r_t)$ is therefore bounded above and below as

$$\underline{u} \equiv u(\underline{\pi}; \bar{r}) \leq u(\pi(x); r_t) \leq u(\bar{\pi}) \equiv \bar{u}$$

where $r_t \leq \bar{r} = \max\{\bar{u}, r_1\}$ by the definition of m. Let $\epsilon > 0$, then

$$\left|F_n\left(\{x_t\}_{t=1}^{\infty}\right) - F\left(\{x_t\}_{t=1}^{\infty}\right)\right| = \left|\sum_{t=n+1}^{\infty} \delta^{t-1} u(\pi(x_t); r_t)\right| \leq \sum_{t=n+1}^{\infty} \delta^{t-1} \max\left\{|\bar{u}|, |\underline{u}|\right\} = \delta^n \frac{\max\left\{|\bar{u}|, |\underline{u}|\right\}}{1 - \delta} < \epsilon$$

for sufficiently large n, for all $\{x_t\}_{t=1}^{\infty} \in \Omega^{\infty}$. Thus, $\left\{F_n\left(\{x_t\}_{t=1}^{\infty}\right)\right\}_{n=1}^{\infty}$ converges uniformly to $F\left(\{x_t\}_{t=1}^{\infty}\right)$.

Thus, the uniform limit theorem implies that $F\left(\{x_t\}_{t=1}^{\infty}\right)$ is continuous on Ω^{∞}. Ω is compact which implies, by Tychonoff's Product Theorem, Ω^∞ is compact. The continuity of $F\left(\{x_t\}_{t=1}^\infty\right)$ in $\{x_t\}_{t=1}^\infty$, the continuity of $u(\pi^D(x_t); r_t) = u(\pi^D(p_t\left(\{x_t\}_{t=1}^\infty\right)); r_t)$ in $\{x_t\}_{t=1}^\infty$,[18] and the continuity of $V^N(r_T)$ in $\{x_t\}_{t=1}^\infty$[19] implies that both sides of the inequality constraints in (5) are continuous. Thus, Ψ is a closed set. Ψ is a closed subset of a compact set Ω^∞ which implies Ψ is compact. $\Psi \neq \emptyset$ by supposition. A solution to (4) exists as the cartel's problem involves maximizing a continuous function $F\left(\{x_t\}_{t=1}^\infty\right)$ over a compact set Ψ. $\square$

The following lemma establishes that the infinite repetition of the Nash equilibrium constitutes a sub-game perfect equilibrium of the dynamic game. This result ensures that Nash competition in the punishment phase is self-enforcing (i.e., no firm wishes to deviate during the punishment phase). Let x^N denote the Nash equilibrium price and let π^N denote Nash equilibrium profits. Let $\tilde{\pi}(x; x^N)$ denote a managers payoff when charging price x when all rivals charge price x^N. Let $W\left(\{\pi_t\}_{t=T}^\infty; r_T\right)$ denote the discounted present value of the payoff, in period T, from a profit sequence of $\{\pi_t\}_{t=T}^\infty$ when the reference point at time T is r_T. Thus,

$$W\left(\{\pi_t\}_{t=T}^\infty; r_T\right) = \sum_{t=T}^{\infty} \delta^{t-T} u(\pi_t; r_t)$$

where $r_t = m(r_{t-1}, u(\pi_{t-1}))$ for $t > T$. The following lemma ensures that the infinite repetition of the static Nash equilibrium constitutes a sub-game perfect Nash equilibrium of the dynamic game.

Lemma 2. *Suppose, for $i \in \{a, b\}$, $W\left(\{\pi_i^N\}_{t=1}^\infty; r\right) \geq W\left(\{\pi_t\}_{t=1}^\infty; r\right)$ for any r and any $\{\pi_t\}_{t=1}^\infty$ where $\pi_t \leq \pi_i^N$ for all t. Then, a strategy profile wherein all managers charge price x^N in all periods (regardless of prior play) constitutes a sub-game perfect equilibrium of the dynamic game.*

Proof. First, I establish that the proposed equilibrium is a Nash equilibrium of the dynamic game. Strategy profiles in the dynamic game are characterized by a function $s : \mathcal{H} \to \Omega$ where $\mathcal{H}$ is the set of possible pricing histories of all managers. The proposed equilibrium strategy profile is s^N where $s^N(H) = x^N$ for all $H \in \mathcal{H}$. Suppose manager 1 deviates to an alternative strategy function $\tilde{s}$. Let $\tilde{x}_1, \tilde{x}_2 ...$ denote the sequence of prices consistent with manager 1 following strategy function $\tilde{s}$ and all rival managers following the strategy function s^N. As x^N is the Nash equilibrium price, manager 1 earns profit less than π^N in each period t. By the supposition, the discounted present value of manager 1's payoff is less than or equal to its payoff from repeated Nash competition. Thus, the manager does not wish to deviate.

Next, I establish that the proposed equilibrium is a sub-game perfect equilibrium. Note that the history of the game (i.e., prices set by all managers in previous periods) impacts a manager's future payoff only

[18] The continuity of $u(\pi^D(x_t); r_t)$ in $\{x_t\}_{t=1}^\infty$ follows from an argument analogous to the above. Note that $\pi^D(x)$ is continuous by Assumption 3(i).

[19] The continuity of $V^N(r_T)$ in $\{x_t\}_{t=1}^\infty$ follows from the continuity of r_T in $\{x_t\}_{t=1}^\infty$, shown earlier in the proof.

through the manager's reference point. As sub-games in the dynamic game correspond to distinct pricing histories, it suffices to establish that no manager wishes to defect for any (current) reference point r. The result follows by the above argument, which holds for any r. ☐

Lemma 3. $V_b^N \geq V_a^N$.

Proof. The following proof holds for a general initial reference point r_1.

$$V_i^N = \sum_{t=1}^{\infty} \delta^{t-1} \left[u(\pi_i^N) - lL(r_t^i - u(\pi_i^N)) \right]$$

for $i \in \{a, b\}$ where r_t^i denotes the sequence of reference points resulting from repeated Nash equilibrium utilities under regime i and an initial reference point of r_1. By Assumption 4(iii), it suffices to show that

$$L(r_t^b - u(\pi_b^N)) \leq L(r_t^a - u(\pi_a^N))$$

for all t, which holds if

$$r_t^b - u(\pi_b^N) \leq r_t^a - u(\pi_a^N)$$

for all t. The above inequality is proven through induction. For period $t = 1$, $r_1^b - u(\pi_b^N) \leq r_1^a - u(\pi_a^N)$ follows from $r_1^b = r_1^a = r_1$ and Assumption 4(iii). Suppose $r_{t-1}^b - u(\pi_b^N) \leq r_{t-1}^a - u(\pi_a^N)$. By the definition of m, $r_t^b - u(\pi_b^N) \leq r_t^a - u(\pi_a^N)$ for $t > 1$ holds if

$$\alpha r_{t-1}^b + (1-\alpha)u(\pi_b^N) - u(\pi_b^N) \leq \alpha r_{t-1}^a + (1-\alpha)u(\pi_a^N) - u(\pi_a^N)$$

$$\iff \alpha r_{t-1}^b - \alpha u(\pi_b^N) \leq \alpha r_{t-1}^a - \alpha u(\pi_a^N)$$

$$\iff r_{t-1}^b - u(\pi_b^N) \leq r_{t-1}^a - u(\pi_a^N)$$

which holds by supposition. Thus, $L(r_t^b - u(\pi_b^N)) \leq L(r_t^a - u(\pi_a^N))$ for all t and $V_b^N \geq V_a^N$. ☐

The following Lemma establishes that $\left\{x_a^M\right\}_{t=1}^{\infty}$ is incentive compatible when l is sufficiently large. Let $\Psi_i(l)$ denote the set of incentive compatible price paths under regime i when the degree of loss aversion is l. Additionally, let $\pi_i^{DM} \equiv \pi_i^D(x_i^M)$ denote the profit a manager earns when defecting from collusion when the collusive price is x_i^M under regime i.

Lemma 4. $\left\{x_a^M\right\}_{t=1}^{\infty} \in \Psi_a(l)$ *if*

$$l \geq \bar{l} = \frac{u(\pi_a^{DM}) + \delta \frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta}}{\delta L\left(r_1 - u(\pi_a^N)\right)}.$$

Proof. It suffices to show that

$$V_a^M(r_T;l) \equiv \sum_{t=T}^{\infty} \delta^{t-T} u(\pi_a^M; r_t, l) \geq u(\pi_a^{DM}; r_T, l) + \delta V_a^N(r_T;l) \tag{6}$$

for all $T \in \{1, 2, 3 \ldots\}$ where $r_t = m(r_{t-1}, u(\pi_a^M))$ for $t > 1$. The definition of m and Assumption 4(ii) imply that $r_t \leq r_{t+1}$ for all t and $r_t \leq u(\pi_a^M)$ for all t. Thus,

$$u(\pi_a^M; r_t, l) = u(\pi_a^M) \tag{7}$$

for all t. Additionally, $u(\pi_a^{DM}; r_T, l) = u(\pi_a^{DM})$ by Assumption 3(iii) and $r_t \leq u(\pi_a^M)$ for all t. Therefore, the inequalities in (6) become

$$\sum_{t=T}^{\infty} \delta^{t-T} u(\pi_a^M) \geq u(\pi_a^{DM}) + \delta V_a^N(r_T;l)$$

for all $T \in \{1, 2, 3 \ldots\}$, or

$$\frac{u(\pi_a^M)}{1-\delta} - u(\pi_a^{DM}) \geq \delta V_a^N(r_T;l).$$

for all $T \in \{1, 2, 3 \ldots\}$. Fix $T \in \{1, 2, 3 \ldots\}$. Let $\tilde{r}_T = r_T$ and let $\tilde{r}_t = m(\tilde{r}_{t-1}, u(\pi_a^N))$ for $t > T$. Note that

$$\begin{aligned}
V_a^N(r_T;l) = \sum_{t=T}^{\infty} \delta^{t-T} u(\pi_a^N; \tilde{r}_t, l) &\leq u(\pi_a^N; r_1, l) + \sum_{t=T+1}^{\infty} \delta^{t-T} u(\pi_a^N; \tilde{r}_t, l) \\
&\leq u(\pi_a^N; r_1, l) + \sum_{t=T+1}^{\infty} \delta^{t-T} u(\pi_a^N) \\
&= u(\pi_a^N; r_1, l) + \sum_{t=1}^{\infty} \delta^{t} u(\pi_a^N) \\
&= u(\pi_a^N; r_1, l) + \delta \frac{u(\pi_a^N)}{1-\delta} \qquad (8)
\end{aligned}$$

where the first inequality follows from $\tilde{r}_T = r_T \geq r_1$ for all T and Assumption 1. The second inequality in Equation (8) follows from $u(\pi; r, l) \leq u(\pi)$ (Assumption 1).

$r_1 > u(\pi_a^N)$ by $r_1 = u(\pi_b^N)$ (Assumption 4(i)) and $\pi_a^N < \pi_b^N$ (Assumption 4(iii)). Thus, $L(r_1 - u(\pi_a^N)) > 0$ by Assumption 1(iv).

Suppose $l \geq \bar{l}$. Then,

$$\begin{aligned}
u(\pi_a^{DM}) + \delta V_a^N(r_T; l) &\leq u(\pi_a^{DM}) + \delta u(\pi_a^N; r_1, l) + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
&= u(\pi_a^{DM}) + \delta u\left(\pi_a^N\right) - \delta l L\left(r_1 - u(\pi_a^N)\right) + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
&\leq u(\pi_a^{DM}) + \delta u\left(\pi_a^N\right) - \delta \bar{l} L\left(r_1 - u(\pi_a^N)\right) + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
&= u(\pi_a^{DM}) + \delta u\left(\pi_a^N\right) \\
&\quad - \left[u(\pi_a^{DM}) + \delta \frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta}\right] + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
&= \delta u\left(\pi_a^N\right) - \delta \frac{u(\pi_a^N)}{1-\delta} + \frac{u(\pi_a^M)}{1-\delta} + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
&= \delta u\left(\pi_a^N\right) - (1-\delta)\, \delta \frac{u(\pi_a^N)}{1-\delta} + \frac{u(\pi_a^M)}{1-\delta} \\
&= \frac{u(\pi_a^M)}{1-\delta}
\end{aligned} \tag{9}$$

where the first inequality follows from (8). The first equality follows from the definition of $u(\pi_a^N; r_1, l)$. The second inequality follows from $l \geq \bar{l}$. The second equality follows from the definition of $\bar{l}$ and $L(r_1 - u(\pi_a^N)) > 0$ (shown earlier in the proof). Thus, Equation 9 implies $\left\{x_a^M\right\}_{t=1}^{\infty} \in \Psi_a(l)$ when $l \geq \bar{l}$. □

The following Lemma establishes that the monopoly price path (after the change in market state) is the unique optimal price path when the monopoly price path is ICC.

Lemma 5. *$\left\{x_a^M\right\}_{t=1}^{\infty}$ is the unique optimal price path when $\left\{x_a^M\right\}_{t=1}^{\infty} \in \Psi_a(l)$.*

Proof. It suffices to show that $\left\{x_a^M\right\}_{t=1}^{\infty}$ generates a strictly larger collusive payoff than any other path when $\left\{x_a^M\right\}_{t=1}^{\infty} \in \Psi_a(l)$. Let $\{x_t\}_{t=1}^{\infty} \in \Omega^{\infty}$ be an alternative path where $x_t \neq x_a^M$ for some t. Let $W(\{\pi_a(x_t)\}_{t=1}^{\infty}; r_1)$ denote the payoff from collusion when the path is $\{x_t\}_{t=1}^{\infty}$ and the initial reference point is r_1.

$$V_a^M(r_1; l) = \frac{u(\pi_a^M)}{1-\delta} > \sum_{t=1}^{\infty} \delta^{t-1} u(\pi_a(x_t)) \geq W(\{\pi_a(x_t)\}_{t=1}^{\infty}; r_1)$$

where the equality follows from $r_t \leq u(\pi_a^M)$ for all t (which follows from the definition of m and Assumption 4(ii)). The first inequality follows from $\pi_a^M > \pi_a(x_t)$ for some t and by Assumption 3(ii). The second inequality follows from from $u(\pi; r, l) \leq u(\pi)$ (Assumption 1). Therefore, $\left\{x_a^M\right\}_{t=1}^{\infty}$ is the unique optimal price path when $\left\{x_a^M\right\}_{t=1}^{\infty} \in \Psi_a(l)$. □

Lemma 6. *$V_a^C(l) = V_a^M(l)$ when $\left\{x_a^M\right\}_{t=1}^{\infty} \in \Psi_a(l)$.*

Proof. The result follows immediately from Lemma 5. □

The following Lemma establishes that a cartel forms when managers are sufficiently loss averse (after the change in market state).

Lemma 7. $V_a^C(l) > V_a^N(l)$ *(thus, a cartel forms) if*

$$l \geq \bar{l} = \frac{u(\pi_a^{DM}) + \delta \frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta}}{\delta L\left(r_1 - u(\pi_a^N)\right)}.$$

Proof. When $l \geq \bar{l}$, $\{x_a^M\}_{t=1}^{\infty} \in \Psi_a(l)$ (by Lemma 4) and $V_a^C(l) = V_a^M(l)$ (by Lemma 6). It remains to establish that $V_a^M(l) > V_a^N(l)$. Note that

$$u(\pi_a^N) < u(\pi_b^N) = r_1 < u(\pi_a^M) \tag{10}$$

where the first inequality follows from Assumption 4(iii), the equality follows from Assumption 4(i), and the second inequality follows from Assumption 4(ii). Thus,

$$\begin{aligned} V_a^N(l) &\leq \sum_{t=1}^{\infty} \delta^{t-1} u(\pi_a^N) \\ &< \sum_{t=1}^{\infty} \delta^{t-1} u(\pi_a^M) \\ &= \frac{u(\pi_a^M)}{1-\delta} = V_a^C(l) \end{aligned}$$

where the first inequality follows from the fact that $u(\pi; r, l) \leq u(\pi)$ (Assumption 1). The second inequality follows from Equation 10. As a result, the cartel forms when $l \geq \bar{l}$. □

Proof of Proposition 1. Suppose $l \geq \bar{l} = \frac{u(\pi_a^{DM}) + \delta \frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta}}{\delta L(r_1 - u(\pi_a^N))}$. Lemmas 4-7 imply that $V_a^C(l) = V_a^M(l) > V_a^N(l)$ (thus, a cartel forms) when $l \geq \bar{l}$. □

Proof of Proposition 2. Suppose $l \geq \bar{l} = \frac{u(\pi_a^{DM}) + \delta \frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta}}{\delta L(r_1 - u(\pi_a^N))}$. Part (i) follows from

$$V_b^C(l) < V_a^M = V_a^C(l)$$

when $l \geq \bar{l}$. The first inequality follows from Condition 2(ii). The equality follows from Lemma 6.

Next, consider Part (ii). Lemma 3 implies $V_b^N \geq V_a^N(l).V_a^N(l) \leq V_b^N$ and $V_b^C(l) < V_a^C(l)$ (from part (i)) imply

$$V_a^C(l) - V_a^N(l) > V_b^C(l) - V_b^N.$$

□

Proof of Proposition 3. The proof follows from Proposition 5 in Technical Appendix C.[20] $\square$

Proof of Proposition 4. Suppose $l \geq \bar{l} = \frac{u(\pi_a^{DM})+\delta\frac{u(\pi_a^N)}{1-\delta}-\frac{u(\pi_a^M)}{1-\delta}}{\delta L(r_1-u(\pi_a^N))}$. By Lemmas 4-7, $\left\{x_a^M\right\}_{t=1}^{\infty} \in \Psi_a(l)$ and a cartel forms when $l \geq \bar{l}$ under regime a. $\left\{x_a^M\right\}_{t=1}^{\infty}$ is the unique optimal price path by Lemma 5. Thus, it suffices to show that $x_a^M > x_{b,t}$ for all t, when $l \geq \bar{l}$. First, note that if a cartel does not form before the change, then $x_{b,t} = x_b^N$ for all t and, by Condition 3(ii), $x_{a,t} = x_a^M > x_b^N = x_{b,t}$ for all t. For the remainder of the proof, assume a cartel forms before the change in market state (i.e., under regime b).

First, note that $u(\pi_b^D(x_a^M))-u(\pi_b(x_a^M)) \geq 0$ by Assumption 3(iii). Suppose $u(\pi_b^D(x_a^M))-u(\pi_b(x_a^M)) = 0$. By Assumption 6(iv), $u(\pi_b^D(x)) - u(\pi_b(x))$ is strictly increasing in x for all $x \geq x_b^N$. Thus, $u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M)) = 0$ implies $u(\pi_b^D(x)) < u(\pi_b(x))$ for $x < x_a^M$ which contradicts Assumption 3(iii). Thus, $u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M)) > 0$. As a cartel forms before the change (by supposition), $\pi_b^M \geq \pi_b^N$ must hold.[21] As $m(r,u) = r$, $r_t = r_1$ for all t.

It follows directly from the derivations in the proof of Proposition 3 that $G(x_a^M) < 0$ (where G is defined in Proposition 3) under Condition 3. $G(x)$ is strictly decreasing in x for $x \geq x_b^N$ by Assumption 6(iv). Thus, $G(x) < 0$ for all $x \geq x_a^M > x_b^N$ (where the last inequality follows from Condition 3(ii)).

Suppose the optimal collusive price under regime b is greater than or equal to x_a^M in at least one period. Let T_1 denote one such period (thus, $x_{b,T_1} \geq x_a^M$).

$$
\begin{aligned}
& u(\pi_b(x_{b,T_1}); r_1) + \sum_{t=T_1+1}^{\infty} \delta^{t-T_1} u(\pi_b(x_{b,t}); r_1) \\
\leq\ & u(\pi_b(x_{b,T_1})) - lL\left(r_1 - u(\pi_b(x_{b,T_1}))\right) + \sum_{t=T_1+1}^{\infty} \delta^{t-T_1} u(\pi_b(x_{b,t})) \\
\leq\ & u(\pi_b(x_{b,T_1})) - lL\left(r_1 - u(\pi_b(x_{b,T_1}))\right) + \delta\frac{u(\pi_b(x_b^M))}{1-\delta} \\
<\ & u(\pi_b^D(x_{b,T_1})) - lL\left(r_1 - u(\pi_b(x_{b,T_1}))\right) + \delta V_b^N \\
\leq\ & u(\pi_b^D(x_{b,T_1})) - lL\left(r_1 - u(\pi_b^D(x_{b,T_1}))\right) + \delta V_b^N \\
=\ & u(\pi_b^D(x_{b,T_1})); r_1) + \delta V_b^N
\end{aligned} \tag{11}
$$

where the first inequality follows from the fact that $u(\pi; r) \leq u(\pi)$ (Assumption 1). The second inequality follows from Assumption 3(ii). The third inequality follows from $G(x_{b,T_1}) < 0$ for $x_{b,T_1} \geq x_a^M$. The fourth inequality follows from Assumption 3(iii) and Assumption 1(iv). (11) implies that the ICC in period t under regime b is not satisfied if $x_{b,t} \geq x_a^M$. Thus, $x_{b,t} < x_a^M = x_{a,t}$ for all t. $\square$

[20]Proposition 5 establishes the conclusion under the more general Assumption 7 and Condition 4.

[21]If $\pi_b^M < \pi_b^N$, then the maximum payoff from collusion is less than the payoff from Nash competition. Therefore, the cartel never forms.

Technical Appendix for "Deteriorating Market Conditions and Cartel Formation under Manager Loss Aversion"

Douglas Turner

B Additional Details Regarding Data Collection

In this section, I provide additional details regarding the collection of data/empirical results presented in Section 2.

To identify which changes in market conditions lead to cartel formation, I review European Commission cartel decisions and record the stated causes of each cartel's formation. European Commission decisions are available at https://competition-policy.ec.europa.eu/index_en. I specifically analyze prohibition decisions. While the US Department of Justice publishes press releases after a cartel member's plea or sentencing, these press releases are typically very brief and do not include detailed information about the cartel (Harrington, 2006). By contrast, European Commission decisions are highly detailed and can be hundreds of pages long. Specifically, I search for changes in the market environment that are alleged to have caused producers to engage in price fixing. For example, the EC decision in the calcium carbide cartel[22] states: "[s]ince the beginning of the 21st century the price of calcium carbide powder for the steel industry has been under pressure, while costs increased and demand shrunk. These developments formed the basis for the meetings between the main European suppliers of calcium carbide powder." In other cases, the decisions explicitly state the producers' motivations for price fixing. For example, the EC decision in the professional videotapes cartel[23] states: "the reasons for participation in the arrangements leading to the first two price increases ... comprised the following: i) the weakness of the Japanese Yen against the Deutsche Mark ... and ii) the fact that the prices ... were comparatively low."

This information is typically contained in the section of the decision pertaining to the origin of the cartel, the cartel's history or a section titled "Description of the Events." I record only changes in the market environment which the decision indicates were relevant for the cartel's formation (i.e., instigated its formation in some way). Often, when reviewing the relevant industry, the decisions discuss general trends in the market environment. For example, the EC decision in the industrial bags cartel (Case COMP/38354 – Industrial bags, 11/30/2005.) states that "[d]uring this period demand for industrial bags as a whole has

[22] COMP/39.396 – Calcium carbide and magnesium based reagents for the steel and gas industries, 7/22/2009, Commission Decision (¶54).

[23] COMP/38.432 – Professional Videotape, 11/20/2007, Commission Decision (¶58).

stagnated or even declined." Unless a change in market conditions is in some way linked to the cartel's formation or the beginning of meetings between producers, this change in market conditions is not recorded.

I do not attempt to distinguish between sudden and gradual changes in, for example, demand. The wording in the EC decisions typically did not clearly indicate whether changes in market conditions were abrupt or more gradual. Gradual deteriorations in market conditions are analyzed formally in Technical Appendix F. Additionally, I do not attempt to determine if a change in market conditions was expected to be permanent or transitory.

The emergence of overcapacity could also be classified as a deterioration in market conditions in Table 1. However, without additional information regarding the cause of excess capacity in the industry, it is difficult to determine if overcapacity should be categorized as a deterioration or improvement in market conditions. For example, overcapacity may be the result of reduced output levels caused by relaxed competition between firms (which could be considered an improvement in market conditions).

I find, in Table 1, a slightly higher likelihood of a cartel being caused by a deterioration in market conditions than prior literature. This difference is likely caused by my focus on changes in market conditions that the EC decision explicitly links to the formation of the cartel, rather than any changes in market environment preceding the cartel's formation. Additionally, Herold and Paha (2018) do not record changes in marginal cost which Table 1 reveals are a frequent deterioration in market conditions in my sample.

When low prices are discussed without reference to a price war, the cause is recorded as low prices. When a price war is explicitly discussed as a cause of the cartel's formation, I record the cause as a price war and not as low prices (recognizing that price wars, by definition, result in relatively low prices). Low prices may also represent a deterioration in market conditions in some cases. However, low prices could also be the result of an improvement in market conditions (a technological advance or reduction in input prices).

There are two important caveats to the empirical analysis presented in this study. First, as mentioned previously, firms anticipating a fine for price fixing activity have an incentive to argue that their industry is/was in a state of turmoil or crisis in order to receive a fine reduction from the European Commission. The Commission has, in some cases, considered an industry's poor financial state as an extenuating circumstance when setting fines. As a result, firms may exaggerate problems in their industry, and de-emphasize prosperous market conditions, when arguing their case with the Commission. If this occurs, the above results may represent an overestimate of the frequency of deteriorations in market conditions as a cause of cartel formation. Second, as with any empirical analysis of illegal cartels, the above sample includes only detected cartels and is therefore not necessarily representative of the entire population of cartels.

C Robustness of Assumption 4(i)

In this section, I examine the robustness of results to Assumption 4(i). Specifically, I consider alternative values for the initial reference point r_1. In place of Assumption 4(i), I place the following restriction on r_1 throughout this section.

Assumption 7. $r_1 > u(\pi_a^N)$

Assumption 7 ensures the Nash equilibrium utility level under regime a is perceived as a loss. If $r_1 \leq u(\pi_a^N)$, then the change in market conditions is not perceived as a loss by managers and the effects highlighted in this study are not applicable. Recall that $r_1 < u(\pi_a^M)$ by Assumption 4(ii). Thus, Assumption 7 and Assumption 4(ii) together imply that $r_1 \in \big(u(\pi_a^N), u(\pi_a^M)\big)$.

Lemmas 4-7 continue to hold under Assumption 7 with only slight modification to the proof of Lemma 7 in Appendix A. Specifically, Equation (10) becomes

$$u(\pi_a^N) < r_1 < u(\pi_a^M) \tag{12}$$

where the first inequality follows from Assumption 7 and the second inequality follows from Assumption 4(ii). Thus, a cartel forms and $V_a^C(l) = V_a^M(l)$ when $l \geq \bar{l} = \frac{u(\pi_a^{DM})+\delta\frac{u(\pi_a^N)}{1-\delta}-\frac{u(\pi_a^M)}{1-\delta}}{\delta L(r_1-u(\pi_a^N))}$ under regime a. Proposition 1, Proposition 2 and their corresponding proofs from Appendix A hold without modification under Assumption 7.

Next, I establish that pricing results from Section 7 continue to hold when Assumption 7 holds (in place of Assumption 4(i)). The following condition generalizes Condition 3 from the main text to reflect Assumption 7, and is equivalent to Condition 3 when $r_1 = u(\pi_b^N)$ (i.e., Assumption 4(i) holds).

Condition 4. i) $\delta < \frac{u(\pi_b^D(x_a^M))-u(\pi_b(x_a^M))}{u(\pi_b^M)-\big(u(\pi_b^N)-lL(r_1-u(\pi_b^N))\big)+u(\pi_b^D(x_a^M))-u(\pi_b(x_a^M))}$, and
ii) $x_b^N < x_a^M$.

Let $\{x_{b,t}\}_{t=1}^{\infty}$ and $\{x_{a,t}\}_{t=1}^{\infty}$ denote optimal price paths under regime b and regime a, respectively. The following proposition establishes that a deterioration in market state can result in higher collusive prices when managers are sufficiently loss averse.

Proposition 5. *Suppose Condition 4 holds. Then, there exists an* $\bar{l}$ *such that if* $l \geq \bar{l}$*, then there exists a* $T \in \{1, 2 \dots\} \cup \{\infty\}$ *such that* $x_{b,t} < x_{a,t}$ *for all* $t \leq T$.

Proof. Suppose $l \geq \bar{l} = \frac{u(\pi_a^{DM})+\delta\frac{u(\pi_a^N)}{1-\delta}-\frac{u(\pi_a^M)}{1-\delta}}{\delta L(r_1-u(\pi_a^N))}$. By Lemma 7, $\{x_a^M\}_{t=1}^{\infty} \in \Psi_a(l)$ and a cartel forms under regime a and $l \geq \bar{l}$. $\{x_a^M\}_{t=1}^{\infty}$ is the unique optimal price path by Lemma 5. Thus, it suffices to show that

$x_{a,1} = x_a^M > x_{b,1}$ when $l \geq \bar{l}$. First, note that if a cartel does not form before the change, then $x_{b,t} = x_b^N$ for all t and, by Condition 4(ii), $x_{a,1} = x_a^M > x_b^N = x_{b,1}$. For the remainder of the proof, assume a cartel forms under regime b.

Note that $u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M)) \geq 0$ by Assumption 3(iii). Suppose $u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M)) = 0$. By Assumption 6(iv), $u(\pi_b^D(x)) - u(\pi_b(x))$ is strictly increasing in x for all $x \geq x_b^N$. Thus, $u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M)) = 0$ implies $u(\pi_b^D(x)) < u(\pi_b(x))$ for $x < x_a^M$ which contradicts Assumption 3(iii). Thus, $u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M)) > 0$. As a cartel forms under regime b (by supposition), $u(\pi_b^M) \geq u(\pi_b^N) - lL(r_1 - u(\pi_b^N))$ must hold.[24] $u(\pi_b^D(x_a^M)) > u(\pi_b(x_a^M))$ and $u(\pi(x_b^M; a_0)) \geq u(\pi_b^N) - lL(r_1 - \pi_b^N)$ imply that the numerator and denominator of the expression in Condition 4(i) are positive.

Let $\hat{r}_1 = r_1$ and let $\hat{r}_{t+1} = m(\hat{r}_t, u(\pi_b^N))$. If $r_1 > u(\pi_b^N)$, then

$$\frac{u(\pi_b^N) - lL(r_1 - u(\pi_b^N))}{1-\delta} \leq V_b^N = \sum_{t=1}^{\infty} \delta^{t-1} \left[u(\pi_b^N) - lL(\hat{r}_t - u(\pi_b^N))\right].$$

as $\hat{r}_t \leq r_1$ for all t by the definition of m. If $r_1 \leq u(\pi_b^N)$, then

$$\frac{u(\pi_b^N) - lL(r_1 - u(\pi_b^N))}{1-\delta} = \frac{u(\pi_b^N)}{1-\delta} = V_b^N = \sum_{t=1}^{\infty} \delta^{t-1} u(\pi_b^N)$$

as $\hat{r}_t \leq u(\pi_b^N)$ for all t by the definition of m. Thus,

$$\frac{u(\pi_b^N) - lL(r_1 - u(\pi_b^N))}{1-\delta} \leq V_b^N. \tag{13}$$

Let

$$G(x) \equiv u(\pi_b(x)) + \delta \frac{u(\pi_b^M)}{1-\delta} - \left(u(\pi_b^D(x)) + \delta V_b^N\right).$$

[24] Otherwise, $V_b^C \leq \frac{u(\pi_b^M)}{1-\delta} < \frac{u(\pi_b^N) - lL(r_1 - u(\pi_b^N))}{1-\delta} \leq V_b^N$ and a cartel does not form.

Condition 4(i) states

$$
\begin{aligned}
& \frac{u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M))}{u(\pi_b^M) - \big(u(\pi_b^N) - lL(r_1 - u(\pi_b^N))\big) + u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M))} > \delta \\
\iff \quad & \delta\left[u(\pi_b^M) - \big(u(\pi_b^N) - lL(r_1 - u(\pi_b^N))\big) + u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M))\right] \\
& \qquad < u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M)) \\
\iff \quad & \delta\left[u(\pi_b^M) - \big(u(\pi_b^N) - lL(r_1 - u(\pi_b^N))\big) + u(\pi_b^D(x_a^M)) - u(\pi_b(x_a^M))\right] \\
& \qquad + u(\pi_b(x_a^M)) - u(\pi_b^D(x_a^M)) < 0 \\
\iff \quad & u(\pi_b(x_a^M)) - u(\pi_b^D(x_a^M)) - \delta\left(u(\pi_b(x_a^M)) - u(\pi_b^D(x_a^M))\right) \\
& \qquad + \delta\left(u(\pi_b^M) - \big(u(\pi_b^N) - lL(r_1 - u(\pi_b^N))\big)\right) < 0 \\
\iff \quad & (1-\delta)\left(u(\pi_b(x_a^M)) - u(\pi_b^D(x_a^M))\right) \\
& \qquad + \delta\left(u(\pi_b^M) - \big(u(\pi_b^N) - lL(r_1 - u(\pi_b^N))\big)\right) < 0 \\
\iff \quad & u(\pi_b(x_a^M)) - u(\pi_b^D(x_a^M)) + \delta\frac{u(\pi_b^M) - \big(u(\pi_b^N) - lL(r_1 - u(\pi_b^N))\big)}{1-\delta} < 0 \\
\iff \quad & u(\pi_b(x_a^M)) - u(\pi_b^D(x_a^M)) + \delta\frac{u(\pi_b^M)}{1-\delta} - \delta\frac{u(\pi_b^N) - lL(r_1 - u(\pi_b^N))}{1-\delta} < 0 \\
\implies \quad & u(\pi_b(x_a^M)) - u(\pi_b^D(x_a^M)) + \delta\frac{u(\pi_b^M)}{1-\delta} - \delta V_b^N < 0 \\
\implies \quad & G(x_a^M) < 0 \qquad (14)
\end{aligned}
$$

where the second to last line follows from Equation (13). $G(x)$ is strictly decreasing in x for $x \geq x_b^N$ by Assumption 6(iv). Thus, Equation (14) implies that $G(x) < 0$ for all $x \geq x_a^M > x_b^N$ (where the last inequality follows from Condition 4(ii)).

If $x_{b,1} \geq x_a^M$, then

$$
\begin{aligned}
& u(\pi_b(x_{b,1}); r_1) + \sum_{t=2}^{\infty} \delta^{t-1} u(\pi_b(x_{b,t}); r_t) \\
& \leq u(\pi_b(x_{b,1})) - lL\left(r_1 - u(\pi_b(x_{b,1}))\right) + \sum_{t=2}^{\infty} \delta^{t-1} u(\pi_b(x_{b,t})) \\
& \leq u(\pi_b(x_{b,1})) - lL\left(r_1 - u(\pi_b(x_{b,1}))\right) + \delta\frac{u(\pi_b^M)}{1-\delta} \\
& < u(\pi_b^D(x_{b,1})) - lL\left(r_1 - u(\pi_b(x_{b,1}))\right) + \delta V_b^N \\
& \leq u(\pi_b^D(x_{b,1})) - lL\left(r_1 - u(\pi_b^D(x_{b,1}))\right) + \delta V_b^N \\
& = u(\pi_b^D(x_{b,1}); r_1) + \delta V_b^N \qquad (15)
\end{aligned}
$$

where the first inequality follows from the fact that $u(\pi; r) \leq u(\pi)$ (Assumption 1). The second inequality follows from Assumption 3(ii). The third inequality follows from $G(x_{b,1}) < 0$ for $x_{b,1} \geq x_a^M$. The fourth inequality follows from Assumption 3(iii) and Assumption 1(iv). (15) implies that the ICC in period 1 is not satisfied if $x_{b,1} \geq x_a^M$. Thus, $x_{b,1} < x_a^M$. □

D Additional Results and Examples

D.1 Naive Managers

Research in behavioral economics and psychology suggests that people are relatively poor predictors of their future tastes. Particularly, people tend to underestimate the extent that their future tastes will differ from their current preferences. Loewenstein, O'Donoghue and Rabin (2003) refers to this cognitive bias as "projection bias." In general, humans adjust to changes in their circumstances, but tend to underestimate their own ability to become acclimatized to gains or losses. A number of experiments[25] have compared individuals' predictions of how a major life event or change in fortune will impact their subjective well-being to actual reports of individuals who have experienced the event. Subjects tend to overestimate the impact of changes in circumstance on their well-being. Kahneman and Snell (1992) find almost no connection between individual's predictions of their changes in preference and their actual changes in preference. For example, Jepson, Loewenstein and Ubel (2001) asked patients waiting for a kidney transplant to predict their (subjective) quality of life one year later if they receive a transplant and if they do not receive a transplant. They then surveyed the same individuals a year later and found that "[p]atients who received transplants predicted a higher quality of life than they ended up reporting, and those who did not predicted a lower quality of life than they ended up reporting" (Jepson, Loewenstein and Ubel, 2001). These results are consistent with individuals failing to entirely anticipate how they would psychologically adapt to receiving or not receiving a transplant. Barkan and Busemeyer (1999) specifically analyze individuals' ability to predict changes in their reference points and, consistent with projection bias, find that individuals do not correctly anticipate changes in their reference point over time.

In the current setting, projection bias suggests that managers may fail to fully anticipate the extent that their reference points will adjust to changes in utility. Naive managers suffer from projection bias. When considering joining a cartel, a naive manager views the relatively high utility levels during collusion as a large gain and may underestimate the extent to which they will become accustomed to those utility levels. Conversely, a naive manager considering defecting from a cartel may underestimate the extent that they will become accustomed to lower utility levels in the punishment phase.

[25] See Helson (1964); Frederick and Loewenstein (1999); Loewenstein, O'Donoghue and Rabin (2003); Wilson and Gilbert (2005) and Wilson and Gilbert (2005) for examples and reviews.

To explore the robustness of results to the possibility of naive managers, suppose that managers do not anticipate any changes in their reference points. Formally, if a manager's reference point in period t is r_t, the manager believes its reference point will remain fixed at r_t in all future periods. Thus, colluding managers set prices expecting their current reference points to prevail in all future periods. In the subsequent period, reference points adjust according to $r_{t+1} = m(r_t, u_t)$. Managers then set prices expecting the updated reference point r_{t+1} to prevail in all future periods. Note that the pricing decisions of naive managers may be dynamically inconsistent. Specifically, a manager may regret a pricing decision made in period t (under the expectation that the reference point in all future periods would remain fixed at r_t) in period $t+1$ because the manager's reference point unexpectedly updates between period t and period $t+1$. While the possibility of dynamic inconsistency complicates the analysis of pricing decisions, the choice to form or not form a cartel can be analyzed more tractably.[26] As in the main text, a cartel forms under regime i if $V_i^C(r_1) > V_i^N(r_1)$ where

$$V_i^C(r_1) = \max_{\{x_t\}_{t=1}^{\infty} \in \Psi} \sum_{t=1}^{\infty} \delta^{t-1} u(\pi_i(x_t); r_1) \tag{16}$$

and

$$\Psi_i(r_1) = \left\{ \{x_t\}_{t=1}^{\infty} : x_t \in \Omega \text{ and } \sum_{t=T}^{\infty} \delta^{t-T} u(\pi_i(x_t); r_1) \geq u(\pi_i^D(x_T); r_1) + \delta V_i^N(r_1) \quad \text{for all } T \in \{1, 2 \dots\} \right\}. \tag{17}$$

Note that, unlike in the main text, $V_i^C(r_1)$ represents the manager's *perceived* discounted present value of collusion in the initial period. Crucially, the manager evaluates all utility levels relative to their current reference point r_1, failing to anticipate future adjustments to their reference point. Thus, whether the manager wishes to form a cartel (in period 1) depends only on the manager's initial reference point. In period 1, the optimization problem in Equation (16) is equivalent to the problem in the main text with $m(r, u) = r$. Thus, the results of the main text regarding cartel formation (in Section 5 and 6) continue to hold for naive managers.

The dynamic inconsistency of naive managers creates an additional complication: managers may, in theory, wish to disband the cartel after their reference points (unexpectedly) change in future periods. Alternatively, no incentive compatible price path may exist in a future period, in light of updated reference points, when a cartel has previously formed. However, dynamic inconsistency of this kind is unlikely to arise

[26]Note that the price path decided upon by managers in period 1 (when forming the cartel) may be revised in subsequent periods when the managers' reference points unexpectedly adjust. Thus, the solution to (16) in the initial period is a sequence of prices which may in fact not occur.

in the present setting. To see this, note that reference points typically increase over time as managers become accosted to higher profits during collusion. As the reference point increases, the punishment for cheating on the cartel is perceived as an increasingly large loss and, as a result, incentives to defect decline over time. Additionally, disbanding the cartel and returning to Nash competition is perceived as an increasingly large loss as reference points adjust upwards. Thus, as managers become accustomed to collusive profits, the cartel becomes more stable, expanding the set of price paths which satisfy the incentive compatibility constraints and reducing the perceived utility of disbanding the cartel and returning to Nash competition.

D.2 S-Shaped Utility and Assumption 1

In this subsection, I explore when a utility function that satisfies Assumption 1 will display risk aversion over gains and risk seeking preferences over losses (i.e., an "S-shape").

For the purposes of this subsection, it facilitates the analysis to define a minimum possible per-period profit level $\underline{\pi}$ and restrict attention to $\pi \geq \underline{\pi}$.[27] For example, $\underline{\pi} < 0$ may represent the maximum possible per-period loss that a manager can incur. For consistency, the reference point r is assumed to satisfy $r > u(\underline{\pi})$. Under Assumption 1, a utility function $u(\pi; r, l)$ is risk averse over gains if base utility $u(\pi)$ is strictly concave for all π such that $u(\pi) > r$. $u(\pi; r, l)$ is risk seeking over losses if $u(\pi; r, l)$ is strictly convex for all π such that $u(\pi) < r$. Assuming $u(\pi; r, l)$ is differentiable in π and L is differentiable, $u(\pi; r, l)$ is strictly convex when $u(\pi) < r$ if

$$u''(\pi; r, l) = u''(\pi) - lL''(r - u(\pi)) \left[u'(\pi)\right]^2 + lL'(r - u(\pi))u''(\pi) > 0$$

for all $\pi \geq \underline{\pi}$. The above inequality holds and $u(\pi; r, l)$ is convex over losses if base utility is linear (i.e., $u''(\pi) = 0$) and L is strictly concave (i.e., $L'' < 0$).

To further explore when utility functions satisfying Assumption 1 are S-shaped, I consider a power utility function (Maggi, 2004) and power loss function for the remainder of this section. Specifically, suppose $u(\pi) = \pi^b$ where $b \in (0, 1)$ and $L(x) = x^a$ where $a \in (0, 1)$. To ensure $u(\pi)$ is real, suppose $\underline{\pi} \geq 0$.

Lemma 8. *$u(\pi; r, l) = u(\pi) - lL(r - u(\pi))$ where $u(\pi) = \pi^b$, $b \in (0, 1)$, $L(x) = x^a$ and $a \in (0, 1)$ is S-shaped if*

i) l is sufficiently large and $\left[(1-a)b + (1-b)\right]\underline{\pi}^b > (1-b)r$, or

ii) b is sufficiently close to 1 and $l > 0$.

Proof. Clearly $u(\pi; r, l)$ exhibits risk aversion over gains as $u(\pi; r, l) = u(\pi) = \pi^b$ when $u(\pi) > r$ and

[27] For example, power utility functions such as $u(\pi) = \pi^b$ for $b \in (0, 1)$ may be undefined for $\pi < 0$.

$b \in (0,1)$. The remainder of the proof establishes that $u(\pi; r, l)$ is convex when $u(\pi) < r$. It suffices to show that the second derivative of $u(\pi; r, l)$ is positive when $u(\pi) < r$:

$$u''(\pi; r, l) = b(b-1)\pi^{b-2} - la(a-1)(r-u(\pi))^{a-2}\left[b\pi^{b-1}\right]^2$$

$$+la(r-u(\pi))^{a-1}b(b-1)\pi^{b-2}$$

$$= b(b-1)\pi^{b-2} - la\pi^{b-2}b\left[(a-1)(r-u(\pi))^{a-2}b\pi^b - (r-u(\pi))^{a-1}(b-1)\right] > 0. \tag{18}$$

i) $[(1-a)b + (1-b)]\,\underline{\pi}^b > (1-b)r$ implies that

$$[(1-a)b + (1-b)]\,\underline{\pi}^b - (1-b)r > 0.$$

As $b \in (0,1)$ and $a \in (0,1)$, it follows that

$$[(1-a)b + (1-b)]\,\pi^b - (1-b)r > 0$$

for all $\pi \geq \underline{\pi}$. Thus,

$$[(1-a)b + (1-b)]\,\pi^b - (1-b)r > 0$$

$$\implies (1-a)b\pi^b - (1-b)\left(r - \pi^b\right) > 0.$$

As $\pi^b < r$, it follows that

$$(1-a)b\pi^b\left(r-\pi^b\right)^{a-2} - (1-b)\left(r-\pi^b\right)^{a-1} > 0$$

$$\implies (a-1)b\pi^b\left(r-\pi^b\right)^{a-2} - (b-1)\left(r-\pi^b\right)^{a-1} < 0.$$

The above inequality implies that the expression in brackets on the left hand side of (18) is negative. Thus, the second term on the left hand side of (18) is positive. Thus, the inequality in (18) holds if l is sufficiently large.

ii) As $b \to 1$, the expression on the left hand side of (18) approaches

$$-la\pi^{-1}(a-1)(r-u(\pi))^{a-2}\pi = la(1-a)(r-u(\pi))^{a-2} \geq la(1-a)(r-u(\underline{\pi}))^{a-2} > 0$$

where the second to last inequality follows from the fact that $a \in (0,1)$. $\square$

Lemma 8 establishes that $u(\pi; r, l)$ is S-shaped when one of two conditions is met. The first condition holds when the manager is sufficiently loss averse. When managers are highly loss averse, the curvature of the utility function is determined primarily by $L(x)$ rather than base utility. As $L(x) = x^a$ where $a \in (0, 1)$, $-L(x)$ is convex and the manager is risking over losses. The concavity of base utility ensures that managers are risk averse over gains. The second condition holds when base utility is sufficiently linear (i.e., b is close to 1). When base utility is highly linear, the curvature of the utility function (for losses) is driven primarily by the curvature of $L(x)$. As $-L(x)$ is convex, the manager is risk seeking over losses.

D.3 Cartel Formation Conditions

$\{x_t\}_{t=1}^{\infty} \in \Psi$ does not imply that the payoff from collusion with a price path $\{x_t\}_{t=1}^{\infty}$ exceeds the payoff from Nash competition. Thus, both conditions are required. To see this, suppose $l = 0$ (i.e., managers are loss neutral), $u(\pi) = \pi$ (i.e., linear utility) and the cartel sets a price x_1 in the first period and a price x in all other periods. Let π^N denote Nash equilibrium profits. $\{x_1, x, x \dots\} \in \Psi$ if

$$\pi(x_1) + \delta \frac{\pi(x)}{1-\delta} \geq \pi^D(x_1) + \frac{\delta}{1-\delta}\pi^N$$

and

$$\frac{\pi(x)}{1-\delta} \geq \pi^D(x) + \frac{\delta}{1-\delta}\pi^N.$$

Both inequalities hold, for example, if $\pi(x) \geq \pi^N$, $\pi^D(x) = \pi(x)$ and $\pi^D(x_1) = \pi(x_1)$.[28] $V^C < V^N$ if

$$V^C = \pi(x_1) + \delta \frac{\pi(x)}{1-\delta} < \frac{\pi^N}{1-\delta} = V^N$$

or

$$\pi(x_1) < \frac{1}{1-\delta}\left(\pi^N - \delta \pi(x)\right)$$

which holds, for example, when δ is sufficiently small and $\pi(x_1) < \pi^N$.

D.4 Sufficient Condition for $V_b^C(l) < V_a^M(l)$ (Condition 2(ii))

In this subsection, let the market state parameter in firms' profit functions be $a \in \Gamma \subset \mathbb{R}$. Thus, $\pi^M(a)$ and $\pi^N(a)$ denote monopoly and Nash equilibrium firm profits, respectively, when the market state is a. Additionally, let a_0 denote the value of the market state parameter prior to the deterioration in market conditions. Thus, $\pi_b^M = \pi^M(a_0)$ and $\pi_b^N = \pi^N(a_0)$. Analogously, let a_1 denote the value of the market

[28] Assumption 3(iii) ensures $\pi^D(x) \geq \pi(x)$ and $\pi^D(x_1) \geq \pi(x_1)$.

state parameter after the deterioration in market conditions. Thus, $\pi_a^M = \pi^M(a_1)$ and $\pi_a^N = \pi^N(a_1)$. Let $\delta_b^M = \frac{u(\pi_b^{DM})-u(\pi_b^M)}{u(\pi_b^{DM})-u(\pi_b^N)}$. δ_b^M represents the smallest discount factor such that $\{x_b^M\}_{t=1}^{\infty} \in \Psi_b(0)$ (i.e., the smallest discount factor such that the monopoly price path is ICC when $l = 0$).

Proposition 6. *Suppose*

i) $\pi^M(a)$ *is continuous in* a *for all* $a \in [c, d] \subset \Gamma$ *where* $a_0 \in (c, d)$*,*

ii) $\delta < \delta_b^M$*, and*

iii) $\pi_b^M > \pi_b^N$*.*

Then, there exists an $\epsilon > 0$ *such that* $V_b^C(l) < V_a^M$ *for* a_1 *such that* $|a_0 - a_1| < \epsilon$*.*

Proof. Note that $V_b^M = \frac{u(\pi_b^M)}{1-\delta}$ by $\pi_b^M > \pi_b^N$ (Assumption (iii)), Assumption 4(i) and the definition of m. First, I show that $V_b^C(l) < V_b^M$ when $\delta < \delta_b^M$. $V_b^C(l) = V_b^M$ if and only if $x_t = x_b^M$ for all t (Assumption 3(ii)). The IC in the first period associated with this path (under regime b) is

$$\frac{u(\pi_b^M)}{1-\delta} \geq u(\pi_b^{DM}) + \frac{\delta}{1-\delta} u(\pi_b^N)$$

which is violated as $\delta < \delta_b^M = \frac{u(\pi_b^{DM})-u(\pi_b^M)}{u(\pi_b^{DM})-u(\pi_b^N)}$. Thus, $V_b^C(l) < V_b^M = \frac{u(\pi_b^M)}{1-\delta}$.

As $\pi^M(a)$ is continuous in a for all $a \in [c, d] \subset \Gamma$, $a_0 \in (c, d)$, $V_b^C(l) < \frac{u(\pi_b^M)}{1-\delta} = \frac{u(\pi^M(a_0))}{1-\delta}$ and $\pi_b^M > \pi_b^N$, there exists an ϵ such that for all a_1 such that $|a_0 - a_1| < \epsilon$, i) $\frac{u(\pi_a^M)}{1-\delta} = \frac{u(\pi^M(a_1))}{1-\delta} > V_b^C(l)$, ii) $\pi_a^M > \pi_b^N$, and iii) $a_1 \in [c, d]$. Thus, for a_1 such that $|a_0 - a_1| < \epsilon$, $V_b^C(l) < \frac{u(\pi_a^M)}{1-\delta} = V_a^M$ where the equality follows from $\pi_a^M > \pi_b^N$, Assumption 4(i) and the definition of m. □

If $V_a^M > V_b^M$, then Condition 2(ii) holds trivially as $V_a^M > V_b^M \geq V_b^C(l)$.

D.5 Target-based Loss Aversion

The following subsections present two distinct models illustrating how the loss averse utility functions analyzed in the main text may reflect the incentives of a manager evaluated relative to pre-defined target performance levels (i.e., target-based loss aversion). As Sullivan and Kida (1995) write "[c]orporate managers typically operate in a decision environment that utilizes targets to both motivate and reward managerial performance." Exceeding or failing to meet a target performance level may determine whether a manager receives a bonus (Healy, 1985; Ely, 1991; Willman et al., 2002), is awarded a promotion, or loses their job (Sullivan and Kida, 1995; Merchant and Manzoni, 1989).

D.5.1 Bonuses

In this subsection, I present a model wherein managers receive a bonus if they meet or exceed a performance target. The model illustrates how the loss averse utility function from the main text can also reflect the incentives of a manager subject to a performance-based bonus structure of this kind.

Suppose managerial compensation includes a fixed wage w, a constant fraction $\alpha \in [0,1]$ of the profit that the manager generates and a possible bonus B. Thus, a managers compensation is $w+\alpha\pi+B$ if they receive the bonus and $w+\alpha\pi$ if they do not receive the bonus. The manager receives the performance-based bonus if they meet or exceed a pre-defined performance target π^T. If the manager fails to meet the performance target, then the manager may or may not receive a bonus.[29] A manager generating profit π receives the bonus with probability $\gamma(\pi^T - \pi)$ where $\gamma(x)$ satisfies the following assumption.

Assumption 8. $\gamma(x) : \mathbb{R} \to [0,1]$ *satisfies the following conditions:*

i) $\gamma(x) = 1$ *if* $x \leq 0$*, and*

ii) $\gamma(x)$ *is continuous and strictly decreasing for* $x > 0$*.*

Assumption 8(i) implies that the manager receives the bonus if they reach their performance target. Assumption 8(ii) reflects the fact that managers are less likely to receive their bonus if they perform significantly below their performance target than if the manager performs only slightly below their target.

A manager's expected wage when generating profit π is therefore

$$\begin{aligned} w + \alpha\pi + \gamma(\pi^T - \pi)B &= w + \alpha\pi + B - \left[1 - \gamma(\pi^T - \pi)\right] B \\ &= w + \alpha\pi + B - \left[1 - \gamma\left(\frac{w + \alpha\pi^T + B - (w + \alpha\pi + B)}{\alpha}\right)\right] B. \end{aligned}$$

The above expression is equivalent to the utility function $u(\pi; r, l)$ from the main text with base utility $u(\pi) = w + \alpha\pi + B$, reference point $r = u(\pi^T)$, $L(x) = 1 - \gamma\left(\frac{x}{\alpha}\right)$ and $l = B$. Note that $L(x) = 1 - \gamma\left(\frac{x}{\alpha}\right)$ satisfies Assumption 1(iv) by Assumption 8.

In summary, the loss averse utility function $u(\pi; r, l)$ from the main text can also reflect the incentives of a manager that earns a bonus when profit exceeds a pre-defined target level and, with a certain probability, does not receive the bonus when performing below the target level.

[29] Even if a manager fails to meet a pre-defined performance target, the manager may be awarded a bonus due to, for example, leniency on the part of higher-level executives awarding bonuses, past service to the firm, or strong overall performance by the firm.

D.5.2 Managerial Job Loss

In this subsection, I present a model wherein managers face the risk of job loss if their performance drops below a pre-specified target level (i.e., target-based loss aversion). I demonstrate how the threat of termination can result in payoffs which closely resemble payoffs from the loss averse utility function $u(\pi; r, l)$ employed in the main text.

Suppose managers are evaluated relative to a target profit level π^T. If the manager fails to generate a level of profit that matches or exceeds the target profit level, then the manager may be immediately fired with a positive probability. Formally, if π represents a manager's profit, then the manager faces a risk of termination if $\pi < \pi^T$ or, in terms of base utility, $u(\pi) < u(\pi^T)$. Let $\phi\left(u(\pi^T) - u(\pi)\right)$ denote the likelihood of termination when the manager generates a profit level of $u(\pi)$.[30] $\phi(x)$ satisfies the following assumption.

Assumption 9. $\phi(x) : \mathbb{R} \to [0, 1]$ *satisfies the following assumptions:*

i) $\phi(x) = 0$ *if* $x \leq 0$*, and*

ii) $\phi(x)$ *is continuous and strictly increasing for* $x > 0$*.*

Additionally, suppose that the performance target π^T is set equal to the previous period's profit/utility level. Formally, performance targets (which will act as reference points in the subsequent analysis) update according to $r_{t+1} = m(r_t, u_t) = u_t$. Intuitively, managers face a risk of termination if their performance drops relative to the previous period.

Each period, the manager's performance is evaluated relative to the performance target. If the manager meets or exceeds their performance target, then the manager retains their employment with probability 1 (see Assumption 9(i)). If the manager performs below their target by a margin $u(\pi^T) - u(\pi) > 0$, the manager is fired with probability $\phi\left(u(\pi^T) - u(\pi)\right)$. If the manager maintains their employment, they receive utility $u(\pi)$ and remain employed for the subsequent period. Additionally, the manager's performance target updates according to $r_{t+1} = m(r_t, u_t) = u_t$. If the manager is terminated, then the manager receives a constant, expected utility u_F in all future periods.[31] By Assumption 5 and 4(i), loss aversion typically impacts incentives to collude strictly through managers' payoffs during the punishment phase (i.e., the discounted present value of utility from repeated Nash competition). The manager's discounted present value of utility from repeated Nash competition (i.e., a constant stream of profits π^N) when the performance target is $u(\pi^T)$

[30] If base utility is $u(\pi) = \pi$, then the likelihood of termination depends only on the difference between the target profit level and the actual profit level.

[31] Note that u_F represents the manager's *expectations* regarding his/her utility levels following termination.

is[32]

$$\begin{aligned}
V^N(u(\pi^T)) &= \left(1 - \phi\left(u(\pi^T) - u(\pi^N)\right)\right) \frac{u(\pi^N)}{1-\delta} + \phi\left(u(\pi^T) - u(\pi^N)\right) \frac{u_F}{1-\delta} \\
&= \frac{u(\pi^N)}{1-\delta} - \phi\left(u(\pi^T) - u(\pi^N)\right) \frac{u(\pi^N)}{1-\delta} + \phi\left(u(\pi^T) - u(\pi^N)\right) \frac{u_F}{1-\delta} \\
&= \frac{u(\pi^N)}{1-\delta} - \phi\left(u(\pi^T) - u(\pi^N)\right) \left[\frac{u(\pi^N)}{1-\delta} - \frac{u_F}{1-\delta}\right] \\
&= \frac{1}{1-\delta}\left[u(\pi^N) - \phi\left(u(\pi^T) - u(\pi^N)\right)\left(u(\pi^N) - u_F\right)\right]. \qquad (19)
\end{aligned}$$

For comparison, recall that the payoff from repeated Nash competition in the model of the main text when the reference point is r and $r_{t+1} = m(r_t, u_t) = r_t$ is

$$V^N(r) = \frac{1}{1-\delta}\left[u(\pi^N) - lL(r - u(\pi^N))\right]. \qquad (20)$$

The payoff in (19) is equivalent to the payoff in (20) with $r = u(\pi^T)$, $L(x) = \phi(x)$ and $l = u(\pi^N) - u_F$. Thus, the loss averse utility function characterized in Assumption 1 can capture target-based loss aversion driven by the threat of managerial job loss.

The preceding analysis involves a number of assumptions, particularly those related to the updating of reference points, which may not hold in practice. Broadly, the preceding discussion is intended to illustrate how the utility functions in the main text can, under certain circumstances, reflect loss aversion driven by the threat of job loss.

D.6 Explanation of $r_1 < u(\pi_a^M)$ (Assumption 4(ii))

The below proposition provides a sufficient condition for $r_1 < u(\pi_a^M)$ (Assumption 4(ii)). For the remainder of this subsection, let the market state parameter in firms' profit functions be $a \in \Gamma \subset \mathbb{R}$. Thus, $\pi^M(a)$ denotes monopoly firm profits when the market state is a. Additionally, let a_0 denote the value of the market state parameter prior to the deterioration in market conditions. Thus, $\pi_b^M = \pi^M(a_0)$. Analogously, let a_1 denote the value of the market state parameter after the deterioration in market conditions. Thus, $\pi_a^M = \pi^M(a_1)$.

Proposition 7. *Suppose i) $\pi^M(a)$ is continuous in a for all $a \in [c,d] \subset \Gamma$ where $a_0 \in (c,d)$, and*

ii) $u(\pi_b^M) > r_1 = u(\pi_b^N)$.

Then, $r_1 = u(\pi_b^N) < u(\pi_a^M)$ if $|a_0 - a_1| < \epsilon$ for some $\epsilon > 0$.

[32] Recall that performance targets/reference points update immediately following a change in performance/profit (i.e., $r_{t+1} = m(r_t, u_t) = u_t$). Thus, the risk of termination occurs only during the period of a drop in performance and not during subsequent periods.

Proof. As $u(\pi)$ is strictly increasing in π (Assumption 1(ii)), it suffices to show that $\pi_b^N < \pi_a^M$.

There are two cases:

Case 1: Suppose $\pi_b^M > \pi_a^M$. By the continuity of $\pi^M(a)$ in a (Assumption 3(i)), there exists an $\epsilon > 0$ such that

$$\left|\pi_b^M - \pi_a^M\right| = \left|\pi^M(a_0) - \pi^M(a_1)\right| < \pi_b^M - \pi_b^N$$

if $|a_0 - a_1| < \epsilon$. Then,

$$\begin{aligned} &\left|\pi_b^M - \pi_a^M\right| < \pi_b^M - \pi_b^N \\ \iff\ &\pi_b^M - \pi_a^M < \pi_b^M - \pi_b^N \\ \iff\ &\pi_a^M > \pi_b^N. \end{aligned}$$

Case 2: If $\pi_b^M \leq \pi_a^M$, then $\pi_b^N < \pi_b^M \leq \pi_a^M$ where the first inequality follows from Assumption (ii) in the statement of the proposition. □

E Presence of an Antitrust Authority

In this section, I introduce an antitrust authority, which may detect and penalize managers engaged in cartel activity, into the model of the main text. The timing of the game proceeds as follows when a cartel is active. First, managers set prices. Second, firms earn profits and managers receive utilities. Third, reference points are updated for the following period according to $r_{t+1} = m(r_t, u_t)$ where r_t is the reference point in period t (the current period) and u_t is utility experienced in period t. Fourth, an antitrust authority detects the cartel with probability $\beta \in (0, 1)$. If the cartel is detected, then collusion ceases and managers experience a stream of utilities u_1^P, $u_2^P \ldots$ where u_t^P denotes manager utility t periods after detection. The discounted present value of manager utility immediately after detection is $V^P(\bar{r}_1) = \sum_{t=1}^{\infty} \delta^{t-1}\left(u_t^P - lL(\bar{r}_t - u_t^P)\right)$ where $\bar{r}_1$ is the reference point in the first period following detection (recall that reference points update prior to the cartel's detection) and $\bar{r}_{t+1} = m(\bar{r}_t, u_t^P)$. Manager utility after detection reflects, among other considerations, criminal and civil punishments such as prison sentences, job loss, reputational damage, and lost income due to diminished career prospects. If a cartel is not detected, collusion continues into the next period. Cartels are only detected when collusion is active (i.e., cartels cannot be detected during defection or punishment phases).

A manager's ex-ante expected collusive payoff in the presence of an antitrust authority is

$$
\begin{aligned}
V^C(r_T) &= u(\pi_T; r_T) + \delta(1-\beta)V^C(r_{T+1}) + \beta\delta V^P(r_{T+1}) \\
&= u(\pi_T; r_T) + \delta(1-\beta)\left[u(\pi_{T+1}; r_{T+1}) + \delta(1-\beta)V^C(r_{T+2}) + \beta\delta V^P(r_{T+2})\right] + \beta\delta V^P(r_{T+1}) \\
&\ \ \vdots \\
&= \sum_{t=T}^{\infty}[\delta(1-\beta)]^{t-T}u(\pi_t; r_t) + \beta\delta\sum_{t=T+1}^{\infty}[\delta(1-\beta)]^{t-(T+1)}V^P(r_t)
\end{aligned}
$$

where reference points update according to m as in the main text. The monopoly payoff is[33]

$$
V^M(r_T) = \sum_{t=T}^{\infty}[\delta(1-\beta)]^{t-T}u(\pi^M; r_t) + \beta\delta\sum_{t=T+1}^{\infty}[\delta(1-\beta)]^{t-(T+1)}V^P(r_t).
$$

Nash payoff is unchanged. The following assumption is assumed to hold for the remainder of this section.

Assumption 10. *i)* $\beta \leq \frac{1}{\delta l}$,
ii) $u_t^P \geq \underline{u}^P$ *for all* t *for some* $\underline{u}^P \in \mathbb{R}$, *and*
iii) $u_t^P < u(\pi_a^M)$ *for all* t.

Assumption 10(i) requires that the probability of detection is sufficiently small relative to the degree of loss aversion. If detection by an antitrust authority is particularly likely, then the following results will not hold as a cartel will never form due to the risk of penalization. Assumption 10(ii) ensures a lower bound on manager utilities following detection by an antitrust authority. Assumption 10(iii) ensures that managers earn a higher level of utility when earning monopoly collusive profits than after detection by an antitrust authority. Put differently, Assumption 10(iii) holds if managers prefer successful collusion to detection.

The following lemma establishes a result analogous to Lemma 4 under the presence of an antitrust authority.

Lemma 9. $\{x_a^M\}_{t=1}^{\infty} \in \Psi_a(l)$ *if*

$$
l \geq \bar{l} = \frac{u(\pi_a^{DM}) + \delta\frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta(1-\beta)} - \beta\delta\frac{V^P(\underline{u}^P) - \frac{1}{\beta\delta}\frac{L(u(\pi_a^M) - \underline{u}^P)}{1-\delta}}{1-\delta(1-\beta)}}{\delta L(r_1 - u(\pi_a^N))}.
$$

[33]Note that a price path of $\{x_i^M\}_{t=1}^{\infty}$ (which results in a payoff of $V_i^M(r_1)$) does not necessarily yield the maximum possible expected payoff under collusion when an antitrust authority is present. For example, managers may wish to reduce their price below the monopoly level in order to reduce their reference points and limit the losses incurred if detected by an antitrust authority. Formally, the term $V^P(r_{T+1})$ in $V^C(r_T)$ may impact the cartel's pricing decisions.

Proof. It suffices to show that

$$\sum_{t=T}^{\infty}\left[\delta\left(1-\beta\right)\right]^{t-T}u(\pi_a^M;r_t,l)+\beta\delta\sum_{t=T+1}^{\infty}\left[\delta\left(1-\beta\right)\right]^{t-(T+1)}V^P(r_t)\geq u(\pi_a^{DM};r_T,l)+\delta V_a^N(r_T) \tag{21}$$

for all $T\in\{1,2,3\ldots\}$ where $r_t=m(r_{t-1},u(\pi_a^M))$ for $t>1$. The definition of m and Assumption 4(ii) imply that $r_t\leq r_{t+1}$ for all t and $r_t\leq u(\pi_a^M)$ for all t. Thus,

$$u(\pi_a^M;r_t,l)=u(\pi_a^M) \tag{22}$$

for all t. Additionally, $u(\pi_a^{DM};r_T,l)=u(\pi_a^{DM})$ by Assumption 3(iii) and $r_t\leq u(\pi_a^M)$ for all t. Thus, the inequalities in 6 simplify to

$$\frac{u(\pi_a^M)}{1-\delta\left(1-\beta\right)}+\beta\delta\sum_{t=T+1}^{\infty}\left[\delta\left(1-\beta\right)\right]^{t-(T+1)}V^P(r_t)\geq u(\pi_a^{DM})+\delta V_a^N(r_T) \tag{23}$$

for all $T\in\{1,2,3\ldots\}$.

Note that reference points after detection never exceed $u(\pi_a^M)$ as $r_t\leq u(\pi_a^M)$ for all t, $u_t^P<u(\pi_a^M)$ for all t (Assumption 10(iii)) and by the definition of m. Thus, for all $T\in\{1,2,3\ldots\}$,

$$\begin{aligned}V^P(r_T)&\geq\sum_{t=1}^{\infty}\delta^{t-1}\left[u_t^P-lL(u(\pi_a^M)-u_t^P)\right]\\&\geq\sum_{t=1}^{\infty}\delta^{t-1}\left[u_t^P-lL(u(\pi_a^M)-\underline{u}^P)\right]\\&=V^P(\underline{u}^P)-l\frac{L(u(\pi_a^M)-\underline{u}^P)}{1-\delta}\\&\geq V^P(\underline{u}^P)-\frac{1}{\beta\delta}\frac{L(u(\pi_a^M)-\underline{u}^P)}{1-\delta}\equiv\underline{V}^P\end{aligned} \tag{24}$$

where the first inequality follows from the fact that $r_t\leq u(\pi_a^M)$ for all t, the fact that post-detection reference points are bounded above by $u(\pi_a^M)$ (shown earlier in the proof) and Assumption 1. The second inequality follows from Assumption 1(iv) and Assumption 10(ii). The equality follows from Assumption 10(ii). The final inequality follows from Assumption 10(i).

Thus, if

$$\frac{u(\pi_a^M)}{1-\delta\left(1-\beta\right)}+\beta\delta\sum_{t=T+1}^{\infty}\left[\delta\left(1-\beta\right)\right]^{t-(T+1)}\underline{V}^P\geq u(\pi_a^{DM})+\delta V_a^N(r_T) \tag{25}$$

holds for all $T \in \{1, 2, 3 \dots\}$, or equivalently,

$$\frac{u(\pi_a^M)}{1-\delta\,(1-\beta)} + \beta\delta \frac{\underline{V}^P}{1-\delta\,(1-\beta)} \geq u(\pi_a^{DM}) + \delta V_a^N(r_T).$$

for all $T \in \{1, 2, 3 \dots\}$, then the inequalities in 23 hold for all $T \in \{1, 2, 3 \dots\}$. Fix $T \in \{1, 2, 3 \dots\}$. Let $\tilde{r}_T = r_T$ and let $\tilde{r}_t = m(\tilde{r}_{t-1}, u(\pi_a^N))$ for $t > T$. Note that

$$\begin{aligned} V_a^N(r_T) = \sum_{t=T}^{\infty} \delta^{t-T} u(\pi_a^N; \tilde{r}_t, l) &\leq u(\pi_a^N; r_1, l) + \sum_{t=T+1}^{\infty} \delta^{t-T} u(\pi_a^N; \tilde{r}_t, l) \\ &\leq u(\pi_a^N; r_1, l) + \sum_{t=T+1}^{\infty} \delta^{t-T} u(\pi_a^N) \\ &= u(\pi_a^N; r_1, l) + \sum_{t=1}^{\infty} \delta^{t} u(\pi_a^N) \\ &= u(\pi_a^N; r_1, l) + \delta \frac{u(\pi_a^N)}{1-\delta} \end{aligned} \tag{26}$$

where the first inequality follows from $\tilde{r}_T = r_T \geq r_1$ for all T and Assumption 1. The second inequality follows from $u(\pi; r, l) \leq u(\pi)$ for all π, r, and l.

$r_1 > u(\pi_a^N)$ by $r_1 = u(\pi_b^N)$ (Assumption 4(i)) and $\pi_a^N < \pi_b^N$ (Assumption 4(iii)). Thus, $L\left(r_1 - u(\pi_a^N)\right) > 0$ by Assumption 1(iv).

Suppose $l \geq \bar{l}$. Then,

$$
\begin{aligned}
& u(\pi_a^{DM}) + \delta V_a^N(r_T) \\
\leq & \, u(\pi_a^{DM}) + \delta u(\pi_a^N; r_1, l) + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
= & \, u(\pi_a^{DM}) + \delta u\left(\pi_a^N\right) - \delta l L\left(r_1 - u(\pi_a^N)\right) + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
\leq & \, u(\pi_a^{DM}) + \delta u\left(\pi_a^N\right) - \delta \bar{l} L\left(r_1 - u(\pi_a^N)\right) + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
= & \, u(\pi_a^{DM}) + \delta u\left(\pi_a^N\right) + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
& - \left[u(\pi_a^{DM}) + \delta \frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta(1-\beta)} - \beta\delta \frac{\underline{V}^P}{1-\delta(1-\beta)} \right] \\
= & \, \delta u\left(\pi_a^N\right) - \delta \frac{u(\pi_a^N)}{1-\delta} + \frac{u(\pi_a^M)}{1-\delta(1-\beta)} + \delta^2 \frac{u(\pi_a^N)}{1-\delta} \\
& + \beta\delta \frac{\underline{V}^P}{1-\delta(1-\beta)} \\
= & \, \delta u\left(\pi_a^N\right) - (1-\delta)\,\delta \frac{u(\pi_a^N)}{1-\delta} + \frac{u(\pi_a^M)}{1-\delta(1-\beta)} + \beta\delta \frac{\underline{V}^P}{1-\delta(1-\beta)} \\
= & \, \frac{u(\pi_a^M)}{1-\delta(1-\beta)} + \beta\delta \frac{\underline{V}^P}{1-\delta(1-\beta)} \\
\leq & \, \frac{u(\pi_a^M)}{1-\delta(1-\beta)} + \beta\delta \sum_{t=T+1}^{\infty} \left[\delta(1-\beta)\right]^{t-(T+1)} V^P(r_t)
\end{aligned} \tag{27}
$$

where the first inequality follows from Equation (26). The first equality follows from the definition of $u(\pi_a^N; r_1, l)$. The second inequality follows from $l \geq \bar{l}$ and $L\left(r_1 - u(\pi_a^N)\right) > 0$. The last inequality follows from Equation (24). Thus, Equation (27) implies the inequalities in 23 are satisfied and $\left\{x_a^M\right\}_{t=1}^{\infty} \in \Psi_a(l)$ when $l \geq \bar{l}$. $\square$

The following lemma establishes a result analogous to Lemma 7 under the presence of an antitrust authority.

Lemma 10. *$V_a^C(l) > V_a^N(l)$ if*

$$
l \geq \bar{l} = \frac{u(\pi_a^{DM}) + \delta \frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta(1-\beta)} - \beta\delta \frac{V^P(\underline{u}^P) - \frac{1}{\beta\delta} \frac{L(u(\pi_a^M) - \underline{u}^P)}{1-\delta}}{1-\delta(1-\beta)}}{\delta L\left(r_1 - u(\pi_a^N)\right)}.
$$

.

Proof. Lemma 9 implies that $V_a^C(l) \geq V_a^M(l)$ when $l \geq \bar{l}$. Thus, it suffices to show $V_a^M(l) > V_a^N(l)$ when $l \geq \bar{l}$.

Note that

$$u(\pi_a^N) < u(\pi_b^N) = r_1 < u(\pi_a^M) \leq u(\pi_a^{DM}) \tag{28}$$

where the first inequality follows from Assumption 4(iii), the equality follows from Assumption 4(i), and the third inequality follows from Assumption 4(ii). The fourth inequality follows from Assumption 3(iii).

Let $r_1 = \tilde{r}_1$ and let $\tilde{r}_t = m(\tilde{r}_{t-1}, u(\pi_a^N))$ for $t > 1$. When $l \geq \bar{l}$,

$$\begin{aligned}
V_a^N(l) &= \sum_{t=1}^{\infty} \delta^{t-1} u(\pi_a^N) - \sum_{t=1}^{\infty} \delta^{t-1} l L(\tilde{r}_t - u(\pi_a^N)) \\
&< u(\pi_a^N) + \sum_{t=2}^{\infty} \delta^{t-1} u(\pi_a^N) - \delta \sum_{t=1}^{\infty} \delta^{t-1} l L(\tilde{r}_t - u(\pi_a^N)) \\
&\leq u(\pi_a^{DM}) + \delta \left[\sum_{t=1}^{\infty} \delta^{t-1} u(\pi_a^N) - \sum_{t=1}^{\infty} \delta^{t-1} l L(\tilde{r}_t - u(\pi_a^N)) \right] \\
&= u(\pi_a^{DM}) + \delta V_a^N(l) \\
&\leq \frac{u(\pi_a^M)}{1 - \delta(1-\beta)} + \beta\delta \sum_{t=2}^{\infty} \left[\delta(1-\beta)\right]^{t-2} V^P(r_t) = V_a^M(l)
\end{aligned}$$

where the first inequality follows from $\delta < 1$, Assumption 1(iv) and $u(\pi_a^N) < u(\pi_b^N) = r_1$ (Assumption 4(i) and Assumption 4(iii)). The second inequality follows from Equation (28). The third inequality follows from Lemma 9. Thus, $V_a^N(l) < V_a^M(l)$. $\square$

Note that there always exists an l that satisfies $l \geq \bar{l}$ and Assumption 10(i) for sufficiently small β. To see this, note that Assumption 10(i) holds if $l \leq \frac{1}{\delta\beta}$. As $\beta \to^+ 0$, $\bar{l} \to \frac{u(\pi_a^{DM}) + \delta \frac{u(\pi_a^N)}{1-\delta} - \frac{u(\pi_a^M)}{1-\delta} + \frac{L(u(\pi_a^M) - u^P)}{(1-\delta)^2}}{\delta L(r_1 - u(\pi_a^N))} < \infty$ and $\frac{1}{\delta\beta} \to \infty$. Therefore, the lower bound on l is finite as $\beta \to^+ 0$ while the upper bound on l approaches infinity.

Proofs of Proposition 1 and 2 follow directly from Lemma 10. Pricing results are more difficult to establish when an antitrust authority is present. This is the case as the threat of detection (and corresponding losses in utility due to penalization) impacts the cartel's pricing decisions. Thus, properly characterizing the optimal price path and comparing price paths before and after the change in market state would require additional assumptions regarding the nature of antitrust penalties and their impact on manager utility (i.e., the stream of utilities $\{u_t^P\}_{t=1}^{\infty}$). However, numerical solutions suggest that pricing dynamics under the presence of an antitrust authority are qualitatively similar to the pricing dynamics outlined in the main text. To illustrate, Figure 5 plots optimal price paths before (red) and after (blue) a reduction in demand. As in the main text, a deterioration in market conditions increases cartel prices in early periods of collusion when managers are sufficiently loss averse.

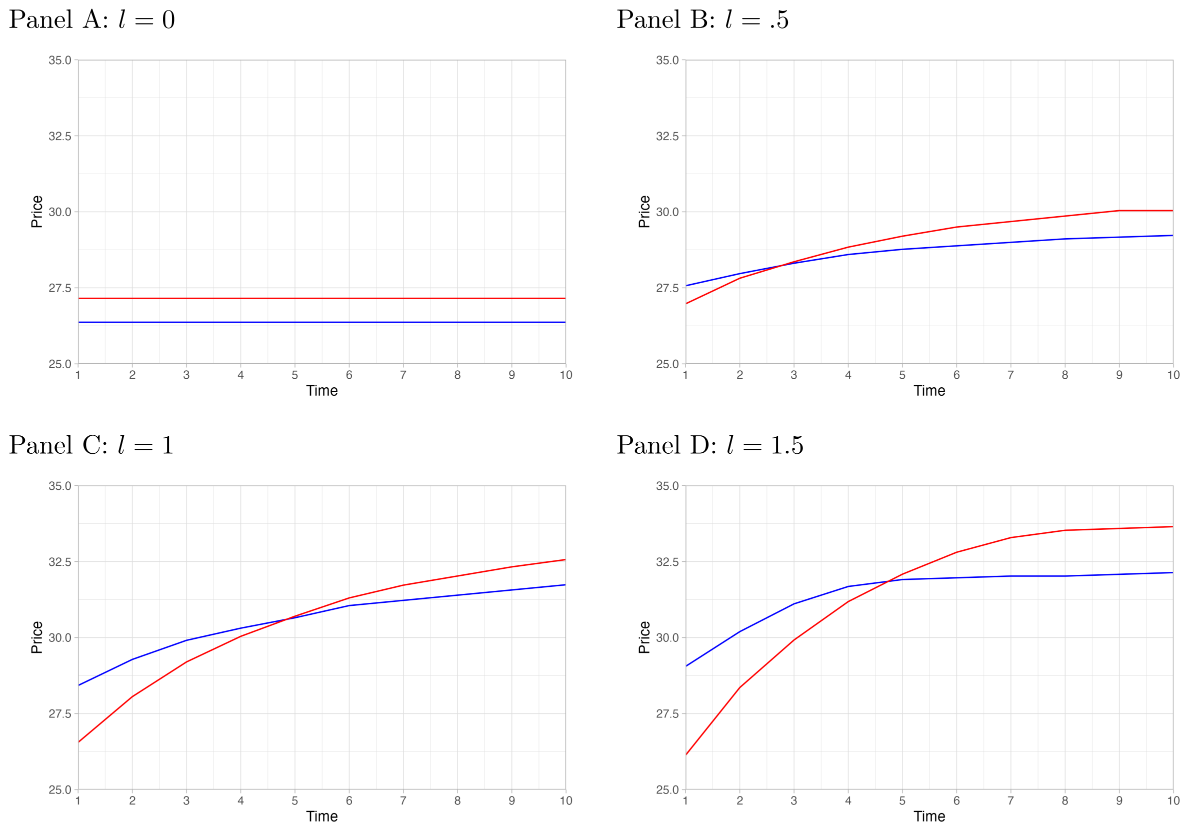


Figure 5: Optimal Price Paths Before (Red) and After (Blue) a 5% Reduction in Demand by Degree of Loss Aversion (with an Antitrust Authority).

Notes: This figure depicts optimal price paths before and after a 5% reduction in the demand parameter a for a variety of degrees of loss aversion with an antitrust authority. Parameters: $b = 2$, $e = 1$, $c = 0$, $N = 4$, $\delta = .25$, $\beta = .125$, $u_t^P = .7 \times \pi_b^N$ for all t, and $\alpha = .75$. Prior to the deterioration in market conditions, $a = 100$. The blue curve depicts the optimal price path after the deterioration in market conditions (i.e., $a = 95$) and the red curve depicts the optimal price path prior to the deterioration in market conditions (i.e., $a = 100$). The Nash equilibrium price before (after) the change is 20 (19). The monopoly price before (after) the change is 50 (47.5).

F Gradual Deteriorations in Market Conditions

In the main text, I restrict attention to abrupt and instantaneous deteriorations in market state. This approach is primarily for analytical tractability and simplicity. Considering gradual changes in market state introduces an additional dynamic variable (i.e., the market state) which complicates the analysis. In this section, I demonstrate that gradual deteriorations in market state can also cause the formation of cartels when managers are loss averse.

Certain types of deteriorations in market state, such as the entry of a new firm or a technological advance that rapidly eliminates a portion of demand for a product, are likely abrupt and are, therefore, best captured by the model presented in the main text. However, other changes in market state may be more gradual (e.g., steadily increasing marginal costs due to rising input prices or a gradual reduction in demand due to changing preferences over time).

Note that a gradual deterioration in market conditions typically hinders the sustainability of collusion between loss neutral agents. For example, gradually declining demand reduces the sustainability of collusion because managers anticipate lower collusive profits in subsequent periods and, thus, have weaker incentives to refrain from cheating on the agreement in the current period (Ivaldi et al., 2007). Put differently, the temptation to cheat in the current period is strong as current demand is high relative to future levels. Thus, when agents are loss neutral, gradual reductions in demand hinder the sustainability of collusion. The subsequent analysis will establish that this is not necessarily the case when managers are loss averse.

Throughout this section, regime b is unchanged from the main text. Under regime a, market conditions begin deteriorating prior to the initial period and, unlike the model in the main text, continue deteriorating gradually in subsequent periods. Managers recognize and anticipate subsequent deteriorations in market state when deciding whether to form a cartel in the initial period.

Formally, let $\pi_{a,t}(x)$ denote profits in period t under regime a when the cartel sets a common price $x \in \Omega$. Assumption 3 holds for each $\pi_{a,t}(x)$. Collusive profits decline over time. Formally, suppose $\pi_{a,t}(x) \geq \pi_{a,t+1}(x)$ for all $t \in \{1, 2 \dots\}$ and $x \in \Omega$. Let $\pi^M_{a,t}$ denote the monopoly profit in period t under regime a.

Let $\pi^D_{a,t}(x)$ denote the profit a manager earns when cheating on collusion in period t when the collusive price is $x \in \Omega$. Assumption 3 is assumed to hold for each $\pi^D_{a,t}(x)$. Assumption 2 holds in each period t. Let $x^N_{a,t}$ denote the Nash equilibrium price in period t. Let $\pi^N_{a,t}$ denote Nash equilibrium profit in period t and suppose $\pi^N_{a,t} \geq \pi^N_{a,t+1}$ for all $t \in \{1, 2 \dots\}$. Thus, the profitability of both collusion and continued competition is gradually declining over time under regime a.

The following assumption replaces Assumption 4(iii).

Assumption 11. $\pi^N_{a,1} < \pi^N_b$

Assumption 11 implies that market conditions have deteriorated prior to the initial period. As Nash equilibrium profits decline over time, Assumption 11 implies that $\pi^N_{a,t} < \pi^N_b$ for all $t \in \{1, 2 \dots\}$. The following assumption ensures that there exists a collusive price path generating a constant stream of utilities that exceeds the initial reference point.

Assumption 12. *There exists a price path* $\{\bar{x}_{a,t}\}_{t=1}^{\infty}$ *such that* $\pi_{a,t}(\bar{x}_{a,t}) = \bar{\pi}_a$ *for all* t *where* $u(\bar{\pi}_a) > r_1$.

$\{\bar{x}_{a,t}\}_{t=1}^{\infty}$ represents a price path wherein managers adjust the collusive price over time, as the market state deteriorates, to ensure a profit of $\bar{\pi}_a$ is earned in each period. Such a price path is not necessarily an optimal price path or incentive compatible, but the existence of such a path is employed in the following proofs. Note that $\bar{\pi}_a \leq \pi^M_{a,t}$ holds, by definition, for all t. Thus, Assumption 12 implies that $\pi^M_{a,t} > r_1$ for all t. Assumption 12 holds if, for example, $\Omega = [c, d]$ for $c < d$ and the size of the deterioration in market state is sufficiently moderate.

Let $\bar{\pi}^D_{a,t} = \pi^D_{a,t}(\bar{x}_{a,t})$ and let $\bar{\pi}^D_a = \sup_t \bar{\pi}^D_{a,t}$.[34] Let $V^N_{a,T}(r_T; l)$ denote the discounted present value of manager utility in period T from competitive play when the current reference point is r_T. Note that $V^N_{a,T}(r_T; l)$ depends on time T as Nash equilibrium payoffs are declining over time.

Lemma 11. *$\{\bar{x}_{a,t}\}_{t=1}^{\infty} \in \Psi_a(l)$ if*

$$l \geq \bar{l} = \frac{u(\bar{\pi}^D_a) + \delta \frac{u(\pi^N_{a,1})}{1-\delta} - \frac{u(\bar{\pi}_a)}{1-\delta}}{\delta L\left(r_1 - u(\pi^N_{a,1})\right)}.$$

Proof. It suffices to show that

$$\sum_{t=T}^{\infty} \delta^{t-T} u(\bar{\pi}_a; r_t, l) \geq u(\bar{\pi}^D_{a,t}; r_T, l) + \delta V^N_{a,T}(r_T) \tag{29}$$

for all $T \in \{1, 2, 3 \ldots\}$ where $r_t = m(r_{t-1}, u(\bar{\pi}_a))$ for $t > 1$. The definition of m and Assumption 12 imply that $r_t \leq r_{t+1}$ for all t and $r_t \leq u(\bar{\pi}_a)$ for all t. Thus,

$$u(\bar{\pi}_a; r_t, l) = u(\bar{\pi}_a) \tag{30}$$

for all t. Additionally, $u(\bar{\pi}^D_{a,t}; r_T, l) = u(\bar{\pi}^D_{a,t})$ by Assumption 3(iii) and $r_t \leq u(\bar{\pi}_a)$ for all t. Therefore, the inequalities in (29) become

$$\sum_{t=T}^{\infty} \delta^{t-T} u(\bar{\pi}_a) \geq u(\bar{\pi}^D_{a,T}) + \delta V^N_{a,T}(r_T)$$

for all $T \in \{1, 2, 3 \ldots\}$. As $\bar{\pi}^D_{a,t} \leq \bar{\pi}^D_a$ for all T, it suffices to show that

$$\sum_{t=T}^{\infty} \delta^{t-T} u(\bar{\pi}_a) \geq u(\bar{\pi}^D_a) + \delta V^N_{a,T}(r_T)$$

for all $T \in \{1, 2, 3 \ldots\}$, or, equivalently,

$$\frac{u(\bar{\pi}_a)}{1-\delta} - u(\bar{\pi}^D_a) \geq \delta V^N_{a,T}(r_T).$$

for all $T \in \{1, 2, 3 \ldots\}$. Fix $T \in \{1, 2, 3 \ldots\}$. Let $\tilde{r}_T = r_T$ and let $\tilde{r}_t = m(\tilde{r}_{t-1}, u(\pi^N_{a,t-1}))$ for $t > T$. Note

[34] The supremum exists if, for example, $\max_x \pi^D_{a,t}(x) \geq \max_x \pi^D_{a,t+1}(x)$ for all $t \in \{1, 2 \ldots\}$.

that

$$V_{a,T}^N(r_T) = \sum_{t=T}^{\infty} \delta^{t-T} u(\pi_{a,t}^N; \tilde{r}_t, l) \leq u(\pi_{a,1}^N; r_1, l) + \sum_{t=T+1}^{\infty} \delta^{t-T} u(\pi_{a,t}^N; \tilde{r}_t, l)$$
$$\leq u(\pi_{a,1}^N; r_1, l) + \sum_{t=T+1}^{\infty} \delta^{t-T} u(\pi_{a,t}^N)$$
$$\leq u(\pi_{a,1}^N; r_1, l) + \sum_{t=T+1}^{\infty} \delta^{t-T} u(\pi_{a,1}^N) \tag{31}$$
$$= u(\pi_{a,1}^N; r_1, l) + \sum_{t=1}^{\infty} \delta^{t} u(\pi_{a,1}^N)$$
$$= u(\pi_{a,1}^N; r_1, l) + \delta \frac{u(\pi_{a,1}^N)}{1-\delta} \tag{32}$$

where the first inequality follows from $\tilde{r}_T = r_T \geq r_1$ for all T and Assumption 1. The second inequality in Equation (32) follows from $u(\pi; r, l) \leq u(\pi)$ (Assumption 1). The third inequality follows from $\pi_{a,t}^N \geq \pi_{a,t+1}^N$.

$r_1 > u(\pi_{a,1}^N)$ by $r_1 = u(\pi_b^N)$ (Assumption 4(i)) and $\pi_{a,1}^N < \pi_b^N$ (Assumption 11). Thus, $L(r_1 - u(\pi_{a,1}^N)) > 0$ by Assumption 1(iv).

Suppose $l \geq \bar{l}$. Then,

$$u(\bar{\pi}_a^D) + \delta V_{a,T}^N(r_T; l) \leq u(\bar{\pi}_a^D) + \delta u(\pi_{a,1}^N; r_1, l) + \delta^2 \frac{u(\pi_{a,1}^N)}{1-\delta}$$
$$= u(\bar{\pi}_a^D) + \delta u\left(\pi_{a,1}^N\right) - \delta l L\left(r_1 - u(\pi_{a,1}^N)\right) + \delta^2 \frac{u(\pi_{a,1}^N)}{1-\delta}$$
$$\leq u(\bar{\pi}_a^D) + \delta u\left(\pi_{a,1}^N\right) - \delta \bar{l} L\left(r_1 - u(\pi_{a,1}^N)\right) + \delta^2 \frac{u(\pi_{a,1}^N)}{1-\delta}$$
$$= u(\bar{\pi}_a^D) + \delta u\left(\pi_{a,1}^N\right) - \left[u(\bar{\pi}_a^D) + \delta \frac{u(\pi_{a,1}^N)}{1-\delta} - \frac{u(\bar{\pi}_a)}{1-\delta}\right] + \delta^2 \frac{u(\pi_{a,1}^N)}{1-\delta} \tag{33}$$
$$= \delta u\left(\pi_{a,1}^N\right) - \delta \frac{u(\pi_{a,1}^N)}{1-\delta} + \frac{u(\bar{\pi}_a)}{1-\delta} + \delta^2 \frac{u(\pi_{a,1}^N)}{1-\delta}$$
$$= \delta u\left(\pi_{a,1}^N\right) - (1-\delta)\,\delta \frac{u(\pi_{a,1}^N)}{1-\delta} + \frac{u(\bar{\pi}_a)}{1-\delta}$$
$$= \frac{u(\bar{\pi}_a)}{1-\delta} \tag{34}$$

where the first inequality follows from (32). The first equality follows from the definition of $u(\pi_{a,1}^N; r_1, l)$. The second inequality follows from $l \geq \bar{l}$. The second equality follows from the definition of $\bar{l}$ and $L(r_1 - u(\pi_{a,1}^N)) > 0$ (shown earlier in the proof). Thus, Equation 34 implies $\{\bar{x}_{a,t}\}_{t=1}^{\infty} \in \Psi_a(l)$ when $l \geq \bar{l}$. □

Lemma 12. $V_a^C(l) > V_a^N(l)$ *(thus, a cartel forms) if*

$$l \geq \bar{l} = \frac{u(\bar{\pi}_a^D) + \delta \frac{u(\pi_{a,1}^N)}{1-\delta} - \frac{u(\bar{\pi}_a)}{1-\delta}}{\delta L \left(r_1 - u(\pi_{a,1}^N)\right)}.$$

Proof. When $l \geq \bar{l}$, $\{\bar{x}_{a,t}\}_{t=1}^{\infty} \in \Psi_a(l)$ (by Lemma 11). It remains to establish that $V_a^C(l) > V_a^N(l)$. Note that

$$u(\pi_{a,1}^N) < u(\pi_b^N) = r_1 < u(\bar{\pi}_a) \tag{35}$$

where the first inequality follows from Assumption 11, the equality follows from Assumption 4(i), and the second inequality follows from Assumption 12. Thus,

$$\begin{aligned} V_a^N(l) &\leq \sum_{t=1}^{\infty} \delta^{t-1} u(\pi_{a,1}^N) \\ &< \sum_{t=1}^{\infty} \delta^{t-1} u(\bar{\pi}_a) \\ &= \frac{u(\bar{\pi}_a)}{1-\delta} \leq V_a^C(l) \end{aligned}$$

where the first inequality follows from the fact that $u(\pi; r, l) \leq u(\pi)$ (Assumption 1) and $\pi_{a,t}^N \geq \pi_{a,t+1}^N$. The second inequality follows from Equation 35. The third inequality follows from Lemma 11. As a result, the cartel forms when $l \geq \bar{l}$. □

Lemma 12 establishes a result analogous to Proposition 1 in the main text. Formally, Lemma 12 demonstrates that a cartel forms during a gradual deterioration in market state when managers are sufficiently loss averse. Recall that the market state begins deteriorating immediately prior to the initial period (see Assumption 11). Managers decide whether to form a cartel in the initial period and, following the initial period, market conditions continue to deteriorate. The effects outlined in the main text also arise when the deterioration in market state is gradual because a reversion to Nash competition is perceived as a loss in both cases. Loss averse managers wish to avoid the painful losses in utility that competitive play would cause. To avoid this outcome, managers turn to collusion.

Note that the analysis in this section has assumed that a cartel forms (or does not form) in the beginning of a gradual deterioration in market state. However, similar effects can arise if the cartel instead decides whether to form after multiple periods of gradual deterioration in the market state. To illustrate, suppose reference points are constant (i.e., $m(r, u) = r$) and suppose managers choose whether to form a cartel in period T. T periods following the beginning of a gradual deterioration in market state, Nash equilibrium profits satisfy $u(\pi_{a,T}^N) \leq u(\pi_{a,1}^N) < r_1$ (where the second inequality follows from Assumption 11). Thus, the cartel's problem is identical to the setting outlined above with period T representing the initial period, period $T+1$ representing the second period etc. The above conclusions will therefore apply without additional modification.

Lemma 12 implies that a deterioration in market conditions can also enhance the value of collusion. Note that the proof of Lemma 12 establishes that $V_a^C(l) \geq \frac{u(\bar{\pi}_a)}{1-\delta}$ when managers are sufficiently loss averse. Thus, if $V_b^C(l) < \frac{u(\bar{\pi}_a)}{1-\delta}$ due to, for example, high fixed costs of collusion or a relatively low discount factor, then a gradual deterioration in market state can also enhance the collusive payoff.

Determining the impact of a gradual change in market state on collusive prices is more challenging. The effects outlined in the main text will, all else equal, result in higher prices (at least, in early periods of collusion). However, other considerations arise when the change in market state is gradual. For example, managers may have a stronger incentive to defect during early periods of collusion because they anticipate relatively low collusive profits in future periods due to the continued deterioration of the market state (Ivaldi et al., 2007). This effect may cause the cartel to reduce prices in early periods of collusion to ensure no manager wishes to defect. Which effect dominates likely depends on, among other factors, the degree of loss aversion and the dynamics of market state changes over time. Generally, if market state changes during collusion are relatively moderate, then it seems likely that the effects outlined in Proposition 2 and Proposition 3 will continue to hold. When changes in the market state during collusion are more extreme or volatile, other considerations may overpower the effects captured in the main text.

G Simulation Results

G.1 Simulation Framework

In this subsection, I provide additional details regarding numerical simulations conducted to generate the figures in the main text (as well as additional figures presented later in this section). All figures reflect outcomes from a setting involving N firms selling symmetrically differentiated products and engaging in price competition. The representative consumer has a utility function of (Singh and Vives, 1984; Harrington, 2004)

$$U(q_1, \ldots q_N) = a\sum_{i=1}^{N} q_i - \left(\frac{1}{2}\right)\left(b\sum_{i=1}^{N} q_i^2 + e\sum_{i=1}^{N}\sum_{j\neq i} q_i q_j\right)$$

where $a > 0$, $b > e > 0$ and q_i is the quantity of firm i's product consumed. If consumers demand a positive quantity of all firms' products, then firm i's demand is

$$D(p_i, p_{-i}) = \left(\frac{a}{b + (N-1)\,e}\right) - \left(\frac{b + (N-2)\,e}{(b + (N-1)\,e)\,(b-e)}\right) p_i + \left(\frac{e\,(N-1)}{(b + (N-1)\,e)\,(b-e)}\right) p_{-i} \tag{36}$$

where p_i is the price set by firm i and p_{-i} is the common price set by all of firm i's rivals. All firms have constant marginal cost c where $0 \le c < a$. In all figures, the base utility function is $u(x) = x$ (i.e., base utility is risk neutral) and $L(x) = x$. Thus, managers have a utility function of

$$u(\pi; r, l) = \pi - l\,(r - \pi)\,.$$

Reference points evolve over time according to $m(r, u) = \alpha r + (1-\alpha)\,u$ where $\alpha \in [0, 1]$. There is a fixed cost of collusion of $F \ge 0$.[35] Thus, collusive profit when all managers charge a common price of p is

$$\pi(p) = (p - c)\,D(p, p) - F.$$

$\pi^D(p)$ denotes profits when defecting from collusion when the collusive price is p.[36] Nash equilibrium profits are

$$\pi^N = \frac{(a-c)^2\,(b + e\,(N-2))\,(b-e)}{(b + (N-1)e)\,(2b + e\,(N-3))^2}.$$

In all figures, the deterioration in market state is either a reduction in the demand intercept (a decrease in a), an increase in marginal cost (an increase in c), or the entry of a new competitor (an increase from N to $N+1$). Note that, as expected, π^N is increasing in a and decreasing in c. Additionally, routine calculations imply that π^N is decreasing in N. Thus, each of the three possible deteriorations in market state are consistent with Assumption 4(iii). Under the differentiated product demand system in Equation (36), a cartel forms for any discount factor unless there is a fixed cost of collusion.[37] The optimal price path and the value of collusion are determined through value function iteration techniques.

To construct figures depicting the critical discount factor as a function of various parameters (i.e., Figure 2, Figure 6 and Figure 7), the following procedure is followed. To illustrate, consider Figure 2 which depicts

[35] Recall that fixed costs of collusion include any moral dis-utilities from participating in an illegal activity, fixed costs of monitoring rivals (e.g., payments made to a third party tasked with monitoring compliance with the collusive scheme), fixed costs involved in concealing collusive activities (including managerial effort) and communicating with other managers involved in the cartel (Klein and Schinkel, 2019), and costs of buying out potential entrants (Ganslandt, Persson and Vasconcelos, 2012). Fixed costs of collusion are paid only when the cartel is active.

[36] In the demand system employed in this section, defection can occur in two distinct ways. The defecting manager can choose a defection price sufficiently low that demand for rival products is 0. Alternatively, the defecting manager can set a higher defection price for which rival demands are positive.

[37] See, for example, the proof of Proposition 2 in Bos et al. (2018).

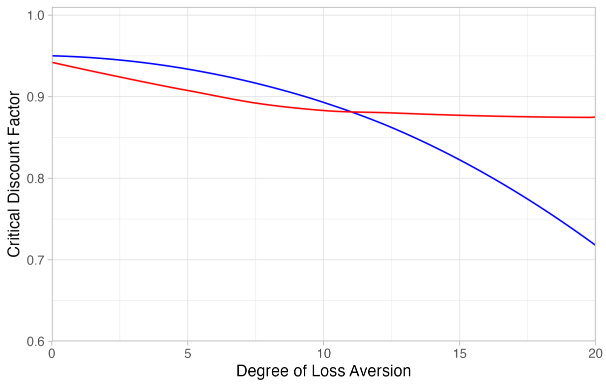
the critical discount factor as a function of the degree of loss aversion l. First, value function iteration is performed on a dense grid of l values and δ values, yielding a value of collusion $V^C(\delta, l)$ for each discount factor δ and degree of loss aversion l in the grid. The critical discount factor is the smallest discount factor for which a cartel forms. Thus, the critical discount factor, for a given value of l, is the smallest δ value such that $V^C(\delta, l) > V^N(\delta, l)$. The preceding procedure determines an approximate critical discount factor for each value of l in the grid. To construct a continuous representation of the critical discount factor, LOESS smoothing is employed to estimate the value of the critical discount factor at intermediate (off-grid) l values. The smoothed curve depicting the critical discount factor as a function of l is presented in Figure 2.

Throughout the main text, parameter values in numerical simulations are selected in order to ensure the relevant effects can be clearly distinguished visually in the figures and, to the greatest extent possible, be representative of patterns observed more generally in alternative parameter settings. Parameter values are not intended to reflect a particular industry/cartel nor empirical estimates from prior literature. Determining realistic values for all parameters (including the rate of reference point adjustment (α), degree of loss aversion (l), and fixed cost of collusion (F)) as well as accurate functional forms for utility $u(\cdot)$ and the loss function $L(\cdot)$ is challenging and beyond the scope of this study. A variety of additional simulations, including variations in the parameter values employed in figures in the main text, are presented in the following subsections.

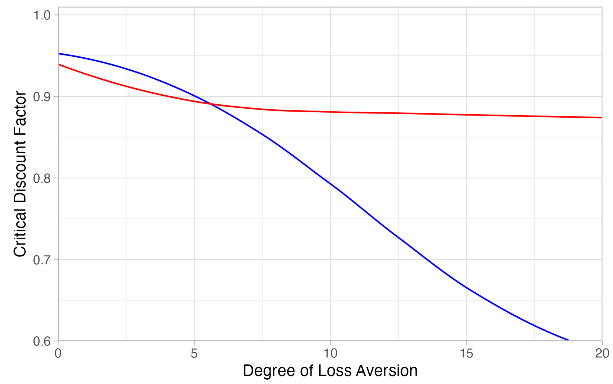
G.2 Cartel Formation

Figure 6 plots the critical discount factor both before and after the entry of a new firm for alternative α values. Note that when $\alpha = 1$ (i.e., reference points are constant), an increase in the degree of loss aversion does not impact the critical discount factor under regime b. This is the case as, when reference points are constant at the Nash equilibrium profit level, a reversion to Nash competition is not, in any period, perceived as a loss by managers. Thus, the degree of loss aversion does not impact manager utility. However, when $\alpha < 1$, reference points adjust upward in response to elevated utilities during collusion. After reference points adjust upwards, a return to Nash competition is perceived as a loss. Thus, an increase in the degree of loss aversion reduces the utility managers experience in the punishment phase, enhancing incentives to collude.

Panel A: $\alpha = 0$



Panel B: $\alpha = .5$



Panel C: $\alpha = 1$

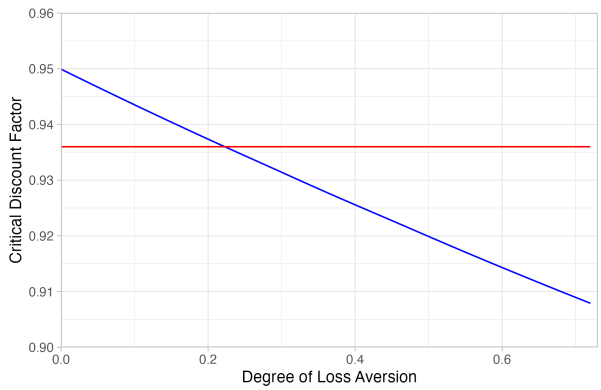
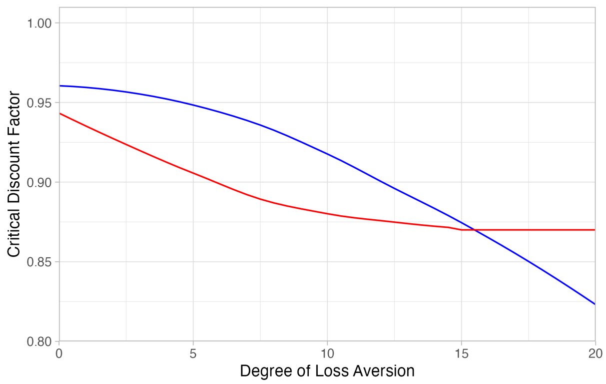


Figure 6: Critical Discount Factor by Degree of Loss Aversion Before (Red) and After (Blue) the Entry of a New Firm.

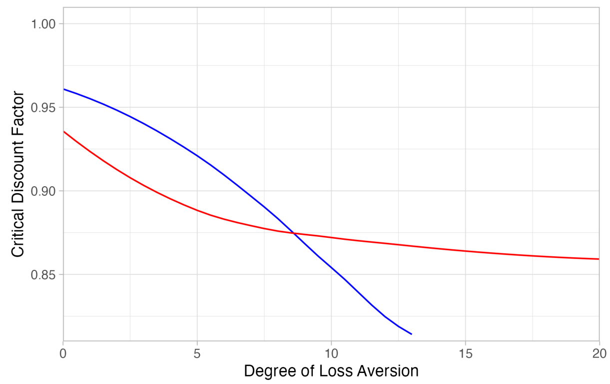
Notes: These figures depict the critical discount factor when $\alpha = 0$ (Panel A), $\alpha = .5$ (Panel B) and $\alpha = 1$ (Panel C) as a function of the degree of loss aversion. The critical discount factor is the smallest discount factor for which a cartel forms. Parameters: $a = 100$, $b = 2$, $e = 1$, $c = 0$, and $F = 125$. The blue curve denotes the critical discount factor after entry (i.e., $N = 6$) and the red curve denotes the critical discount factor absent entry (i.e., $N = 5$).

Figure 7 depicts the critical discount factor before (red) and after (blue) a 50% increase in marginal cost for a parameter configuration involving a fixed cost of collusion. When managers are loss neutral, an increase in marginal cost increases the critical discount factor and reduces incentives to collude. This is the case as an increase in marginal cost reduces the variable profits each firm earns during collusion which makes it more difficult to cover fixed costs of collusion. However, when managers are sufficiently loss averse, an increase in marginal cost can reduce the critical discount factor and increase the range of discount factors for which a cartel forms.

Panel A: $\alpha = 0$



Panel B: $\alpha = .5$



Panel C: $\alpha = 1$

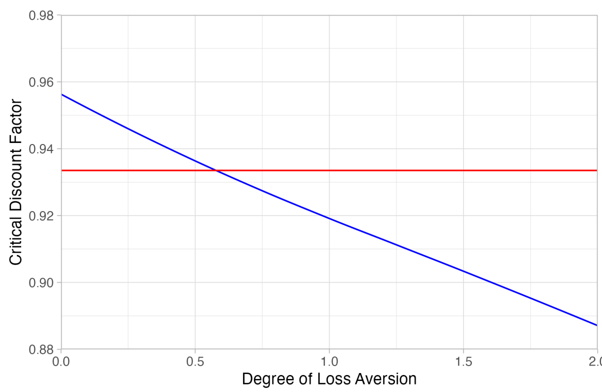


Figure 7: Critical Discount Factor by Degree of Loss Aversion Before (Red) and After (Blue) a 50% Increase in Marginal Cost.

Notes: These figures depict the critical discount factor when $\alpha = 0$ (Panel A), $\alpha = .5$ (Panel B) and $\alpha = 1$ (Panel C) as a function of the degree of loss aversion. The critical discount factor is the smallest discount factor for which a cartel forms. Parameters: $a = 100$, $b = 2$, $e = 1$, $N = 5$ and $F = 100$. The blue curve denotes the critical discount factor after a 50% increase in marginal cost (i.e., $c = 15$) and the red curve denotes the critical discount factor absent an increase in marginal cost (i.e., $c = 10$).

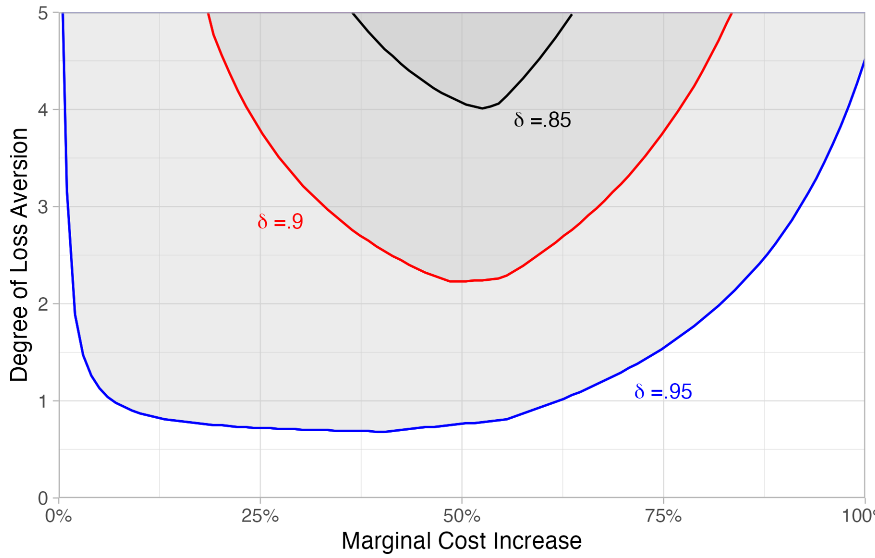


Figure 8: Threshold Degree of Loss Aversion $\bar{l}$ by Size of Marginal Cost Increase.

Notes: This figure depicts the threshold degree of loss aversion $\bar{l}$ by the size of a marginal cost increase for various discount factors. The grey shaded region depicts marginal cost increases and degrees of loss aversion for which a cartel forms. Parameters: $a = 100$, $b = 2$, $e = 1$, $N = 5$, $\alpha = 1$ and $F = 105$. Absent a deterioration in market conditions, all firms have a marginal cost of $c = 10$. Thus, a 50% increase in marginal cost corresponds to a marginal cost of 15.

Figure 8 depicts the threshold degree of loss aversion $\bar{l}$ (as in Proposition 1) when the deterioration in market conditions is an increase in marginal cost, for various discount factors. When the degree of loss aversion exceeds $\bar{l}$, a cartel forms after the marginal cost increase. Thus, the shaded grey regions in Figure 8 denote combinations of the degree of loss aversion (l) and the size of the marginal cost increase that result in the formation of a cartel. Note that, for the discount factors depicted in Figure 8, a cartel does not form prior to the marginal cost increase (i.e., regime *b*)) for any degree of loss aversion (thus, Condition 1 holds for all l).

Figure 8 indicates that a moderate increase in marginal cost causes a cartel to form for the widest range of l values (i.e., $\bar{l}$ is smallest for moderate increases in marginal cost). This finding reflects two considerations. First, small increases in marginal cost do not significantly reduce Nash equilibrium profit and, therefore, are not perceived as significant losses, which weakens the effects outlined in the preceding discussion. Second, pronounced increases in marginal cost substantially reduce each firm's variable profits and limit the firm's ability to cover fixed costs of collusion, weakening incentives to collude.[38]

[38] If the increase in marginal cost is exceptionally large, collusive profits may be less than Nash equilibrium profits prior to the deterioration in market conditions (i.e., a violation of Assumption 4(ii)). In this case, collusion would be perceived as a loss and the effects outlined in Section 5 do not occur.

G.3 Collusive Payoff

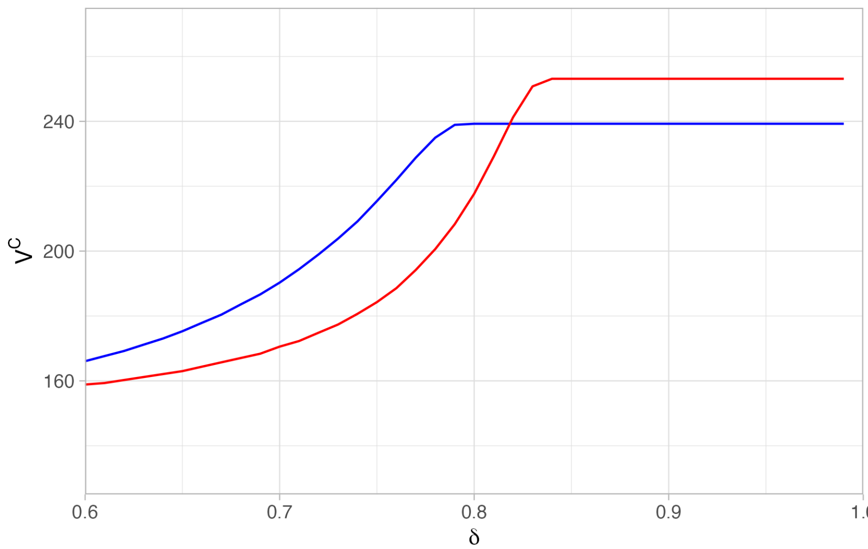


Figure 9: Collusive Payoff by Discount Factor Before (Red) and After (Blue) a 25% Increase in Marginal Cost.

Notes: This figure depicts $1-\delta$ times the collusive payoff as a function of the discount factor. Parameters: $a = 100$, $b = 2$, $e = 1$, $N = 7$, $F = 0$, $\alpha = 1$ and $l = 7.5$. Absent the deterioration in market conditions, all firms have a marginal cost of $c = 10$. The blue curve depicts $(1-\delta)\,V_a^C$ after a 25% increase in marginal cost (i.e., $c = 12.5$) and the red curve depicts $(1-\delta)\,V_b^C$ absent an increase in marginal cost (i.e., $c = 10$).

Figure 3 depicts the (normalized) collusive payoff after an increase in marginal cost (in blue) and absent an increase in marginal cost (in red) as a function of the discount factor when managers are loss averse. An increase in marginal cost enhances the collusive payoff for moderate discount factors. If the discount factor is instead sufficiently close to 1, then the cartel can, prior to the increase in marginal cost, set monopoly prices in all periods (i.e., $V_b^C(l) = V_b^M$). If this is the case, then a deterioration in market state will never enhance the collusive payoff because no price path can generate a collusive payoff exceeding the monopoly payoff. Formally, Condition 2(ii) is violated when the discount factor is relatively large.

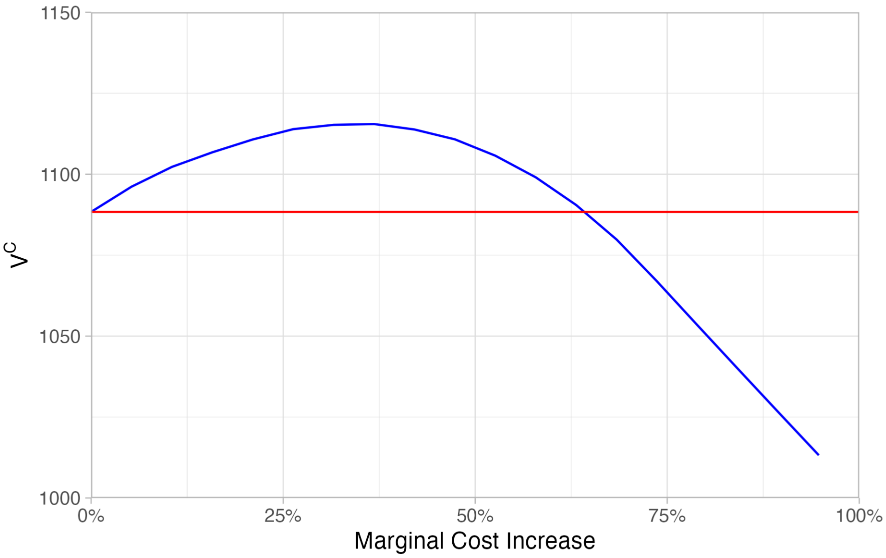


Figure 10: Collusive Payoff Before (Red) and After (Blue) Deterioration by Size of Marginal Cost Increase.

Notes: This figure depicts the collusive payoff as a function of the size of the marginal cost increase. Parameters: $a = 100$, $b = 2$, $e = 1$, $N = 7$, $\alpha = 1$, $F = 0$, $l = 2$ and $\delta = .8$. Prior to the deterioration in market conditions, all firms have a marginal cost of $c = 10$. Thus, a 50% increase in marginal cost results in a marginal cost of $c = 15$. The blue curve denotes V_a^C and the red curve denotes V_b^C.

Figure 10 depicts the collusive payoff after an increase in marginal cost (in blue) and absent an increase in marginal cost (in red), as a function of the size of the marginal cost increase when managers are loss averse. Moderate increases in marginal cost increase the collusive payoff (which is consistent with Proposition 2). However, drastic increases in marginal cost reduce the collusive payoff. To see this, note that substantial increases in marginal cost reduce the maximal/monopoly payoff (i.e., V_a^M) below the collusive payoff absent the marginal cost increase (i.e., V_b^C). If this is the case, managers cannot, for any degree of loss aversion, obtain a collusive payoff after the cost increase which exceeds the collusive payoff absent the increase. Formally, Condition 2(ii) is violated when the deterioration in market conditions is sufficiently severe.

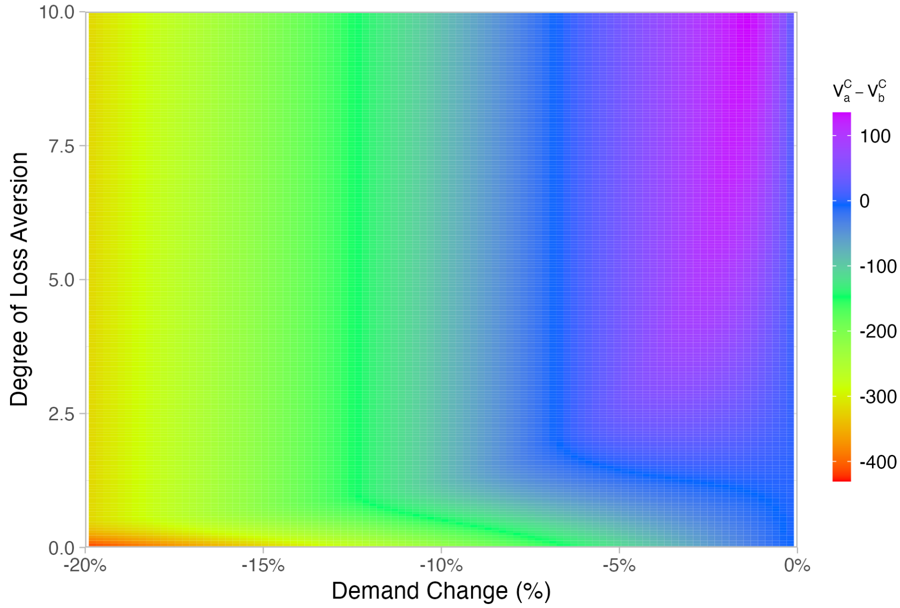


Figure 11: Difference in Collusive Payoff by Size of Demand Decrease and Degree of Loss Aversion.

Notes: This figure depicts $V_a^C - V_b^C$ as a function of the size of the decrease in the demand parameter a and the degree of loss aversion. Parameters: $b = 2$, $e = 1$, $N = 7$, $\alpha = 1$, $F = 0$ and $\delta = .8$. Prior to the deterioration in market conditions, $a = 100$. Thus, a 10% decrease in a implies $a = 90$.

Figure 12 depicts the difference between the collusive payoff after a decrease in the demand parameter a and the collusive payoff absent a change in demand (i.e., $\left(V_a^C - V_b^C\right)$) as a function of the size of the demand reduction and the degree of loss aversion l. A reduction in demand enhances the collusive payoff when 1) the reduction in demand is moderate and 2) managers are loss averse. Recall that extreme reductions in demand violate Condition 2(ii).

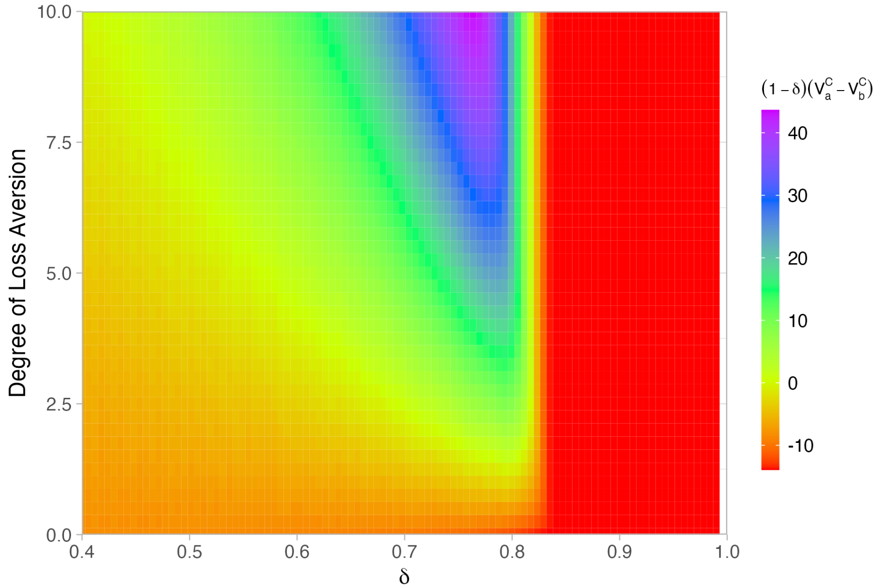


Figure 12: Change in Collusive Payoff from a 25% Increase in Marginal Cost by Discount Factor and Degree of Loss Aversion.

Notes: This figure depicts $(1-\delta)\left(V_a^C - V_b^C\right)$ as a function of the discount factor and the degree of loss aversion. Parameters: $a = 100$,$b = 2$, $e = 1$, $N = 7$, $F = 0$ and $\alpha = 1$. The deterioration in market conditions is an increase in marginal cost of 25%. Prior to the deterioration in market conditions, all firms have a marginal cost of $c = 10$. Thus, a 25% increase in marginal cost results in a marginal cost of $c = 12.5$.

Figure 12 depicts the normalized difference between the collusive payoff after a 25% increase in marginal cost and the collusive payoff absent a change in marginal cost (i.e., $(1-\delta)\left(V_a^C - V_b^C\right)$) as a function of the discount rate δ and the degree of loss aversion l. Consistent with Figure 3, an increase in marginal cost reduces the collusive payoff for relatively high discount factors. However, an increase in marginal cost can enhance the collusive payoff for moderate discount factors when managers are sufficiently loss averse.

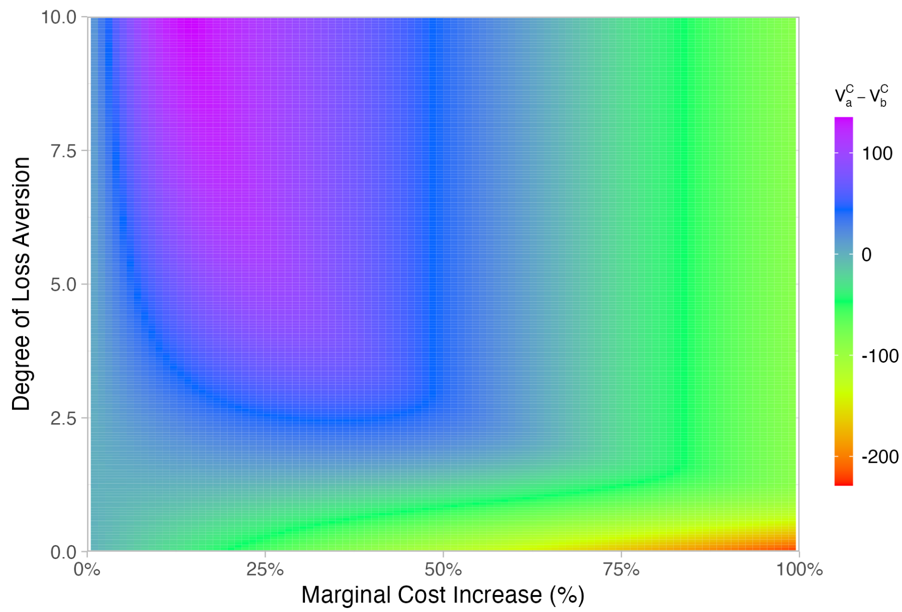


Figure 13: Difference in Collusive Payoff by Size of Marginal Cost Increase and Degree of Loss Aversion.

Notes: This figure depicts $V_a^C - V_b^C$ as a function of the size of the marginal cost increase and the degree of loss aversion. Parameters: $a = 100$, $b = 2$, $e = 1$, $N = 7$, $\alpha = 1$, $F = 0$ and $\delta = .8$. Prior to the deterioration in market conditions, all firms have a marginal cost of $c = 10$. Thus, a 50% increase in marginal cost results in a marginal cost of $c = 15$.

Figure 13 depicts the difference between the collusive payoff after an increase in marginal cost and absent a change in marginal cost (i.e., $\left(V_a^C - V_b^C\right)$) as a function of the size of the marginal cost increase and the degree of loss aversion l. Consistent with Figure 10, moderate increases in marginal cost increase the collusive payoff when managers are loss averse. Alternatively, large increases in marginal cost reduce the collusive payoff, regardless of the degree of loss aversion. The simulations in this subsection illustrate that a moderate deterioration in market state (e.g., a moderate increase in marginal cost) can enhance the collusive payoff when the discount factor is relatively low and managers are loss averse.

G.4 Gain From Collusion

Figure 14 depicts the gain from collusion after an increase in marginal cost by the size of the marginal cost increase and the degree of loss aversion. The gain from collusion is increasing in the degree of loss aversion and, when managers are sufficiently loss averse, increasing in the size of the marginal cost increase. Recall that increases in loss aversion impact the gain from collusion in two distinct ways. First, increases in loss aversion stabilize collusion and enhance the collusion payoff. Second, increases in loss aversion reduce V_a^N, the payoff from Nash competition, which also increases the gain from collusion in this simulation.

Figure 15 depicts the normalized gain from collusion after a 25% increase in marginal cost for various

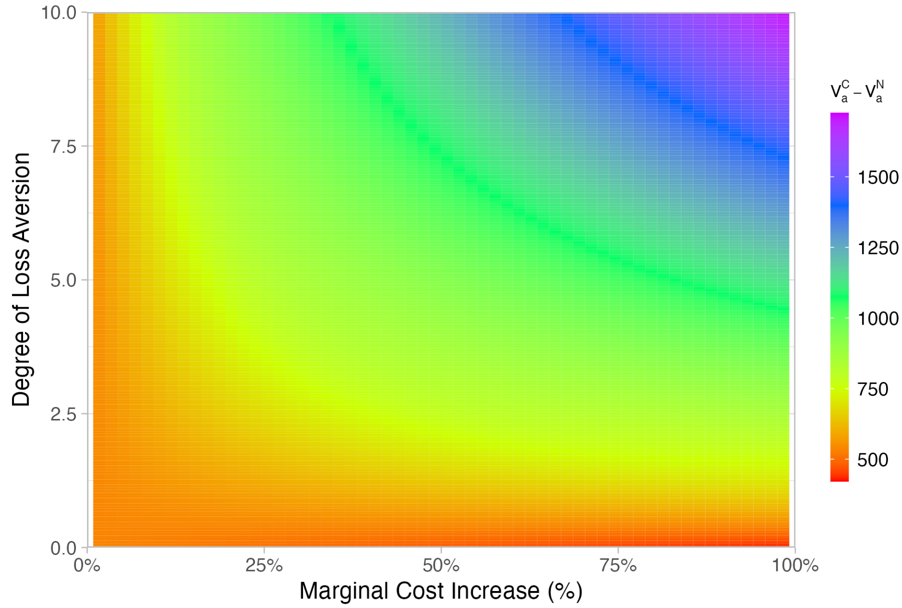


Figure 14: Gain from Collusion by Size of Marginal Cost Increase and Degree of Loss Aversion.

Notes: This figure depicts $V_a^C - V_a^N$ as a function of the increase in marginal cost and the degree of loss aversion. Parameters: $a = 100$, $b = 2$, $e = 1$, $N = 7$, $\alpha = 1$, $F = 0$ and $\delta = .8$. Prior to the deterioration in market conditions, all firms have a marginal cost of $c = 10$. Thus, a 25% increase in marginal cost results in a marginal cost of $c = 12.5$.

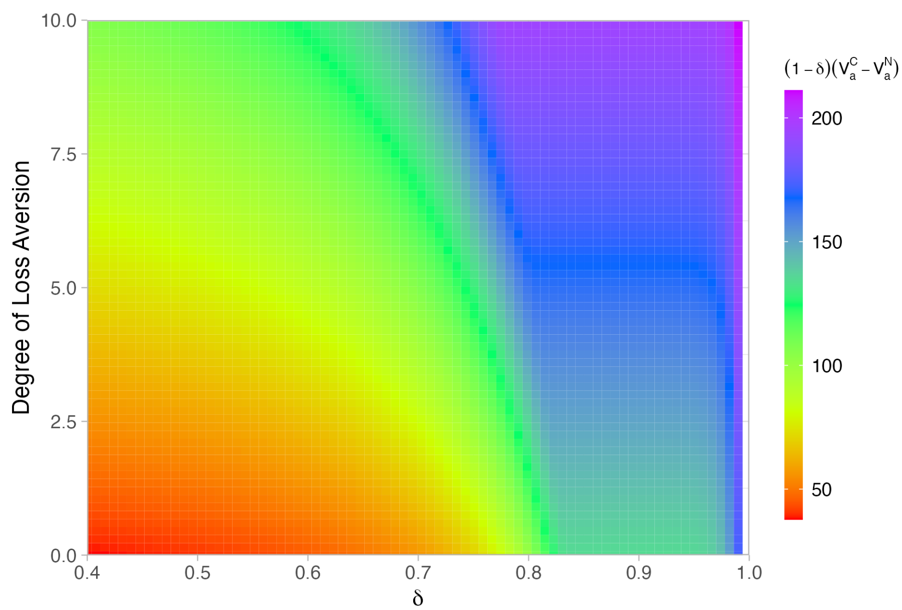


Figure 15: Gain from Collusion by Discount Factor and Degree of Loss Aversion After a 25% Increase in Marginal Cost.

Notes: This figure depicts $(1-\delta)\left(V_a^C - V_a^N\right)$ as a function of the discount factor and the degree of loss aversion. Parameters: $a = 100$, $b = 2$, $e = 1$, $c = 10$, $N = 7$, $F = 0$ and $\alpha = 1$. The deterioration in market conditions is a 25% increase in marginal cost. Prior to the deterioration in market conditions, all firms have a marginal cost of $c = 10$. Thus, a 25% increase in marginal cost results in a marginal cost of $c = 12.5$.

discount factors and degrees of loss aversion. Increases in the degree of loss aversion enhance the gain from collusion for all discount factors. Recall that a cartel forms for any discount factor as there are no fixed costs of collusion in this simulation.

G.5 Pricing Results

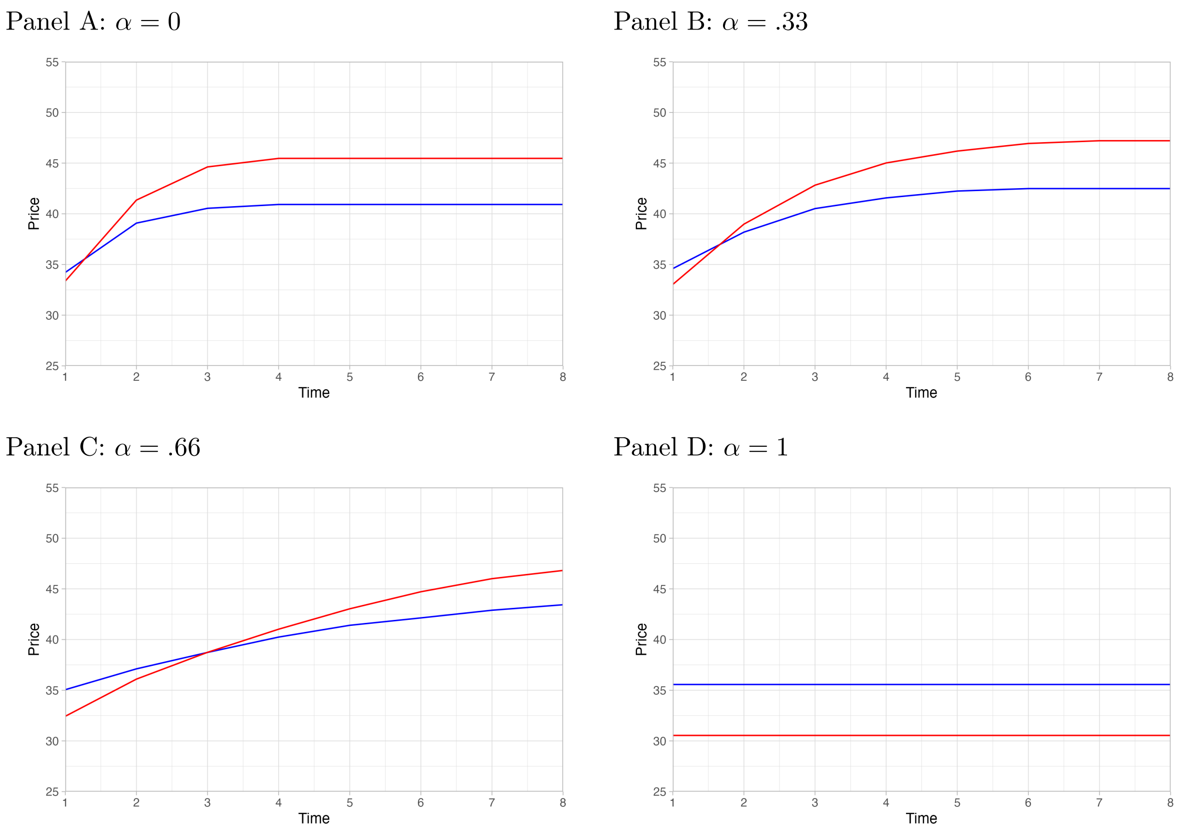


Figure 16: Optimal Price Paths Before (Red) and After (Blue) a 10% Reduction in Demand by α.

Notes: This figure depicts optimal price paths before (red) and after (blue) a 10% reduction in the demand parameter a for a variety of α values. Parameters: $b = 2$, $e = 1$, $c = 0$, $N = 4$, $\delta = .25$, $F = 0$ and $l = 10$. Prior to the deterioration in market conditions, $a = 100$. The blue curve depicts the optimal price path after the deterioration in market conditions (i.e., $a = 90$) and the red curve depicts the optimal price path prior to the deterioration in market conditions (i.e., $a = 100$).

Figure 16 presents the optimal price path before (red) and after (blue) a reduction in demand for a range of α values. In these simulations, $m(r, u) = \alpha r + (1 - \alpha) u$ where $\alpha \in [0, 1]$ determines the speed of reference point adjustment. $\alpha = 1$ corresponds to constant reference points that do not adjust in response to experienced utility levels (i.e., $m(r, u) = r$ as in Proposition 4). Figure 16 illustrates that the length of time that collusive prices under regime a exceed collusive prices under regime b increases with α. Consistent with Proposition 4, the reduction in demand enhances collusive prices in all periods when reference points are constant ($\alpha = 1$).

Panel A: $l = 0$

Panel B: $l = 5$

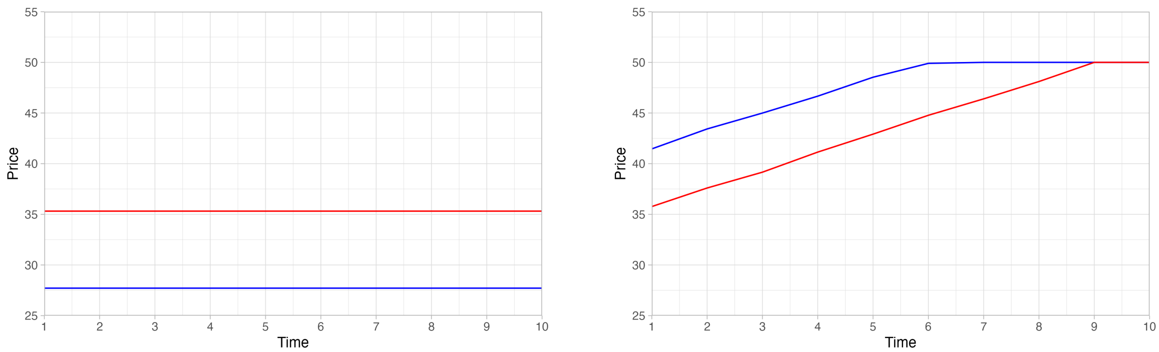


Figure 17: Optimal Price Paths Before (Red) and After (Blue) the Entry of a New Firm by Degree of Loss Aversion.

Notes: This figure depicts optimal price paths before and after the entry of a new firm for various degrees of loss aversion. Parameters: $a = 100$, $b = 2$, $e = 1$, $c = 0$, $\delta = .5$ and $\alpha = .9$. Prior to the entry of a new firm, $N = 4$. The blue curve depicts the optimal price path after the deterioration in market conditions (i.e., $N = 5$) and the red curve depicts the optimal price path prior to the deterioration in market conditions (i.e., $N = 4$). The Nash equilibrium price before (after) the change is 20 (16.67). The monopoly price before and after the change is 50.

Figure 17 presents the optimal price paths before and after the entry of a new firm. When managers are loss neutral, prices under regime b always exceed prices under regime a. However, when managers are loss averse, prices under regime a exceed prices under regime b in early periods of collusion. In later periods of collusion, prices under both regimes equal the monopoly price (50).