NBER WORKING PAPER SERIES

COLLUSION AT THE EXTENSIVE MARGIN

Martin C. Byford
Joshua S. Gans

Working Paper 20163
http://www.nber.org/papers/w20163

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2014

We would like to thank John Asker, Kathryn Spier, two anonymous referees and seminar participants at Harvard University, Harvard Law School, the US Department of Justice, New York University, Northeastern University, Northwestern University, University of Texas (Austin), University of Melbourne, Australian National University and University of Colorado at Boulder for comments on earlier drafts of this paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w20163.ack

Collusion at the Extensive Margin
Martin C. Byford and Joshua S. Gans
NBER Working Paper No. 20163
May 2014
JEL No. C73,L41

ABSTRACT

We augment the multi-market collusion model of Bernheim and Whinston (1990) by allowing for firm entry into, and exit from, individual markets. We show that this gives rise to a new mechanism by which a cartel can sustain a collusive agreement: Collusion at the extensive margin whereby firms collude by avoiding entry into each other's markets or territories. We characterise parameter values that sustain this type of collusion and identify the assumptions where this collusion is more likely to hold than its intensive margin counterpart. Specifically, it is demonstrated that Where duopoly competition is fierce collusion at the extensive margin is always sustainable. The model predicts new forms of market sharing such as oligopolistic competition with a collusive fringe, and predatory entry. We also provide a theoretic foundation for the use of a proportional response enforcement mechanism.

Martin C. Byford
Department of Economics
RMIT University GPO Box 2476
Melbourne, Victoria 3001 Australia
martin.byford@rmit.edu.au

Joshua S. Gans
Rotman School of Management
University of Toronto
105 St. George Street
Toronto ON M5S 3E6
and NBER
joshua.gans@gmail.com

1 Introduction

It has long been understood that the existence of multiple markets creates the potential for *market sharing*; a collusive agreement in which each member of a cartel is assigned monopoly rights over a territory (see for example Edwards, 1955; Stigler, 1964). A common feature of market sharing models is that, when a firm deviates, it captures a share of its rival's market before its rival has the opportunity to respond (see for example Bernheim and Whinston, 1990; Gross and Holahan, 2003; Belleflamme and Bloch, 2008; Bond and Syropoulos, 2008). This seems a reasonable assumption for industries in which products are manufactured in a firm's home market, before being transported for sale in foreign markets. If, as is the case for many commodities, transportation is quicker than production, a firm's arrival in a market could catch incumbents off guard.

There are, however, many industries in which a firm cannot contest a market without first establishing a presence in that market. If the process of entry is observable and takes a sufficiently long time, incumbents will have the opportunity to adjust their behaviour within the market in anticipation of the entrant's arrival. This fundamentally alters both the incentives for a firm to deviate from a cartel agreement, and the mechanisms by which the cartel can punish deviations.

In this paper, we augment the multi-market collusion model of Bernheim and Whinston (1990) (henceforth BW) by incorporating an explicit mechanism for firm entry into, and exit from, individual markets. In each period, firms decide which markets to contest before selecting their behaviour within each market. While a firm can surprise its rivals by its decision to enter a market, this action is observable. From the perspective of collusion, participation in a market, rather than simply actions within markets, can form the basis of histories that enforce collusive outcomes in a repeated non-cooperative games. Cartels can assign different markets to different firms with a deviation being entry by a firm into a market not assigned to it. That such entry could trigger counter-entry by rivals is what disciplines cartel behaviour. We term such outcomes *collusion at the extensive margin* to distinguish it from *collusion at the intensive margin*, based on firms' behaviour within markets (in terms of price setting and quantity restrictions) that has been the focus of most of the formal literature to date.

Our model is relevant to a number of industries that are subject to ongoing regulatory scrutiny. As an example, consider the antitrust case against Rural Press and Waikerie that was adjudicated by the High Court of Australia. Rural Press marketed a newspaper, *The Murray River Standard*, in the towns of Murray Bridge and Mannum (among others) while Waikerie operated another newspaper, *The River News* in Waikerie; all along the Murray River in South Australia. When Waikerie started selling and marketing (to advertisers), *The River News* in Mannum, Rural Press responded with a (draft) letter:

> The attached copies of pages from The River News were sent to me last week. The Mannum advertising was again evident, which suggests your Waikerie operator, John Pick, is still not focussing on the traditional area of operations.
>
> I wanted to formally record my desire to reach an understanding with your family in terms of where each of us focuses our publishing efforts.
>
> If you continue to attack in Mannum, a prime readership area of the Murray Valley Standard, it may be we will have to look at expanding our operations into areas that we have not traditionally services [sic].
>
> I thought I would write to you so there could be no misunderstanding our position. I will not bother you again on this subject.[1]

Waikerie promptly exited Mannum. The Australian courts found that this was an anti-competitive agreement and fined both parties (see Gans, Sood and Williams, 2004). Note that this did not involve attempted collusion within the Mannum area but instead a division of geographic markets along the Murray

[1] *Rural Press Ltd v Australian Competition and Consumer Commission*; *Australian Competition and Consumer Commission v Rural Press* (2003) 203 ALR 217; 78 ALJR 274; [2003] ATPR 41-965; [2003] HCA 75 (Rural Press decision).

River. Note also that the antitrust violation resulted from the enforcement of a deviation from an implied 'agreement' and, indeed, the newspapers exist in their separate markets today.

Interestingly, Stigler (1964) briefly considered this type of collusion but dismissed it, writing:

> ... the conditions appropriate to the assignment of customers will exist in certain industries, and in particular the geographical division of the market has often been employed. Since an allocation of buyers is an obvious and easily detectable violation of the Sherman Act, we may again infer that an efficient method of enforcing a price agreement is excluded by the antitrust laws. (p.47)

However, today, it is more likely that, absent evidence of an explicit agreement or a 'smoking gun' letter, such as existed in the Australian case, collusion at the extensive margin would be difficult to prosecute. Specifically, the successful prosecution in the Australian case is likely an exception rather than the rule with the investigation being triggered by off the equilibrium path behaviour rather than the collusive outcome itself. Indeed, in 2007, in *Bell Atlantic v. Twombly*[2] the US Supreme Court examined the complaint that Baby Bell telephone companies violated Section 1 of the Sherman Act by refraining from entering each other's geographic markets. The Court recognized that "sparse competition among large firms dominating separate geographical segments of the market could very well signify illegal agreement." However, they did not consider that an unwillingness on the part of Baby Bells to break with past behaviour and compete head to head was necessarily a conspiracy. The Court concluded that the implicit refraining of competition was a natural business practice; placing an evidentiary burden on off the equilibrium path behaviour.[3]

[2] *Bell Atlc v. Twombly* 550 U.S. 544 (2007).

[3] Our model also shows how a system of mutual forbearance can be sustained when each firm operates in a different product market. For instance, accounts of Apple and Google's recent falling out have indicated that this arose when Google entered into the mobile phone industry (with hardware as well as software) challenging Apple's iPhone (Stone and Helft, 2010). It was reported that Apple's response (possibly restricting Google applications on the iPhone as well as acquiring a mobile advertising start-up) was the result of Google's violation of a 'gentleman's agreement.'

The paper proceeds as follows. The model is detailed in section 2. There we add an explicit participation stage (where firms choose which markets to enter and/or exit) to the stage game in BW. We also state conditions under which a maximal competitive outcome can arise in equilibrium. Collusion utilising a grim-trigger strategy is considered in section 3. Significantly, we show that mutual avoidance outperforms multi-market contact if duopoly profits are sufficiently small. That is, a more intense baseline level of market competition makes collusion at the extensive margin stable at discount factors where collusion at the intensive margin cannot be sustained. In addition, we discuss the role of entry costs and also asymmetries between markets in terms of their value to the cartel. Uncertainty is introduced into the model in section 4. We show that reducing the length of the punishment phase increases expected profits at the expense of cartel stability. Moreover, expected profits may be further improved if cartel punishments are target at the deviating firm and scale with the size of the initial deviation. Finally, in section 5 we describe two novel forms of market sharing agreement predicted by the model, *oligopoly with a collusive fringe* and *predatory entry.* These are of interest because they involve observable behaviour that is distinct from the normal indicia of anti-competitive behaviour examined by anti-trust authorities. A final section concludes.

2 The multi-market model

The seminal paper in the multi-market collusion is BW (1990). In their model, firms tacitly collude over the levels of 'within market' actions such as price and quantity. Here we preserve that possibility but add another dimension for collusion based on 'market participation.' Specifically, rather than taking the choice of market presence as a costless one for firms, we assume that entry involves costs and takes some time. Consequently, while it may be that those deviating from a collusive agreement on 'within market' actions can profit prior to a reaction by others, when it comes to collusion based on participation,

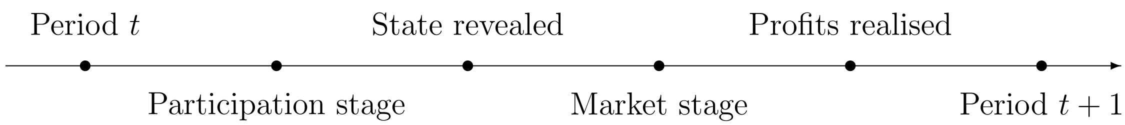

Figure 1: Timing

others can react fully if a deviation is observed.[4]

While our model allows for collusion over market actions *and* market participation, it will be useful to distinguish between two distinct types of collusive agreement. When multiple firms coordinate their behaviour within a single market, we say that they are colluding at the *intensive margin* of that market. If instead, the firms coordinate their participation across a set of markets, each acting as a monopolist in a subset of markets, we say that they are colluding at the *extensive margin*.

2.1 Preliminaries

Consider an infinite-horizon game in which a set I of identical firms interact repeatedly over a set N of discrete markets. It is assumed that $\|I\| \geq 2$ while $\|N\| \geq \|I\|$.[5] All firms discount the future by the common discount factor $\delta \in (0, 1)$.

The timing of the game is set out in figure 1. Any given period, t, begins with the *participation stage* in which firms decide which markets they will contest. Formally, firm i's participation stage action is a subset $a_i^t \subseteq N$. The inclusion of a market $n \in a_i^t$ indicates that firm i will contest market n in period t, while $n \notin a_i^t$ indicates that i will be absent from n.

Entry and/or exit occurs when a firm's participation stage action differs across consecutive periods. Specifically, firm i is said to *enter* (resp. *exit*)

[4] BW (1990) (see also Belleflamme and Bloch, 2008) have a variant of their model where the costs of producing in a given market involve some fixed costs for the firm. However, they assume that if a firm is merely present in a market but does not produce, its costs are zero. By contrast, we assume that being present in a market requires an observable step and investment even though once a firm is present in a market, collusion over the precise level of output is possible.

[5] The notation $\|N\|$ refers to the cardinality of the set N.

market n in period t if $n \in a_i^t$ and $n \notin a_i^{t-1}$ (resp. $n \in a_i^{t-1}$ and $n \notin a_i^t$). Entry by firm i into market n costs an amount $c_{i,n} \geq 0$. The entry cost is only incurred in the period in which entry occurs. As in BW, the cost of maintaining a presence in a market following entry is assumed to be accounted for in the market's profit function outlined below. If a firm exits and subsequently reenters a market the entry cost must be paid again. Importantly, this entry cost means that firms must commit to be present in a market and can also commit not to be present.

Following the participation stage, the profile of firm participation $a^t = \{a_i^t\}_{i \in I}$ is revealed to the market. Thus, all firms know the number and identities of their rivals in each market, when they compete in the subsequent market stage.

In the *market stage* firms choose actions for each of the markets they are contesting. Formally, firm i selects an action $x_{i,n}^t$ for each market $n \in a_i^t$. The nature of the action $x_{i,n}^t$ depends on the competitive environment of the market in question. For example, $x_{i,n}^t$ could represent a choice of price, quantity or quality.

Profits are realised at the conclusion of the market stage. The profit that firm i earns in market n is given by the function $\pi_n(x_{i,n}^t, x_{-i,n}^t)$.[6] Because firms are identical, the profit function does not vary between firms (although it may vary between markets).

Firms seek to maximise the present value of their lifetime profits. For firm i, this is,

$$\Pi_i = \sum_{t=0}^{\infty} \delta^t \left(\sum_{n \in a_i^t} \pi_n(x_{i,n}^t, x_{-i,n}^t) - \sum_{n \in a_i^t \setminus a_i^{t-1}} c_{i,n} \right), \tag{1}$$

where $a_i^t \setminus a_i^{t-1}$ is the set of markets firm i enters in period t.

2.2 The profit function

For any given number of participating firms, the (oligopolistically) competitive outcome in a market is the Nash equilibrium of a one-shot *market stage* game;

[6] The term $x_{-i,n}^t$ represents the actions of all firms other than i in market n, and so is contingent on a^t.

the equilibrium outcome where each participating firm selects an action $x_{i,n}^t$ and receives a payoff $\pi_n(x_{i,n}^t, x_{-i,n}^t)$.

Let ϕ_n^t represent the number of firms contesting market n in period t. For all $\phi_n^t \in \{1, \ldots, \|I\|\}$ we assume that there exists an action $x_n^*(\phi_n^t)$, such that all firms choosing $x_n^*(\phi_n^t)$ is the unique Nash equilibrium of the one-shot *market stage* game in market n with ϕ_n^t participants. The corresponding Nash equilibrium profit of a participating firm is written $\pi_n^*(\phi_n^t)$.

The monopoly action and profit for market n are, thus, $x_n^*(1)$ and $\pi_n^*(1)$ respectively, while $x_n^*(I)$ and $\pi_n^*(I)$ represent Nash equilibrium actions and profits when all firms in I contest market n. The following assumption describes the relationship between the number of firms contesting a market and equilibrium profits.

Assumption 1 (Expansion incentive) *For all* $n \in N$ *and* $\phi_n^t \in \{1, \ldots, \|I\| - 1\}$,

$$\phi_n \pi_n^*(\phi_n^t) > (\phi_n^t + 1)\pi_n^*(\phi_n^t + 1). \tag{2}$$

Moreover,

$$\frac{1}{1-\delta}\pi_n^*(I) > c_{i,n}. \tag{3}$$

(2) states that an increase in the number of firms contesting a market reduces the equilibrium profit of each incumbent firm. Moreover, the reduction in each firm's equilibrium profit is sufficiently large that total market profits also fall. A lower bound on equilibrium profits is established in (3). It states that, even under the most competitive conditions, the discounted sum of lifetime Nash equilibrium profits is sufficient to recoup the cost of entry into the market.

2.3 Participation strategies and equilibria

In the infinite-horizon game, the history in period t is a record of the actions taken in the preceding periods $h^t = \{(a^0, x^0), \ldots, (a^{t-1}, x^{t-1})\}$, where the vector $x^\tau = \{x_{i,n}^\tau\}_{i \in I, n \in N}$ represents the actions of all firms, in all markets, in period τ. All prior actions are observable so there is no uncertainty regarding past behaviour. The history h_t can be decomposed into the participation history $h_P^t = \{a^0, \dots, a^{t-1}\}$ and the market stage history $h_M^t = \{x^0, \dots, x^{t-1}\}$.

BW (1990) focus exclusively on collusion at the intensive margin where the participation history is not relevant. In their model, a cartel cares only about the market stage history as it already knows that all firms are present in all markets.

By contrast, we want to focus on collusion at the extensive margin, or what we will call *participation equilibria*, where strategies are conditioned on the participation history and not on the market stage history.[7] Formally, we say that firm i is employing a *participation strategy* if a_i^t and $\{x_{i,n}^t\}_{n \in a_i^t}$ are independent of the market stage history h_M^t for all h^t and $t \in \{0, 1, \dots\}$. We call a sub-game perfect equilibrium (SPE) in which all firms employ participation strategies, a *participation equilibrium*.

It is important to note that while, by definition, all firms employ participation strategies in a participation equilibrium, participation equilibria must be robust against unilateral deviations by any firm, at any history, to any strategy, including non-participation strategies. Thus, the set of participation equilibria is a subset of the set of all SPEs, and the concept should be viewed as an equilibrium refinement rather than an alternative solution concept.

This fact notwithstanding, the following lemma establishes that a profile of participation strategies constitutes a participation equilibrium if, and only if, it is robust to deviations within the class of participation strategies.

Lemma 1 *Suppose that all firms $j \neq i$ play* participation strategies. *Firm i has a best-response that is likewise a participation strategy.*

Proof. Because a_j^t and $\{x_{j,n}^t\}_{n \in a_j^t}$ are independent of h_M^t for all h^t, $t \in \{0, 1, \dots\}$ and $j \neq i$, the set of firm i's best responses is likewise independent of h_M^t for all h^t and $t \in \{0, 1, \dots\}$. ■

Another characteristic of a participation equilibrium is that market stage actions and profits must be identical to the Nash equilibrium actions and profits of a one-shot market stage game. The following lemma established

[7] Thus, they can arise even if market stage history where unobserved.

this result. Moreover, it allows us to omit market stage actions when we characterise participation equilibria.

Lemma 2 *In a participation equilibrium* $x_{i,n}^t = x_n^*(\phi_n^t)$ *for all* $n \in a_i^t$, $i \in I$, h^t *and* $t \in \{0, 1, \dots\}$.

Proof. Consider the market stage of period t. Because each firm's strategy is independent of the market stage history, a firm's only concern is to maximise the its profits in each of the markets it is contesting. By assumption, this occurs when $x_{i,n}^t = x_n^*(\phi_n^t)$ for all $n \in a_i^t$ and $i \in I$. ■

2.4 The competitive baseline

To conclude this section we establish the existence of a *competitive equilibrium.* This equilibrium is a participation equilibrium and exists under very general conditions. The competitive equilibrium serves as the baseline against which the various collusive equilibria can be compared.

Proposition 1 (Competitive equilibrium) *There exists a participation equilibrium in which* $a_i^t = N$ *for all* $i \in I$, h^t *and* $t \in \{0, 1, \dots\}$.

Proof. Given that all firms seek to enter every market regardless of the history, the expansion incentive (assumption 1) makes expanding into every market a best response. ■

In a competitive equilibrium, firms employ strategies that are independent of the entire history of the game. Along the equilibrium path, all firms enter and remain in every market. In a given market n, all firms take the competitive action $x_n^*(I)$ and receive the competitive profit $\pi_n^*(I)$. The competitive equilibrium is very robust, existing for all $\delta \in (0, 1)$. Proposition 1 is significant for our analysis as it implies that wherever firms collude at the extensive margin, they do so in an environment in which there exists a competitive participation equilibrium (and hence a competitive SPE) which is at least as robust.

3 Collusion by mutual avoidance

Consider the case of a cartel that assigns to each of its members, monopoly control over a subset of markets. If $N_i \subset N$ represents the set of markets assigned to firm i, the cartel agreement can be represented by the partition $P = \left\{N_\varnothing, \{N_i\}_{i\in I}\right\}$. Included in the partition P is the (possibly empty) set $N_\varnothing \subset N$ representing unassigned markets that all firms are free to contest. The profile of participation that implements P is $a^P = \{a_i^P\}_{i\in I}$, where $a_i^P = N_i \cup N_\varnothing$ for all $i \in I$.

A cartel agreement of this type is simple and easy to monitor. The cartel has no interest in how a firm behaves in any given market, it only matters which markets each firm contests. A deviation from the agreement occurs if one firm enters a market assigned to a another firm. Absent a response, such entry is profitable by assumption 1.

In order to prevent deviations, a cartel must employ an enforcement mechanism. The only way to punish a firm is to enter some or all of its markets, competing away its monopoly profits. In this section, we consider the strongest available enforcement mechanism, a grim-trigger strategy. Temporary and targeted punishments are considered in section 4.

3.1 The grim-strategy equilibrium

In a perfect information setting, the most robust collusive equilibrium is the equilibrium with the strongest enforcement. The greatest punishment that can be imposed by participation strategies is for any transgression to cause the game to permanently revert to the competitive equilibrium set out in proposition 1.

The following proposition characterises the necessary and sufficient conditions for the existence of a grim-strategy equilibrium:

Proposition 2 (Grim-strategy equilibrium) *Consider the following participation strategy profile: For all $i \in I$, if there exists $\tau < t$ and $k \neq l$ such that $a_k^\tau \cap N_l \neq \varnothing$ then $a_i^t = N$; otherwise $a_i^t = a_i^P$. This strategy profile constitutes* *a participation equilibrium if and only if $\delta \geq \delta_{\mathrm{G}}$ where,*

$$\delta_{\mathrm{G}} = \max_{i \in I} \left[\frac{\sum_{n \in \bigcup_{j \neq i} N_j} \left(\pi_n^*(2) - c_{i,n} \right)}{\sum_{n \in N_i} \pi_n^*(1) - \sum_{n \in N \setminus N_\varnothing} \pi_n^*(I) + \sum_{n \in \bigcup_{j \neq i} N_j} \left(\pi_n^*(2) - c_{i,n} \right)} \right]. \quad (4)$$

Proof. From the proof of proposition 1, it follows that if a firm triggers the punishment phase by entering a rival's market in period t ($a_k^t \cap N_l \neq \varnothing$ for some $k \neq l$), reversion to the competitive equilibrium from period $t+1$ onward is an equilibrium of the sub-game.

From assumption 1 it follows that the worst case deviation is for a firm i to enter the set of all markets assigned to rival firms ($\bigcup_{j \neq i} N_j$). This deviation is not profitable if,

$$\frac{1}{1-\delta} \left(\sum_{n \in N_i} \pi_n^*(1) + \sum_{n \in N_\varnothing} \pi_n^*(I) \right) \geq \sum_{n \in N_i} \pi_n^*(1) + \sum_{n \in N_\varnothing} \pi_n^*(I) + \sum_{n \in \bigcup_{j \neq i} N_j} \left(\pi_n^*(2) - c_{i,n} \right) + \frac{\delta}{1-\delta} \sum_{n \in N} \pi_n^*(I).$$

Solving for δ yields (4). ■

Proposition 2 establishes the relationship between the critical discount factor δ_{G} and the structure of the cartel agreement. From (4) it follows that a necessary condition for the grim-strategy equilibrium is,

$$\sum_{n \in N_i} \pi_n^*(1) > \sum_{n \in N \setminus N_\varnothing} \pi_n^*(I), \quad (5)$$

for all $i \in I$. That is, the profits firm i receives as a result of retaining exclusive control of the markets in N_i must be higher than the profits it receives in a competitive equilibrium, from all markets assigned to cartel members ($N \setminus N_\varnothing$). Indeed, if (5) is satisfied for all $i \in I$ then there exists a $\delta \in (0, 1)$ satisfying (4).

More generally, (4) highlights the way in which cartel stability depends on the symmetry of the partition. Minimising δ_{G} requires an even distribution of monopoly profits between the cartel members. The contested component of the partition can play a useful role here. By assigning asymmetrically valuable markets to $N_\varnothing$, the cartel removes a powerful incentive for rival firms to deviate. Note that profits from the markets in $N_\varnothing$ do not appear anywhere in (4). All firms receive $\pi_n^*(I)$ from each market $n \in N_\varnothing$, regardless of whether the game is on the equilibrium path or has reverted to the competitive equilibrium.

Proposition 2 also illustrates the impact of entry costs on cartel stability. It follows from (4) that δ_G is decreasing in $\sum_{n \in \bigcup_{j \neq i} N_j} c_{i,n}$. This term represents the cost to firm i of initiating a deviation. As the cost of entry increases, the net return from entering a rival's market decreases, and cartel stability improves.

3.2 Identical markets and costless entry

It is instructive to consider cartel stability in a special case where (a) entry costs are zero (a worse case for stability) and (b) markets are identical (allowing straightforward allocations). We will here assume that $c_{i,n} = 0$ for all $i \in I$ and $n \in N$ and also that $\pi_m(\cdot) = \pi_n(\cdot)$ for all $m, n \in N$; dropping the market subscripts for simplicity. It is also useful to define $n_i = \|N_i\|$ and $n_\varnothing = \|N_\varnothing\|$.

With identical markets and zero entry costs (4) reduces to,

$$\delta_\mathrm{G} = \max_{i \in I} \left[\frac{\sum_{j \neq i} n_j \pi^*(2)}{n_i\big(\pi^*(1) - \pi^*(I)\big) + \sum_{j \neq i} n_j \big(\pi^*(2) - \pi^*(I)\big)} \right]. \tag{6}$$

Note that the fraction in (6) is increasing in the number of markets assigned to rival firms ($\sum_{j \neq i} n_j$) and decreasing in the number of markets over which firm i has monopoly control (n_i).

When markets are identical, the firm i that maximises (6), and, therefore, determines the level of the critical discount factor, is the firm with the smallest partition. It follows that cartel stability is maximised when all firms are assigned the same number of markets, with any remaining markets assigned to $N_\varnothing$. The minimum critical discount factor for a cartel of $\|I\|$ firms is thus,

$$\delta_\mathrm{G} = \frac{(\|I\| - 1)\pi^*(2)}{\pi^*(1) + (\|I\| - 1)\pi^*(2) - \|I\|\pi^*(I)} < 1, \tag{7}$$

where the inequality follows from (2). The RHS of (7) is unambiguously increasing in $\|I\|$. It follows that the stability of the cartel falls with each additional member.

3.3 Mutual avoidance versus multi-market contact

We now turn to compare mutual avoidance that arises in collusion purely at the extensive market with multi-market contact as examined by BW (1990). Here we discuss three broad substantive areas where the stability of each type of collusion depends on distinct structural characteristics of the environment: they are symmetries across markets, the level of entry costs and the time taken for entry.

First, with respect to markets themselves, BW (1990) provide an irrelevance result and show that, if markets are identical, the cartel will be no more stable than a cartel colluding in a single representative market. Intuitively, while multi-market contact does increase the magnitude of the punishments that may be imposed on a deviating firm, it proportionately increases the incentive to deviate. This implies that it is asymmetries between markets that facilitate stability in intensive margin collusion. By contrast, as demonstrated above, stability in extensive margin collusion is facilitated by symmetry so that markets of equivalent total value can be divided up amongst firms.

Second, for multi-market contact, entry costs, if any,[8] are incurred in the collusive equilibrium across all markets whereas under collusion at the extensive margin, entry costs are only incurred by firms in the markets they are allocated. Consequently, suppose that under collusion of both types, monopoly profit outcomes arise across markets. Then, net of any entry costs, the profits accruing to each firm would be the same but total entry costs would be correspondingly higher under collusion at the intensive margin. Thus, from the cartel's perspective, collusion at the extensive margin would Pareto dominate with the outcomes being equivalent as entry costs went to zero.

Third, as noted earlier, a key difference between BW (1990) and the model

[8]BW do not consider entry costs but they are potentially present in our model here. Specifically, in order for a firm to react quickly to a rival's deviation within a market, entry (which takes time and costs) would have to occur ex ante.

presented here is not so much that entry costs exist but that entry can be observed and takes time. Thus, while it is the case that deviating from intensive margin collusion allows a firm to capture profits that arise when their rivals continue to choose their collusive within market actions, when deviating from extensive margin collusion, rivals can observe a deviation and react to it in the market where it occurs. Of course, it takes time (one period) for the reaction to occur across all markets (just as it does for BW). In what follows, we consider in detail precisely when this difference between the models implies that one type of collusion is more stable than another, employing the assumptions of identical markets and zero entry costs so as to remove other reasons for a difference in stability outcomes between collusion types.

For extensive margin collusion, we assume an outcome where markets are allocated equally amongst firms and are enforced via a grim-trigger mechanism considered earlier. For the intensive margin outcome, consider the following cartel agreement in which $\|I\|$ identical firms collude at the intensive margin of $\|N\|$ identical markets. Under the cartel agreement, each firm establishes a permanent presence in every market. Firms coordinate market stage actions such that each firm receives an equal share of the monopoly profit $(\pi^*(1)/\|I\|)$ from each market.

The cartel employs the grim-trigger strategy as its enforcement mechanism. If a firm reneges on the cartel agreement, it receives a profit of π^{dev} from each market in which it deviates from the agreed behaviour. Any deviation is punished by permanent reversion to the competitive equilibrium.

Proposition 3 *A cartel that colludes at the extensive margin has a lower critical discount factor that a cartel colluding at the intensive margin if, and only if,*

$$\pi^{\text{dev}} - \frac{\pi^*(1)}{\|I\|} > \frac{\|I\|-1}{\|I\|}\pi^*(2). \tag{8}$$

Proof. The lowest possible critical discount factor for a cartel colluding at the extensive margin is given by (7).

Now consider a cartel colluding at the intensive margin. The worst case for this type of cartel occurs when a firm deviates in all markets simultaneously.

Such a deviation is not profitable if,

$$\|N\|\frac{1}{1-\delta}\cdot\frac{\pi^*(1)}{\|I\|} \geq \|N\|\pi^{\text{dev}} + \|N\|\frac{\delta}{1-\delta}\pi^*(I).$$

It follows that the critical discount factor is,

$$\delta_{\text{crit}} = \frac{\|I\|\pi^{\text{dev}} - \pi^*(1)}{\pi^*(1) + \big(\|I\|\pi^{\text{dev}} - \pi^*(1)\big) - \|I\|\pi^*(I)}. \tag{9}$$

Note that δ_{crit} is increasing in the difference $\|I\|\pi^{\text{dev}} - \pi^*(1)$. Comparing (7) and (9) it is clear that $\delta_{\text{G}} < \delta_{\text{crit}}$ if and only if (8) holds. ■

Proposition 3 has a straightforward interpretation. The LHS of (8) represents the net gain (per market) of deviating when a cartel coordinates behaviour within markets. The RHS of (8) represents the net gain (averaged across all markets) of deviating when a cartel coordinates participation across markets. The most stable cartel is the cartel with the lowest return to a deviation.

What is interesting here is that cartel stability is influenced by the intrinsic intensity of competition within markets. To see this, suppose that firms are selling close substitutes and competing by setting prices; i.e., competition within the market is intrinsically highly competitive. If a cartel is colluding at the intensive margin, a firm can capture an entire market by marginally undercutting its rivals. A firm thus earns $\pi_{\text{dev}} \approx \pi^*(1)$ from each market in which it initiates a deviation. In contrast, the duopoly profit $\pi^*(2)$ is close to zero. With these values (8) holds for all $\|I\| \geq 2$.[9]

This Bertrand case highlights the key difference between the two collusive mechanisms. When a firm deviates under intensive margin collusion, it surprises its rivals in the market stage, undercutting them and stealing the entire market. In contrast, in order to deviate from an extensive margin collusive agreement, a firm must first enter markets assigned to its rivals. Entry is observable and provides the remaining cartel members with sufficient time to adjust their own prices in anticipation of the entrant's arrival. Thus, the best

[9] The critical discount factor for intensive margin collusion in this example is $\delta_{\text{crit}} = (\|I\| - 1)/\|I\|$, while for extensive margin collusion δ_{G} is close to zero.

the deviating firm can hope for is duopoly profits. In this example, duopoly profits are negligible and hence, extensive margin collusion is more stable.

A counter example is the case of Cournot competition with constant marginal cost and a linear demand curve (generally, considered a less intense competitive environment than Bertrand). Suppose that the cartel agreement requires each firm to produce a fraction $1/\|I\|$ of the monopoly quantity in each market. In this example,

$$\pi_{\text{dev}} = \frac{(\|I\|+1)^2}{4\|I\|^2}\pi^*(1) \quad \text{and} \quad \pi^*(2) = \frac{4}{9}\pi^*(1).$$

Substituting these values into (8) violates the inequality for all $\|I\| \in \{2, 3, \dots\}$.

In the case of Cournot competition, intensive margin collusion is more stable than extensive margin collusion. When a firm deviates from an intensive margin collusive agreement it produces $(\|I\|+1)/2\|I\|$ of the monopoly quantity. The remaining firms in the market produce $(\|I\|-1)/\|I\|$ of the monopoly quantity, delivering the deviating firm a net gain of,

$$\pi_{\text{dev}} - \frac{\pi^*(1)}{\|I\|} = \frac{(\|I\|-1)^2}{4\|I\|^2}\pi^*(1) < \frac{\pi^*(1)}{4},$$

per market. In contrast, entry into a monopoly controlled market is attractive as Cournot duopolists earn substantial profits.

In summary, while it is the case that collusive outcomes are generally more stable the more competitive is the non-collusive environment, collusion at the intensive margin becomes less stable relative to collusion at the extensive margin, the more competitive the non-collusive environment becomes. Indeed, in the limit, collusion at the extensive margin is a possible equilibrium regardless of the discount factor.

4 Uncertainty

Since Green and Porter (1984) it has been understood that temporary punishments have advantages over the grim-trigger strategy where uncertainty triggers punishments along the equilibrium path. In this section we show that the scale and scope of punishments are also important.

4.1 The structure of uncertainty

For the purposes of this section we assume that the source of uncertainty is the possibility that a firm makes an error in the participation stage, contesting more markets than the firm intended. Such an error might occur if an overzealous manager, unaware of the existence of the cartel, overstepped their authority and initiated entry into a market without seeking permission from his or her superiors.

Suppose that firm i makes a mistake and contests markets assigned to a subset of firms $K \subseteq I \setminus \{i\}$. Let $d_k \in \{1, 2, \dots\}$ represent the number of firms belonging to a firm $k \in K$ that firm i contests. The complete profile of firm i's mistake is written $d_K = \{d_k\}_{k \in K}$. The probability that this mistake occurs is written $\sigma(i, K, d_K) > 0$. It will be useful to define the following probabilities,

$$\sigma(i, K) = \sum_{d_K} \sigma(i, K, d_K) \quad \text{and} \quad \sigma(i) = \sum_{K \subseteq I \setminus \{i\}} \sigma(i, K).$$

Here $\sigma(i, K)$ is the overall probability that firm i enters markets belonging to a subset of firms $K \subseteq I \setminus \{i\}$, while $\sigma(i)$ is the probability of firm i making a mistake of any sort.

In order to simplify the analysis we make two further assumptions concerning the structure of uncertainty. Neither of these assumptions is necessary, however each reduces the notation considerably. First, the probability of two or more firms making an error in the same period is arbitrarily close to zero. Second, an error can only occur if all firms implement the cartel agreement in the participation stage.

In contrast with Green and Porter (1984), the errors made by firms in our model are real. When a firm makes a mistake and enters a rival's market it profits from that mistake. The expected gains to firm i from mistakes it makes, less the expected losses due to mistakes by other cartel members, are,

$$\varepsilon_i = \sum_{\substack{K \subseteq I \setminus \{i\} \\ d_K \\ k \in K}} \sigma(i, K, d_K) d_k \pi^*(2) - \sum_{\substack{j \neq i \\ i \ni K \subseteq I \setminus \{j\} \\ d_K}} \sigma(j, K, d_K) d_i \big(\pi^*(1) - \pi^*(2)\big).$$

4.2 Un-targeted enforcement

The first temporary enforcement mechanism we consider is *un-targeted enforcement*. Under un-targeted enforcement, all firms respond to a deviation by reverting to the competitive equilibrium for τ periods. Following the conclusion of the punishment phase, firms withdraw from markets assigned to rival firms, restoring the collusive partition. The punishment phase of un-targeted enforcement is the analogue of a price war in intensive margin collusion.

Proposition 4 *Let $\sigma_{\mathrm{U}} = \sum_{i\in I}\sigma(i)$ represent the probability of an error occurring if all firms adhere to the collusive agreement. If $\delta \geq \delta_{\mathrm{U}}^{\infty}$ where,*

$$\delta_{\mathrm{U}}^{\infty} = \max_{i\in I}\left[\frac{\sum_{j\neq i} n_j\pi^*(2) - \varepsilon_i}{(1-\sigma_{\mathrm{U}})\big[n_i\big(\pi^*(1)-\pi^*(I)\big) + \sum_{j\neq i} n_j\big(\pi^*(2)-\pi^*(I)\big)\big]}\right],$$

then there exists a punishment length $\tau \in \{1, 2, \dots\}$ for which the cartel will be stable. If $\delta \geq \delta_{\mathrm{U}}^{1}$ where,

$$\delta_{\mathrm{U}}^{1} = \left(\frac{1}{\delta_{\mathrm{U}}^{\infty}} - 1\right)^{-1} > \delta_{\mathrm{U}}^{\infty},$$

then the cartel is stable with a single period punishment phase. The expected profit to firm i from participating in the cartel is,

$$\Pi_{i,\mathrm{U}} = \frac{1}{1-\delta}\left(\frac{n_i\pi^*(1) + \varepsilon_i + \frac{\delta-\delta^{\tau+1}}{1-\delta}\sigma_{\mathrm{U}}\sum_{j\in I} n_j\pi^*(I)}{1 + \frac{\delta-\delta^{\tau+1}}{1-\delta}\sigma_{\mathrm{U}}} + n_{\varnothing}\pi^*(I)\right).$$

Proof. From the proof to proposition 1 it follows that firms have no incentive to deviate during a punishment phase. Moreover, in each period of the punishment phase each firm earns the competitive profit $\|N\|\pi^*(I)$.

Outside of a punishment phase, the continuation value of the game to firm i from selecting the participation stage action a_i^P is,

$$V_i^+ = n_i\pi^*(1) + \varepsilon_i + n_{\varnothing}\pi^*(I) + \delta(1-\sigma_{\mathrm{U}})V_i^+ + \delta\sigma_{\mathrm{U}}\left(\frac{1-\delta^{\tau}}{1-\delta}\Big(\sum_{j\in I} n_j\pi^*(I) + n_{\varnothing}\pi^*(I)\Big) + \delta^{\tau}V_i^+\right).$$

Solving for V_i^+ yields the expected profit $\Pi_{i,\mathrm{U}}$. If, instead, firm i deviates, entering every market belonging to a rival firm, the continuation value of the game is,

$$V_i^- = n_i\pi^*(1) + \sum_{j\neq i} n_j\pi^*(2) + n_\varnothing\pi^*(I) \\ + \frac{\delta - \delta^{\tau+1}}{1-\delta}\Big(\sum_{j\in I} n_j\pi^*(I) + n_\varnothing\pi^*(I)\Big) + \delta^{\tau+1}V_i^+.$$

The cartel is stable if $V_i^+ - V_i^- \geq 0$ for all $i \in I$ implying,

$$\min_{i\in I}\left[\frac{1-\delta^{\tau+1}}{1-\delta}\left(\frac{n_i\pi^*(1) + \varepsilon_i + \frac{\delta-\delta^{\tau+1}}{1-\delta}\sigma_\mathrm{U}\sum_{j\in I} n_j\pi^*(I)}{1 + \frac{\delta-\delta^{\tau+1}}{1-\delta}\sigma_\mathrm{U}}\right) \right.\\ \left. - \left(n_i\pi^*(1) + \sum_{j\neq i} n_j\pi^*(2) + \frac{\delta-\delta^{\tau+1}}{1-\delta}\sum_{j\in I} n_j\pi^*(I)\right)\right] \geq 0.$$

Taking the limit as $\tau \to \infty$, and substituting for $\tau = 1$, yields δ_U^∞ and δ_U^1 respectively. ■

Proposition 4 characterises the critical discount factor and expected profits for a cartel employing un-targeted enforcement. In common with Green and Porter (1984), increasing the length of the punishment phase increases cartel stability at the expense of expected profits. Unlike Green and Porter (1984), it is possible for a cartel to be stable when punishments last for a single period.

Once again, the critical discount factors depend on the profits a firm earns from a deviation. If duopoly competition is fierce and duopoly profits are close to zero, both δ_U^∞ and δ_U^1 will likewise tend to zero. Under these circumstance single-period punishments are more than adequate to ensure cartel stability.

It is important to note that the efficacy of this enforcement mechanism, as well as the others discussed in this section, is predicated on the ability of a firm to reenter a market it has previously withdrawn from. If exiting a market damages a firm's reputation in a manner that makes reentry difficult, the credibility of temporary punishment is diminished and the cartel would have to rely on the grim-trigger strategy.[10]

[10]We are grateful to an anonymous referee for pointing this out.

4.3 Targeted enforcement

If there is slack between the discount factor δ, and the critical discount factor for single-period un-targeted enforcement δ^1_{U}, the cartel may be able to further improve expected profits by employing *targeted enforcement*.

Targeted enforcement treats deviations as bilateral disagreements. Suppose that firm i deviates from the cartel agreement in period t, entering markets belonging to a subset of firms $K \subseteq I \setminus \{i\}$. The responsibility for punishing firm i falls on its victims. In period $t+1$ all firms in K enter all markets in N_i competing away firm i's profits, while firm i contests every market in $\sum_{j \in K} N_j$. The punishment phase concludes after a single period with all firms withdrawing to their respective markets in period $t+2$.

Notice that, under targeted enforcement, the only firms involved in the punishment phase are the deviating firm and its victims. Moreover, the victims only retaliate against the aggressor, they do not target each other. The remaining cartel members play no part.

The probability of firm i either making a mistake, or being a victim of an error, is,

$$\sigma_i = \sigma(i) + \sum_{\substack{j \neq i \\ i \ni K \subseteq I \setminus \{j\}}} \sigma(j, K). \leq \sum_{j \in I} \sigma(j),$$

with strict inequality where there are three or more firms in the cartel. Expected profits and the critical discount factor are characterised in the following proposition.

Proposition 5 *The expected profit to firm i from participating in a cartel employing targeted enforcement is,*

$$\Pi_{i,\mathrm{T}} = \frac{1}{1-\delta} \left(\frac{n_i \pi^*(1) + \varepsilon_i + \delta \sigma_i z_{i,\mathrm{T}}}{1 + \delta \sigma_i} + n_\varnothing \pi^*(I) \right),$$

where,

$$z_{i,\mathrm{T}} = \frac{1}{\sigma_i} \sum_{K \in I \setminus \{i\}} \sigma(i,K)\Big(n_i \pi^*(\|K\|+1) + \sum_{k \in K} n_k \pi^*(2)\Big) \\ + \frac{1}{\sigma_i} \sum_{\substack{j \neq i \\ i \ni K \in I \setminus \{j\}}} \sigma(j,K)\big(n_i \pi^*(2) + n_j \pi^*(\|K\|+1)\big).$$

The critical discount factor δ_{T} *is continuous in the probability of each error. Moreover, in the limit as the probabilities of all error tend to zero, the critical discount factor becomes,*

$$\lim_{\sum_{i \in I} \sigma(i) \to 0} \delta_{\mathrm{T}} = \max_{i \in I} \left[\frac{\sum_{j \neq i} n_j \pi^*(2)}{n_i\big(\pi^*(1) - \pi^*(I)\big) - \sum_{j \neq i} n_j \pi^*(2)} \right]. \tag{10}$$

Proof. First note that $z_{i,\mathrm{T}}$ is the expected profit of firm i in a period in which firm i is a party to a punishment phase that occurs along the equilibrium path, excluding the profits it earns from the markets in $N_\varnothing$.

Outside of a punishment phase, the continuation value of the game to firm i from selecting the participation stage action a_i^P is,

$$V_i^+ = n_i \pi^*(1) + \varepsilon_i + n_\varnothing \pi^*(I) + \delta(1-\sigma_i)V_i^+ + \delta\sigma_i\big(z_{i,\mathrm{T}} + n_\varnothing \pi^*(I) + \delta V_i^+\big).$$

Solving for V_i^+ yields the expected profit of firm i. If, instead, firm i deviates, entering every market belonging to a rival firm, the continuation value of the game is,

$$V_i^- = n_i \pi^*(1) + \sum_{j \neq i} n_j \pi^*(2) + n_\varnothing \pi^*(I) \\ + \delta\Big(n_i \pi^*(I) + \sum_{j \neq i} n_j \pi^*(2) + n_\varnothing \pi^*(I)\Big) + \delta^2 V_i^+.$$

A necessary condition for cartel stability is $V_i^+ - V_i^- \geq 0$ or,

$$(1+\delta)\frac{n_i \pi^*(1) + \varepsilon_i + \delta\sigma_i z_{i,\mathrm{T}}}{1+\delta\sigma_i} \\ - \Big(n_i \pi^*(1) + \delta n_i \pi^*(I) + (1+\delta)\sum_{j \neq i} n_j \pi^*(2)\Big) \geq 0. \tag{11}$$

Taking the limit of this inequality as $\sigma(i, K, d_K) \to 0$ for all $i \in I$, $K \subseteq I \setminus \{i\}$ and d_K (and therefore $\varepsilon_i \to 0$ and $\sigma_i z_{i,\mathrm{T}} \to 0$), and solving for δ yields the limit of δ_{T}. Notice that the difference $V_i^+ - V_i^-$ is continuous in each of its arguments and therefore δ_{T} must likewise be a continuous function of the continuous variables.

It only remains to establish that firm i will not deviate during a punishment phase if $V_i^+ - V_i^- \geq 0$. Suppose that a punishment phase occurs in period t. Firm i has no incentive to contest fewer markets than the enforcement mechanism dictates as each market contested delivers firm i a positive profit. Moreover, if firm i enters a market that it was not supposed to contest, this action triggers a new punishment phase in period $t + 1$. The gain to firm i from such a deviation must be less than the gain from entering every market belonging to a rival firm when all remaining firms implement the cartel agreement. ■

Proposition 5 demonstrates that where the probability of each error is small, the critical discount factor is in the neighbourhood of (10). This means that a cartel utilising targeted enforcement is less stable than a cartel utilising single-period un-targeted enforcement as,

$$\lim_{\sum_{i \in I} \sigma(i) \to 0} \delta_{\mathrm{U}}^1 = \max_{i \in I} \left[\frac{\sum_{j \neq i} n_j \pi^*(2)}{n_i\big(\pi^*(1) - \pi^*(I)\big) - \sum_{j \neq i} n_j \pi^*(I)} \right].$$

Although, it should be noted that the two mechanisms are equivalent if there are only two firms in the cartel. A cartel may prefer targeted enforcement if it delivers firms higher expected profits. This is a possibility because, along the equilibrium path, firms are less likely to be involved in a punishment phase under targeted enforcement.

4.4 Proportional response enforcement

An obvious alternative to targeted enforcement is *proportional response enforcement*. Proportional response enforcement functions in exactly the same way as targeted enforcement except that the magnitude of punishments scale with the size of the initial deviations. Whereas, under targeted enforcement, a deviation results in a complete breakdown in the bilateral relationship between instigator and victim for the duration of the punishment phase, under proportion response enforcement punishments 'fit the crime'.

Suppose that firm i enters d_j markets belonging to firm j. In the punishment phase, firm j is required to retaliate proportionally, entering $n_i d_j / n_j$ of firm i's markets. If $n_i d_j / n_j$ is not an integer, firm j can enter the required number of markets in expectation, conditioning entry into the final market on a random variable that is observed by all cartel members. During the punishment phase firm i continues its presence in the d_j markets it entered in the previous period but expands no further. Any expansion of firm i's presence in firm j's markets during the punishment phase is regarded as a fresh deviation and results in another punishment phase in the subsequent period.

The following proposition demonstrates that proportional response enforcement outperforms targeted enforcement in terms of both expected profits and stability. For the purposes of the following proposition it is useful to relabel the victims of a deviation (the firms in K) such that $\frac{d_1}{n_1} \geq \frac{d_2}{n_2} \geq \cdots \geq \frac{d_{\|K\|}}{n_{\|K\|}}$.

Proposition 6 *The expected profit to firm i from participating in a cartel employing proportional response enforcement is,*

$$\Pi_{i,\mathrm{PR}} = \frac{1}{1-\delta}\left(\frac{n_i \pi^*(1) + \varepsilon_i + \delta\sigma_i z_{i,\mathrm{PR}}}{1+\delta\sigma_i} + n_\varnothing \pi^*(I)\right) \geq \Pi_{i,\mathrm{T}},$$

with strict inequality if $n_j \geq 2$ for some $j \in I$, and where,

$$z_{i,\mathrm{PR}} = \frac{1}{\sigma_i} \sum_{\substack{K \in I \setminus \{i\} \\ d_K}} \sigma(i, K, d_k) \Bigg[n_i \Bigg(\left(1 - \frac{d_1}{n_1}\right) \pi^*(1)$$

$$+ \sum_{k=1}^{\|K\|-1} \left(\frac{d_k}{n_k} - \frac{d_{k+1}}{n_{k+1}}\right) \pi^*(k+1) + \frac{d_{\|K\|}}{n_{\|K\|}} \pi^*(\|K\|+1) \Bigg) + \sum_{k \in K} d_k \pi^*(2) \Bigg]$$

$$+ \frac{1}{\sigma_i} \sum_{\substack{j \neq i \\ i \ni K \in I \setminus \{j\} \\ d_K}} \sigma(j, K, d_k) \Bigg[(1 - d_i) \pi^*(1) + d_i \pi^*(2)$$

$$+ n_j \left(\sum_{k=i}^{\|K\|-1} \left(\frac{d_k}{n_k} - \frac{d_{k+1}}{n_{k+1}}\right) \pi^*(k+1) + \frac{d_{\|K\|}}{n_{\|K\|}} \pi^*(\|K\|+1) \right) \Bigg].$$

The critical discount factor $\delta_{\mathrm{PR}} \leq \delta_{\mathrm{T}}$, with strict inequality if $n_j \geq 2$ for some $j \in I$. Moreover,

$$\lim_{\sum_{i\in I}\sigma(i)\to 0} \delta_{\mathrm{PR}} = \lim_{\sum_{i\in I}\sigma(i)\to 0} \delta_{\mathrm{T}}.$$

Proof. Along the equilibrium path, proportional response enforcement only differs from targeted enforcement in the expected profit of a firm i, in a period in which firm i is a party to a punishment phase. For proportional response enforcement this is $z_{i,\mathrm{PR}}$. Note that $z_{i,\mathrm{PR}} \geq z_{i,\mathrm{T}}$, with strict inequality if $n_j \geq 2$ for some $j \in I$. The value and lower bound of $\Pi_{i,\mathrm{PR}}$ then follows from the proof of proposition 5.

In common with targeted enforcements, the worst case deviation occurs when firm i enters every market belonging to a rival firm. It follows from the proof of proposition 5 that a necessary condition for cartel stability is,

$$(1+\delta)\frac{n_i\pi^*(1)+\varepsilon_i+\delta\sigma_i z_{i,\mathrm{PR}}}{1+\delta\sigma_i} - \left(n_i\pi^*(1)+\delta n_i\pi^*(I)+(1+\delta)\sum_{j\neq i} n_j\pi^*(2)\right) \geq 0. \quad (12)$$

Taking the limit of this inequality as $\sigma(i, K, d_K) \to 0$ for all $i \in I$, $K \subseteq I\setminus\{i\}$ and d_K (and therefore $\varepsilon_i \to 0$ and $\sigma_i z_{i,\mathrm{PR}} \to 0$), and solving for δ yields the limit of δ_{PR}.

Given that $z_{i,\mathrm{PR}} \geq z_{i,\mathrm{T}}$, if (11) is satisfied for some δ then (12) must also be satisfied. Moreover, if $n_j \geq 2$ for some $j \in I$, and therefore $z_{i,\mathrm{PR}} > z_{i,\mathrm{T}}$, there must exist values of δ for which (12) is satisfied and (11) is not. Finally, from the proof of proposition 5 it follows that if (12) is satisfied, no firm has an incentive to deviate during a punishment phase. ■

Proposition 6 is significant as, to the best of our knowledge, this is the first game theoretic justification for the use of proportional response in self-enforcing contracts. Intuitively, the probability of a firm being involved in a punishment phase along the equilibrium path is the *same* as under targeted enforcement — under both mechanisms participation is confined to the deviating firm and its victims — however the expected profits from the punishment phase will be higher if there is a chance that the punishment will not encompass all of a firm's markets.

5 Applications

In this section we consider two examples in which the existence of an atypical market gives rise to forms of market sharing than would not usually be predicted by models of collusion.

5.1 Oligopolistic competition with a collusive fringe

Suppose that a single large market L exists alongside the set of identical small markets N. Let $\pi^*_L(\cdot)$ denote the profit function of the large market, while $\pi^*(\cdot)$ is the profit function for each of the small markets. We assume that competitive profits in the large market satisfy,

$$\pi^*_L(I) > \frac{\|N\|}{\|I\|-1}\pi^*(1) - \|N\|\pi^*(I).$$

It follows from (5) that if L is assigned to a firm by the cartel, extensive margin collusion will not be stable for any $\delta \in (0,1)$.

The large market creates a problem for the cartel because any firm that is excluded from the large market has an overwhelming incentive to enter. In order to form a stable agreement, L must be assigned to the contested component of the partition $N_\varnothing$. This neutralises the large market as a determinant of cartel stability, because the markets in $N_\varnothing$ are contested by every firm regardless of history. The cartel can then partition the small markets in N between its members, as it would have in the absence of the large market.

The lesson here is that we cannot use the degree of competition in a large market as an indicator of whether or not collusion is occurring in small peripheral markets. It is entirely possible to have *oligopolistic competition with a collusive fringe.*

There are a number of market structures that may display a collusive fringe. Consider, for example, the market for beer or sodas. The firms in these markets tend to compete vigorously with one and other, selling their products through supermarkets and liquor stores. At the same time these same firms sign exclusive deals with restaurant chains, sporting venues and entertainment venues; effectively partitioning the small client relationships peripheral to the main consumer market. Another environment in which a collusive fringe may be found is where a major population centre is surrounded by a number of small regional centres. A collusive fringe may exist where a number of firms compete within the major population centre while avoiding contact in the smaller regional markets.

Of course, neither exclusive dealing nor geographic monopoly necessarily imply the existence of a collusive fringe. The key to detecting a collusive fringe lies in identifying the duopoly profit from the small markets. If the duopoly profit less discounted entry cost is positive, the partitioning of these markets is not consistent with competitive behaviour and we can conclude that we are observing collusion at the extensive margin.[11]

5.2 Predatory entry

Thus far, we have assumed that all firms have an expansion incentive; that is, absent a collusive agreement they would expand to all markets. Here we relax this assumption. Consider a multi-market game with two firms ($I = \{1, 2\}$) and two markets ($N = \{m, d\}$). Market m is assumed to be a natural monopoly ($\pi^*_m(1) > 0 > \pi^*_m(2)$), while market d is a natural duopoly ($\pi^*_d(1) > \pi^*_d(2) > 0$). Suppose that firm 1 is located in the natural monopoly market while firm 2 is located in the natural duopoly. Under what circumstances is this partition of markets stable?

The presence of the natural monopoly market introduces asymmetric incentives into the game. Firm 1 has an incentive to enter market d in order to

[11] An earlier version of this paper showed that entering and colluding at the intensive margin in such larger markets would aid in the sustainability of collusion overall. Thus, there are circumstances where extensive and intensive margin collusion can be complements. Nonetheless, the notion of oligopolistic competition with a collusive fringe is highlighted here to demonstrate that just because one might observe firms competing in some markets (especially a large and prominent one), does not mean that collusion is absent in peripheral markets.

attain duopoly profits. In contrast, firm 2 has no interest in entering market m as doing so forces the profits in market m below zero. Nevertheless, so long as punishments are temporary firm 2 may be able to use the threat of *predatory entry* into market m to enforce the cartel agreement.

We define *predatory entry* to be entry by a firm into a market with the purpose of reducing the profits of that market below zero. In contrast to predatory pricing, the goal of predatory entry is not to force rival firms out of the market in which the losses are occurring, but rather to force a rival to exit a second market in which both firms can coexist profitably.

Suppose that firm 2's strategy is to punish each deviation by firm 1 with a single period of predatory entry, consistent with the rules of the targeted enforcement discussed in section 4. From (10) it follows that the threat of predatory entry is sufficient to deter firm 1 from entering market d so long as,

$$\delta \geq \frac{\pi_d^*(2)}{\pi_m^*(1) - \pi_m^*(2) - \pi_d^*(2)}.$$

Here the fact that $\pi_m^*(2) < 0$ enhances cartel stability as it increases the cost of the punishment that follows entry. We do not have to establish an equivalent condition for firm 2 as the return to entering m is negative.

It is, however, necessary to verify that firm 2's threat is credible. Firm 2 must weigh the cost of entering market m as a duopolist for one period against the permanent loss of monopoly profits in market d. It follows that predatory entry is a credible threat if,

$$\pi_m^*(2) + \pi_d^*(2) + \frac{\delta}{1-\delta}\pi_d^*(1) \geq \frac{1}{1-\delta}\pi_d^*(2),$$

implying,

$$\delta \geq \frac{-\pi_m^*(2)}{\pi_d^*(1) - \pi_d^*(2) - \pi_m^*(2)} \in (0,1),$$

which must be satisfied if firms are sufficiently patient.

6 Conclusion

This paper has taken the standard approach to modelling tacit collusion where firms might compete in multiple markets and added a distinct market participation choice in the stage game. Consequently, collusion can be based on the history of participation decisions alongside any history of within market actions taken by firms. In the process, we have characterised collusive equilibria where firms in a cartel allocate markets amongst themselves and engage in mutual avoidance as opposed to multi-market contact. Significantly, we have demonstrated that such collusive outcomes can be more stable as the intensity of stage game duopoly competition becomes high; in the limit, the critical discount factor converges to zero regardless of the number of firms in a cartel. We have also demonstrated how enforcement mechanisms that are more proportionate to the scope of deviations may permit the same prediction of cartel stability than broader enforcement mechanisms but be more forgiving of misinterpretations or errors.

There, of course, remain directions for future research that would establish whether the form of collusion highlighted in this paper is of relevance. The challenge for empirical work is identifying firms that could compete but have chosen not to as part of a tacit agreement. This might also be identified by examining the patterns of participation in response to an increase in firm ownership and asymmetries between firms following mergers.

Another avenue for investigation would be to embed the model of collusion here in a model of antitrust enforcement (such as Harrington, 2008; Spagnolo, 2008; Choi and Gerlach, 2009). Those models consider the fact that antitrust authorities often rely on whistleblowers to identify cartels and that, under tacit collusion at the intensive margin, the number of people who might have knowledge of a conspiracy may be large. For example, in a model of multi-market contact, the responsibility for implementing the cartel agreement within each firm will likely fall on members of senior management who are in a position to direct the firm's activities in each market covered by the cartel agreement. In addition, the firm's management in each market must either be party to the agreement, or have knowledge of its existence. These local managers will be required to move the firm's actions away from its best response, and refrain from any activities that would see the firm steal business from its cartel partners.

Compare this with an extensive margin collusive agreement across the same set of markets. The involvement of senior management remains. However, because the cartel agreement requires each firm to compete as oligopolists in contested markets, and as a monopolist where no rival firm is present, market level management need never know of the cartel's existence. Of course, it could also be the case that the very act of avoidance of what would seem to be obvious market opportunities for firms (as was the trigger to action in *Twombly*), could itself bring a set of firms under scrutiny. A model that embedded enforcement with cartel stability could tease these effects out.

Finally, we have not modelled how collusive agreements come to be formed. As is well know, the coordination problem with repeated games is a challenge for explaining how collusion at the intensive margin arises. It strikes us that collusion at the extensive margin may arise in an uncoordinated fashion. For example, two chains may start on separate parts of the country and slowly expand. Just as they are about to overlap, they understand the potential consequences of such competition — perhaps through head to head competition in a small set of areas. Those areas may remain competitive while the historic locations are monopolized. The issue of the evolution of collusion is something that we leave for future research.

References

Belleflamme, P. and F. Bloch (2008), Sustainable Collusion on Separate Markets, *Economic Letters* 99 (2), pp. 384–386.

Bernheim, B. D. & M. D. Whinston (1990), Multimarket Contact and Collusive behaviour, *RAND Journal of Economics* 21 (1), pp. 1–26.

Bond, E. W. and C. Syropoulos (2008), Trade Costs and Multimarket Collusion, *RAND Journal of Economics* 39 (4), pp. 1080–1104.

Choi, J. P. and H. Gerlach (2009), International Antitrust Enforcement and Multimarket Contact, *mimeo.*, Michigan.

Edwards, C. (1955), Conglomerate Bigness as a Source of Power, *Business Concentration and Price Policy* (G. Stigler ed.), Princeton: NJ, pp. 331–360.

Gans, J. S., R. Sood and P.L. Williams (2004), The Decision of the High Court in Rural Press: How the literature on credible threats may have materially facilitated a better decision, *Australian Business Law Review* 33 (5), pp. 337–344.

Green, E. J. and R. H. Porter (1984), Noncooperative Collusion Under Imperfect Price Information, *Econometrica* 52 (1), pp. 87–100.

Gross, J. and W. Holahan (2003), Credible Collusion in Spatially Separated Markets, *International Economic Review* 44 (1), pp. 299–312.

Harrington, J. E. (2008), Detecting Cartels, *Handbook of Antitrust Economics* (P. Buccirossi Ed.), MIT Press, Chapter 6.

Spagnolo, G. (2008), Leniency and Whistleblowers in Antitrust, *Handbook of Antitrust Economics* (P. Buccirossi Ed.), MIT Press, Chapter 7.

Stigler, G. (1964), A Theory of Oligopoly, *Journal of Political Economy*, 72 (1), pp. 44–61.

Stone, B. and M. Helft (2010), Apple's Spat with Google is Getting Personal, *New York Times*, March 12, (http://www.nytimes.com/2010/03/14/technology/14brawl.html?src=tptw&pagewanted=all).