Collusion at the Extensive Margin*

Martin C. Byford†
Joshua S. Gans‡

Abstract

This paper is the first to examine collusion at the extensive margin (whereby firms collude by avoiding entry into each other's markets or territories). We demonstrate that such collusion offers distinct predictions for the role of multiple markets in sustaining collusion such as the use of proportionate response enforcement mechanisms, the possibilities of oligopolistic competition with a collusive fringe, and predatory entry. We argue that collusion at the extensive margin poses difficult issues for antitrust authorities relative to its intensive margin counterpart.

Release
This Version: March 2011

J.E.L. Classification: C73, L41.
Keywords: Collusion, Credible Threats, Proportional Response, Segmented Markets, Multi-Market Contact, Predatory Entry, Markov Perfect Equilibrium.

* We would like to thank John Asker and seminar participants at Harvard University, the US Department of Justice, New York University, Northeastern University, ANU and The University of Colorado at Boulder for comments on earlier drafts of this paper.

† School of Economics, Finance and Marketing, RMIT University. Building 108 Level 12, 239 Bourke Street, Melbourne VIC 3000, Australia. Phone: +61 3 9925 5892. Fax: +61 3 9925 5986. Email: martin.byford@rmit.edu.au.

‡ Melbourne Business School, University of Melbourne.

1 Introduction

The standard treatment of collusion in economics involves examining the sustainability of attempts by firms in the same market to coordinate on high prices or to restrict quantity in that market. That is, where firms collude at the *intensive* margin. The main issue in sustaining such collusion is one the detectability of deviations (Stigler, 1964) and the incentives to punish off the equilibrium path (Friedman, 1971; see Shapiro, 1989 for a review). Even where interaction across multiple markets is considered, the focus remains on collusion at the intensive margin of each market (Bernheim and Whinston, 1990). Given this theoretical base, most policy discussions and empirical analyses are focused on collusion of this type (Jacquemin and Slade, 1989; Feuerstein, 2005; Porter, 2005).

An alternative form of collusion is one that takes place at the *extensive* margin. In this case, firms collude by coordinating *participation* across markets or market segments in order to avoid contact; leaving each firm as a monopolist in one or more markets. 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]

[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).

Waikerie promptly exited Mannum. The Australian courts found that this was an anti-competitive agreement and fined both parties.[2] Note that this did not involve attempted collusion within the Mannum area but instead a division of geographic markets along the Murray 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 anti-trust 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 behavior rather than the collusive outcome itself. Indeed, in 2007, in *Bell Atlantic v. Twombly*[3] 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 behavior 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 behavior. Indeed, we would go further and argue that identifying collusion at the extensive margin is a significant challenge for antitrust enforcers as it can be implemented across multiple markets with little but 'top level' managerial knowledge. As a consequence, it is likely to be an area of actual practice by firms. For that reason, it deserves explicit study by economists.

In this paper, we develop a framework for understanding collusion at the extensive margin. Like other analyses of collusion our focus is on enforcement of the collusive

[2]For an account see Gans, Sood and Williams (2004).

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

agreement. To provide a clear point of contrast with intensive margin collusion, we utilize the Markov perfect equilibrium requirement to screen out elements of such collusion.[4] We do this by assuming that there are (possibly infinitesimally small) costs of entering each individual market and that such decisions are observable but take time, permitting rival firms to implement a response in market before entry is completed. This stands in contrast to other treatments, such as Fershtman and Pakes (2000), who include a dummy variable as a proxy for history. Our approach allows firms to condition their Markov strategies on the profile of firm participation in markets alone. We argue that this is a realistic representation of possibilities in many markets.

We assume that there exists several, clearly-defined markets that firms can participate in. The most natural interpretation of such markets is geographic but distinctions might also be on the basis of other characteristics such as product category. 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.' Of course, it is also possible that markets might be divided up on a buyer-by-buyer basis with exclusive supply agreements being signed and unchallenged. Here, we take the set of markets as given, although it is useful to note that their definition may well be endogenous in reality.

In a natural first result, we find that for a sufficiently high discount factor, there exists a Markov perfect equilibrium whereby firms divide up the markets between them into individual monopolies so long as each individual firm earns more from their set of allocated monopolies than they would if there was competition across all markets. This collusive outcome coexists with other potential equilibria including a competitive one that itself acts as the sustained grim trigger punishment mechanism enforcing the collusive equilibrium. While the conditions for the existence of such an equilibrium share properties associated with collusion at the intensive margin (including patience as well as the strength of the competitive equilibrium), it also identifies the need for *balance*: that is, each firm must be allocated at least one market for themselves. Consequently, there must be at least as many markets as firms for the strongest collusive equilibrium to exist.

[4]Although we do revisit the model to consider the interaction between intensive and extensive margin collusion below.

At this point, it is useful to place this result within the context of the existing literature on collusion. Collusion at the extensive margin requires that there exists multiple markets that can be partitioned and allocated amongst participants. There is, in fact, a long-standing literature that identifies firm interactions across multiple markets as making collusion more likely. The insight began with Edwards (1955):

> The interests of great enterprises are likely to touch at many points, and it would be possible for each to mobilize at any one of these points a considerable aggregate of resources. The anticipated gain to such a concern from unmitigated competitive attack upon another large enterprise at any point of contact is likely to be slight as compared with the possible loss from retaliatory action by that enterprise at many other points of contact ... Hence, the incentive to live and let live, to cultivate a cooperative spirit, and to recognize priorities of interest in the hope of reciprocal recognition. (p.335)

Edwards was, in fact, arguing why larger firms may find it more likely to refrain from competing with each other rather than with smaller firms who possess less retaliatory power to keep larger firms at bay. However, many have interpreted this notion of "mutual forbearance" as an argument as suggesting that contact across multi-markets can soften competition or facilitate tacit collusion rather than multiplicity giving rise to the potential for avoidance (Feinberg, 1984, 1985; Bernheim and Whinston, 1990). Indeed, Edwards appeared to be considering the latter:

> Those attitudes support such policies as refraining from sale in a large company's home market below whatever price that company may have established there; refraining from entering into the production of a commodity which a large company has developed; not contesting the patent claims of a large company even when they are believed to be invalid; abstaining from an effort to win away the important customers of a large rival; and sometimes refusing to accept such customers even when they take the initiative. (p.335)

That is, collusion may arise where firms explicitly *avoid* contact rather than when they are observed to engage in such contact.[5]

[5] In some situations, it may be difficult to distinguish between avoidance and actual contact. For instance, in their study of collusion in the Ohio school milk market, Porter and Zona (1999) found that two suppliers in the Cincinatti area likely colluded by refraining from bidding aggressively for

The seminal study of multi-market contact and collusion is Bernheim and Whinston (1990) – hereafter BW. They ask whether participation in multiple markets is likely to lead to more sustainable collusion at the intensive margin by pooling incentive constraints not to deviate. BW demonstrates that when firms and markets are symmetric, multi-market contact does not assist in sustaining collusion as a firm that is considering deviating in one market should deviate in all markets as the punishment would be the same. Under symmetry, markets and hence, incentives are separable. This, of course, identifies asymmetry as a key reason why multi-market contact may facilitate collusion.

In contrast, symmetry actually facilitates collusion at the extensive margin. Indeed, we demonstrate that collusion is sustainable and facilitated by the presence of multiple markets even under the symmetry conditions underpinning BW's irrelevance result. Alternatively, say, when there is a single large 'desirable' market, its allocation to a single firm makes it more difficult for others to satisfy incentive constraints not to contest it. In contrast, BW show that in this situation, the monopoly rents available from collusion at the intensive margin in the large market can be used to relax incentive constraints in other markets. In this respect, the two forms of collusion are distinct.[6] Nonetheless, as we demonstrate below, there are situations in which the ability to use both forms of collusion can complement one another.

In addition to providing some distinct empirical predictions, examining collusion at the extensive margin explicitly allows us to analyze enforcement mechanisms that target the transgressor. Specifically, an all out war across many markets involving all firms following an isolated transgression is something that, while game theoretically justified, is an enforcement mechanism that many would find unrealistic. Consequently, it is interesting to examine 'weaker' enforcement mechanisms to examine their effectiveness in sustaining collusion. We demonstrate that, in some circumstances, a proportionate response (i.e., I enter one of your markets if you enter one of mine) is as sustainable as a broader enforcement mechanism. When there is uncertainty, we demonstrate that expected payoffs are higher under proportionate response

their rival's customers. In this case, the fact that they participated at all in bids outside of their designated area can be considered as contact although a weak bid may equally be considered as avoidance.

[6] BW (1990) do provide some examples whereby collusion takes the form of refraining from activity (and perhaps participation) in a each other's markets. This arises when firms have asymmetric advantages of operating in other markets — say, due to the existence of transportation costs. Multiple markets allows the forms to specialize and gain productive efficiencies that would not be possible if say, one firm was unable to participate in the rival's market. In this respect, the existence of transportation costs can facilitate both intensive and extensive margin collusion.

than other mechanisms providing the first game-theoretic justification for this natural punishment outcome.

Finally, we consider implications for antitrust policy. One issue is that the existence of a large or valuable market might destablize the collusive division of markets. We show that the addition of such a market has no impact on the scope of collusion over the remainder with firms competing in the large market but monopolizing the others. The end result is best characterized as an oligopoly with a collusive fringe, something that would make anti-trust detection difficult. A second issue is where firms do not have an incentive to enter all markets perhaps because some markets are a natural monopoly. In this situation, punishing firms who have monopolies in such markets may appear difficult but can occur via a process of predatory entry. That is, as punishment, a firm enters the natural monopoly market temporarily with below-cost pricing until such time as withdrawl from their own markets occurs. A third issue is whether mergers might act to de-stablize a cartel. We demonstrate that all mergers that do not involve the 'smallest' participant in the cartel have this potential. Finally, we argue that extensive margin collusion requires less 'middle manager' buy in than intensive margin collusion across multiple markets.

The paper proceeds as follows: Section 2 sets out the structure of a *multi-market game*; the framework through which we examine collusion at the extensive margin. Section 3 examines collusion in a multi-market game under perfect information. Here we show that irrelevance result of BW (1990) does not extend to collusion at the extensive margin.

Uncertainty is introduced into the model in section 4. We introduce three tit-for-tat enforcement mechanisms and show that proportional response dominates a disproportionate response both in terms of profits and parameter values. Several applications for the framework are considered in section 5 including the possibility that oligopolistically competitive large markets may display a *collusive fringe*, the potential for *predatory entry* and the potential for anti-trust enforcement to enhance cartel stability. The paper concludes with a discussion of the model and results.

2 The Model

This section sets out the multi-market framework that serves as the general setting for our analysis of collusion at the extensive margin. We augment the multi-market game developed by BW (1990) by including an explicit mechanism for firm entry into and withdrawal from individual markets.

Consider an infinite horizon, discrete time, dynamic game in which a finite set I of

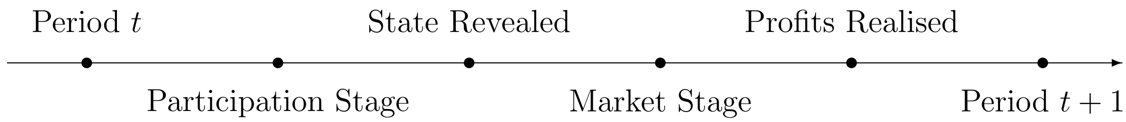

FIGURE 1: Timing

firms interact repeatedly over a finite set N of discrete markets (or market segments). It is assumed that $I \geq 2$ while $\|N\| \geq \|I\|$.[7] All firms discount the future by the common discount factor $\delta \in (0,1)$.

Each period of the game begins with the *participation stage*. Formally, in the participation stage of period t each firm i selects an action $a_i^t \subseteq N$. 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 firm i will absent itself from market n in the current period.

Firm i is said to *enter* (resp. *exit*) 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 into market n costs firm i an amount $c_{i,n} > 0$. The entry cost is only incurred in the period in which entry occurs, the cost of maintaining a presence in a market following entry is assumed to be accounted for in the market's instantaneous profit function outlined below. If a firm exits and subsequently reenters a market the entry cost must be paid again.

Following the participation stage the profile of firm participation $a^t = \{a_i^t\}_{i \in I}$ is revealed to the market. The participation profile represents the state of the world and belongs to the state space $a^t \in 2^{N^I}$, the I-fold cartesian product of the power set of N.

Competition between firms occurs in the *market stage*. Markets are modelled as a possibly interdependent simultaneous moves games. In the market stage of a period t each firm i selects an action $x_{i,n}^t \in X_{i,n}$ for all $n \in a_i^t$. In the interests of expositional simplicity it is assumed that the set of actions available to firm i in market n is independent of the number and identities of rival firms engaged in the market. Aggregating across markets, the actions of firm i in the market stage are represented by the vector $x_i^t = \{x_{i,n}^t\}_{n \in N}$ with $x_{i,n}^t = \varnothing$ for all $n \notin a_i^t$ while $x^t = \{x_i^t\}_{i \in I}$.

Following their choice of actions firms receive instantaneous profits from each market in which they are a participant. The instantaneous profit to a firm i from market n in period t is given by the function $\pi_{i,n}(x^t)$. In general we permit the instantaneous profits in one market to experience externalities from actions taken in

[7]The notation $\|N\|$ refers to the cardinality of the set N. The sets of firms and markets are assumed to be finite.

other markets. The present value of firm i's lifetime profits is the sum,

$$\Pi_i = \sum_{t=0}^{\infty} \delta^t \Bigg(\sum_{n \in a_i^t} \pi_{i,n}(x^t) - \sum_{n \in a_i^t \setminus a_i^{t-1}} c_{i,n} \Bigg),$$

where $a_i^t \setminus a_i^{t-1} \subseteq N$ is the set of markets firm i enters in period t. The timing of the model is set out in figure 1.

2.1 The Intuition Behind the Model

In BW (1990) a number of examples are developed in which optimal multi-market collusion requires one or other of the colluding firms to refrain from trading in any given market. Along the equilibrium path the consequence of one firm's inaction is that the other member of the cartel effectively gains monopoly control of the market. However, the inactive firm does not truly exit the market. Rather, the inactive firm *lurks* in the market, maintaining a passive presence and with it the ability to rapidly ramp up production, possibly stealing the entire market, before the active firm has the opportunity to respond.

The key technical innovation of the present paper is the inclusion of the participation stage in the dynamic game. By explicitly including participation decisions within our framework we distinguish between a firm lurking within a market and a firm that is absent from a market entirely.

This distinction is important as a firm that is initially absent from a market cannot simply appear, catching all incumbent firms by surprise. To the contrary, the firm must first undertake the process of entry.

Entry into a market is typically a complex process, observed by incumbent firms. In order to enter a market a firm may need to construct or acquire production and retail premises, hire a local workforce, acquire market specific licences and regulatory approval, and initiate marketing activities in order to connect to customers. Such is the complexity of entry that in many markets it is reasonable to assume that incumbent firms will be able to adjust their strategic behaviour within the market faster than an outsider firm can complete the process of entry. In such a market incumbent firms have the opportunity to adjust their strategic behaviour in anticipation of the entrant's arrival and post-entry behaviour. The model developed here captures this timing by assuming that the outcomes of the participation stage decisions are revealed prior to the selection of market stage actions.

2.2 Refining the Set of Equilibria

The multi-market framework typically produces a large number of sub-game perfect equilibria. Moreover, any given sub-game perfect equilibrium may include elements of both collusion at the intensive margin (coordination of actions within markets) and collusion at the extensive margin (mutual forbearance across markets). In order to facilitate the analysis of collusion at the extensive margin we wish to refine the set of equilibria, screening out all equilibria in which firms employ strategies that are contingent on the history firm behaviour in the market stage. Fortunately, the Markov perfect equilibrium (MPE) refinement has exactly this effect.

An MPE is a sub-game perfect equilibrium in which all firms employ Markov strategies; strategies that depend only on the payoff relevant information of the game. Note that while firms must employ Markov strategies in an MPE, the equilibrium must be robust against a unilateral deviation by any firm to any feasible strategy, including non-Markov strategies.

The payoff relevant information in the market stage of period t is the prevailing profile of firm participation a^t that indicates the number and identities of firms participating in each market. Given that firm i does not incur the entry cost $c_{i,n}$ for participating in market n if $n \in a_i^{t-1}$ the payoff relevant information in the participation stage of period t is the profile of firm participation from the preceding period. It follows that a Markov strategy for firm i can be written as a pair of functions $\big(a_i(a^{t-1}), x_i(a^t)\big)$.

In order to simplify our analysis we make the following assumption regarding the one shot Nash equilibrium to the market stage game.

Assumption 1: For all $a^t \in 2^{NI}$ there exists a (possibly mixed) strategy profile $x^*(a^t)$ that constitutes the unique Nash equilibrium for the one-shot game in which all markets are resolved simultaneously. The corresponding (expected) Nash equilibrium instantaneous profits are written $\pi^*_{i,n}(a^t)$.

MPE firm behaviour in the market stage is now completely characterised by the following lemma.

Lemma 1: *In an MPE* $x_i(a^t) = x_i^*(a^t)$ *for all* $a^t \in 2^{NI}$ *and* $i \in I$. *In words, in an MPE firms select their static Nash equilibrium strategies in the market stage.*

Proof. In an MPE the outcomes of the market stage do not affect the state of the game. It follows that in an MPE firms must select their market stage actions to maximise instantaneous profits given the current state. □

An immediate consequence of lemma 1 is that the static Nash equilibrium (expected) profits $\pi^*_{i,n}(a^t)$ are exactly the (expected) MPE instantaneous profits resulting from the market stage of each period.[8] The following assumption provides further structure to the MPE instantaneous profits:

ASSUMPTION 2 (Long-Run Expansion Incentive): For all $i \in I$, $m \in N$, and $a^t \in 2^{N^I}$ such that $m \notin a^t_i$,

$$\sum_{n \in a^t_i} \pi^*_{i,n}(a^t) < \sum_{n \in a^t_i \cup \{m\}} \pi^*_{i,n}\big(a^t_{-i}, a^t_i \cup \{m\}\big) - (1-\delta)c_{i,m}.$$

Moreover, for all $j \neq i$ such that $m \in a^t_j$,

$$\sum_{n \in a^t_j} \pi^*_{j,n}(a^t) > \sum_{n \in a^t_j} \pi^*_{j,n}\big(a^t_{-i}, a^t_i \cup \{m\}\big)$$

Intuitively, assumption 2 states that holding the participation of rival firms constant, it is always profitable for firm i to participate in an additional market m. Moreover, the increase in the present value of lifetime profits from entering m are more than sufficient to compensate for the cost of entry. It follows that firms never experience diseconomies of scale that would prohibit further expansion. However, any firm j that is present in the market m suffers a reduction in profits as a result of firm i's entry. The expansion incentive destabilises a cartel by providing colluding firms with an incentive to deviate. We consider the consequences of relaxing this assumption in section 5.

3 Collusion

In this section we characterise the strongest class collusive MPE that may arise in our multi-market setting: a grim-strategy equilibrium that closely parallels a grim-strategy equilibrium in super-game collusion. This collusive equilibrium is contrasted with a baseline competitive equilibrium which exists wherever assumptions 1 and 2 hold. The section concludes by contrasting the grim-strategy equilibrium with the multi-market contact model of BW (1990).

[8] Lemma 1 implies that in an MPE the structure of the market stage is irrelevant so long as outcomes are deterministic in expectation. For example, markets could take the form of auctions, bargaining or coalitional games. For the case of a coalitional game we require that the core is not empty and firms have consistent expectations as to which core outcome will arise for each participation profile. This structure is consistent with an infinitely repeated *bi-form* game (See Brandenburger and Stuart, 2007).

3.1 Simple Equilibria

We begin by establishing the existence of a baseline oligopolistically competitive MPE in which all firms enter and remain in all markets indefinitely; regardless of the actions of rival firms.

Proposition 1 (Competitive Equilibrium): *Consider a dynamic multi-market game satisfying assumptions 1 and 2. There exists an MPE, which we term the* competitive equilibrium*, such that the equilibrium Markov strategy satisfies* $a_i^*(a^{t-1}) = N$ *for all* $i \in I$, $a^{t-1} \in 2^{N^I}$ *and* $\delta \in (0, 1)$.

Proof. Given that all rival firms seek to enter every market regardless of the actions of rival firms, the long-run expansion incentive (assumption 2) makes expanding into every market a best response. □

Proposition 1 establishes the existence under general conditions of an MPE in which all firms contest every market. Formally, in a competitive equilibrium firm participation satisfies $a^t = N^I$ for all $t \in \{1, 2, \dots\}$ where N^I is the I-fold cartesian product of N. From lemma 1 it follows that within each market all firms behave in an oligopolistically competitive manner yielding firm i the oligopolistically competitive instantaneous profit $\sum_{n \in N} \pi_{i,n}^*(N^I)$ in each period. Proposition 1 is significant in a dynamic oligopoly setting as it implies that wherever firms implement a collusive equilibrium, they do so in an environment in which there exists a competitive equilibrium which is at least as robust.[9]

A direct corollary of proposition 1 is that a necessary condition of any non-competitive MPE is that at least two firms play strategies in the participation phase that are contingent on the past participation of rival firms in the game. Moreover, any firm that does not play a strategy that is contingent on the participation of rival firms, must play the strategy set out in proposition 1.

Definition 1 (Collusive Equilibrium): A *collusive equilibrium* is defined to be a steady state MPE in which the equilibrium strategy profile $a^*(a^{t-1})$ induces a partition of the set of markets $P = \left\{N_\varnothing, \{N_i\}_{i \in I}\right\}$. The partition is defined such that steady state equilibrium participation, denoted a^P, satisfies $a_i^P = N_i \cup N_\varnothing$ for all $i \in I$.

In words, a collusive equilibrium is an MPE in which firms divide up the markets between them. Along the equilibrium path firm i acts as a monopolist in all markets

[9] In contrast, the model dynamic oligopoly with sequential moves developed by Maskin and Tirole (1988a,b) may have MPE's that produce profits for firms that exceed competitive levels, however in their model an oligopolistically competitive outcome is not an MPE.

$n \in N_i$ while all firms contest the markets in the component $N_\varnothing$.[10] As lemma 1 dictates the MPE behaviour of all firms during the market stage, a collusive equilibrium is completely defined by the partition P and the participation stage component of the Markov strategy $a^*(\cdot)$.

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 an enforcement mechanism within this model is for any transgression to cause the game to permanently revert to the competitive equilibrium set out in proposition 1.

Proposition 2 (Grim-Strategy Equilibrium): *Consider multi-market game satisfying assumptions 1 and 2, a partition $P = \{N_\varnothing, \{N_i\}_{i\in I}\}$, and the participation stage strategy $a^{GS}(\cdot)$ such that $a^{GS}(a^{t-1}) = N$ if there exists $k \neq l$ such that $a_k^{t-1} \cap N_l \neq \varnothing$ and $a^{GS}(a^{t-1}) = N_i \cup N_\varnothing$ otherwise. The pair (P, a^{GS}) defines a collusive equilibrium if and only if δ satisfies,*

$$\delta \geq \delta^{GS} = \max_{i\in I}\left[\frac{\sum_{n\in N}\pi^*_{i,n}(a^P_{-i}, N) - \sum_{n\in\bigcup_{j\neq i}N_j} c_{i,n} - \sum_{n\in N_i\cup N_\varnothing}\pi^*_{i,n}(a^P)}{\sum_{n\in N}\pi^*_{i,n}(a^P_{-i}, N) - \sum_{n\in\bigcup_{j\neq i}N_j} c_{i,n} - \sum_{n\in N}\pi^*_{i,n}(N^I)}\right]. \quad (1)$$

Proof. From the proof of proposition 1 it is clear that once a single firm triggers a punishment phase by entering a rival's market ($a_k^{t-1} \cap N_l \neq \varnothing$ for some $k \neq l$) transition to the competitive behaviour outline in proposition 1 is sub-game perfect. Suppose that $a^{t-1} = a^P$. If a single firm i deviates in period t — selecting to enter a non-empty set of markets $Q \subseteq \bigcup_{j\neq i} N_j$ — the consequent progression of participation profiles becomes $a^t = (a^P_{-i}, a^P_i \cup Q)$ and $a^\tau = N^I$ for all $\tau \geq t+1$ as firms revert to the competitive equilibrium in periods $t+1$ onward. From assumption 2 it follows that the worst case deviation occurs where firm i enter the set of markets $Q = \bigcup_{j\neq i} N_j$ which is not profitable where,

$$\frac{1}{1-\delta}\sum_{n\in a^P_i}\pi^*_{i,n}(a^P) \geq \sum_{n\in N}\pi^*_{i,n}(a^P_{-i}, a^P_i \cup Q) - \sum_{n\in\bigcup_{j\neq i}N_j} c_{i,n} + \frac{\delta}{1-\delta}\sum_{n\in N}\pi^*_{i,n}(N^I).$$

Solving for δ yields (1). □

The grim-strategy mechanism requires that as soon as any firm k, is observed entering a rival firm l's market, all firms respond by entering and remaining in every

[10] In the interests of expositional simplicity we confine our attention to a class of partitions that are both relevant to applied work and tractable. The model can be readily extended to accommodate collusive agreements with richer structures. For example, we might define one or more additional components of the partition N_J such that the subset $J \subset I$ of firms share control of each market $n \in N_J$ while excluding the remaining firms. Alternatively, a subset of firms could rotate ownership of a given market. This extension is left as a topic for future research.

market in the game. Once mass entry occurs the game reverts to the competitive equilibrium outlined in proposition 1. It follows from (1) that a necessary condition for the stability of a grim-strategy equilibrium is,

$$\sum_{n \in a_i^P} \pi^*_{i,n}(a^P) > \sum_{n \in N} \pi^*_{i,n}(N^I). \tag{2}$$

That is, the profits each firm i receives as a result of retaining exclusive control of the markets in N_i must be higher than the profits firm i receives in a competitive equilibrium. Indeed, wherever (2) is satisfied there exists a $\delta \in (0, 1)$ satisfying (1).

Implicitly, proposition 2 provides an insight into the form of those partitions that may arise in a grim-strategy equilibrium. An immediate corollary of proposition 2 is that a necessary condition for the existence of a grim-strategy equilibrium is $N_i \neq \varnothing$ for all $i \in I$. Moreover, asymmetrically valuable markets may need to reside in $N_\varnothing$ in order to prevent creating an overwhelming incentive for rival firms to deviate.

The grim-strategy equilibrium requires each firm to punish every other firm in response to a single observed transgression. This is despite the fact that both the initial transgression, and the consequent punishments, may be targeted. Intuitively, a persistent punishment reduces the returns from the targeted market as well as reducing the ability of rival firms to inflict further discipline. Consequently, targeting a punishment on the transgressor alone increases the incentive for both transgressor and victim to engage in subsequent deviations. By applying a punishment to every firm, firms in a grim-strategy equilibrium avoid this problem by reducing the value of all markets simultaneously. Nonetheless, below we demonstrate that temporary targeted punishments are capable of relatively straightforward examination in supporting collusion at the extensive margin and, in the presence of uncertainty, may support higher equilibrium payoffs for cartel participants.

Finally, consider the role of entry costs in determining cartel stability. Where (2) holds δ^{GS} is decreasing in $\sum_{n \in \bigcup_{j \neq i} N_j} c_{i,n}$ as the cost of entry erodes the returns a firm receives from entering a rival's market. In the extreme case where,

$$\sum_{n \in \bigcup_{j \neq i} N_j} c_{i,n} \geq \sum_{n \in N} \pi^*_{i,n}(a^P_{-i}, N) - \sum_{n \in N_i \cup N_\varnothing} \pi^*_{i,n}(a^P) > 0$$

for all $i \in I$, deviating is not profitable for any $\delta \in (0, 1)$.

If increasing the entry costs enhances cartel stability then the worst case for a cartel is where $\sum_{n \in \bigcup_{j \neq i} N_j} c_{i,n} \to 0$ for all $i \in I$. In this case condition (1) becomes,

$$\delta^{GS} = \max_{i \in I} \left[\frac{\sum_{n \in N} \pi^*_{i,n}(a^P_{-i}, N) - \sum_{n \in N_i \cup N_\varnothing} \pi^*_{i,n}(a^P)}{\sum_{n \in N} \pi^*_{i,n}(a^P_{-i}, N) - \sum_{n \in N} \pi^*_{i,n}(N^I)} \right].$$

Nevertheless, where (2) holds we continue to have $\delta^{GS} \in (0,1)$ and as such the collusive agreement (P, a^{GS}) remains viable for sufficiently patient firms.

3.2 Regularity Conditions

In order to simplify the analysis from this point onwards we will sometimes impose one or more of the following three regularity conditions:

Definition 2 (Regularity Conditions): A set of markets are termed:

1. *Separable* if and only if for all $i \in I$ and $n \in N$ instantaneous profits $\pi_{i,n}(a^t)$ are independent of firm participation in all remaining markets $N \setminus \{n\}$.
2. *Identical* if and only if relabelling markets does not alter the entry costs and instantaneous profits to the firms participating in those markets.
3. *Symmetric* if and only if relabelling firms does not alter their entry costs and instantaneous profits for any given market.

Identicality and symmetry are regularity conditions which imply the independence of entry costs and MPE profits from the identity of markets and firms respectively. Markets are separable if the MPE instantaneous profits that a firm receives for participating in a market depends only on the identities of the firms who are currently active in that market. Separability implies that in equilibrium, participation decisions do not create externalities in other markets.

Formally, the MPE profits of markets that are identical, separable and symmetric depend only on the number of participants in the market. It follows that we can define a function $\pi^*(\cdot)$ such that $\pi^*(q)$ is the MPE instantaneous profit to each firm participating in a market when the total number of firms participating in the market is $q \in \{1, 2, \dots\}$. With a slight abuse of notation we write $\pi^*(I)$ to refer to the MPE instantaneous profits from a market with $\|I\|$ participants. The entry costs for separable, identical and symmetric markets take a value $c > 0$ that does not vary across firms or markets. Given the regularity conditions it is also useful to define $n_i = \|N_i\|$ and $n_\varnothing = \|N_\varnothing\|$.

Applying the three regularity conditions allows the constraint (1) from proposition 2 to be simplified revealing an important requirement of extensive margin collusion.

Proposition 3: *Consider an identical, separable and symmetric multi-market game satisfying assumptions 1 and 2. The partition* $P = \{N_\varnothing, \{N_i\}_{i \in I}\}$ *can be supported* *in a grim-strategy equilibrium wherever δ satisfies,*

$$\delta \geq \delta^{GS} = \max_{i \in I} \left[\frac{\sum_{j \neq i} n_j \big(\pi^*(2) - c\big)}{n_i \big(\pi^*(1) - \pi^*(I)\big) + \sum_{j \neq i} n_j \big(\pi^*(2) - \pi^*(I) - c\big)} \right]. \tag{3}$$

Moreover, a necessary condition for a grim-strategy equilibrium is,

$$\pi^*(I) < \frac{1}{\|I\|} \pi^*(1). \tag{4}$$

Proof. For (4) note that for $i = \text{argmin}_{k \in I}\, n_k$,

$$\pi^*(I) < \frac{n_i}{\sum_{j \in I} n_j} \pi^*(1) \leq \frac{1}{\|I\|} \pi^*(1),$$

where the first inequality follows from (2) by substitution and rearrangement, and the second inequality is a consequence of the fact that in a grim strategy equilibrium $n_i \geq 1$ for all $i \in I$. Condition (3) is derived from (1) by substitution. □

Condition (4) is intuitively appealing. It implies that in order for collusion at the intensive margin to be possible, increasing the intensity of competition in a market must reduce the aggregate profits received by firms. This is a common feature in models of oligopoly competition. Moreover, given that wherever $N \geq I$ we can define a partition P such that $n_i = 1$ for all $i \in I$ and $n_\varnothing = N - I$, it follows that for sufficiently high $\delta \in (0, 1)$ there exists a partition P that can be supported by a grim-strategy equilibrium wherever condition (4) is satisfied.

Condition 3 sheds more light on one key determinant of cartel stability. The firm i that maximizes 3, and therefore determines the level of the critical discount factor, will be the firm with the smallest partition ($n_i \leq n_j$ for all $j \in I$). This insight generalizes beyond the identical, separable and symmetric case. The more valuable are the markets in a firm's component of the partition, the less the firm has to gain and the more the firm has to lose if it instigates a deviation. Conversely, firms granted monopoly control over markets with a low aggregate value have the greatest incentive to violate the agreement.

3.3 Multi-Market Contact vs. Multi-Market Avoidance

BW (1990) examine the benefits that firms colluding at the intensive margin can derive from coming into contact across multiple markets.[11] For the purposes of the

[11] In the context of the present paper the equilibria examined by BW (1990) require all firms to participate in every market in every period ($a_i^t = N$ for all $t \in \{1, 2, \ldots\}$ and $i \in I$) while colluding on the actions taken within these markets.

present paper BW produce two key results: First, they prove that collusion at the intensive margins of multiple identical markets is no more stable than collusion at the intensive margin of a single representative market.[12] Intuitively, while multi-market contact does increase the magnitude of the punishments that may be imposed in the game, multi-market contact also results in a proportionate increase in the incentive to initiate a deviation. However, where markets are asymmetric, multi-market contact provides colluding firms with the possibility of smoothing participation constraints; utilizing the slack in the participation constraints in one market to facilitate collusion in a second market where the incentive constraints would not otherwise be satisfied.

In common with the model of multi-market contact, collusion at the extensive margin permits the characteristics of asymmetric markets to be smoothed. For each component of the partition P the constraint (1) aggregates the profits and entry costs of the constituent markets. Under grim-strategy enforcement the return that a firm derives from an individual markets is inconsequential so long as aggregate profits on and off the equilibrium path satisfy the participation constraints.[13]

In contrast to BW (1990), extensive margin collusion does derive stability from the presence of multiple identical markets. Extensive margin collusion requires multiple markets and as proposition 3 demonstrates these markets can be identical. Moreover, adding markets to the game tends to increase the set of partitions that can be be supported by grim-strategy enforcement as increasing the number of markets also increases the fineness with which the markets can be divided between firms as captured by the ratio $n_i / \sum_{j \in I} n_j$.

Under certain parameter values, collusion at the extensive margin may be more stable than collusion at the intensive margin. In such cases, the existence of multiple identical markets facilitates collusion by providing firms with the option to employ a more stable mechanism. The following example illustrates this phenomenon.

Example 1. Consider a two firm, two market game in which the markets are identical, separable and symmetric. Suppose that $\frac{1}{2}\pi^*(1) > \pi^*(2) > 0$ and consider grim-strategy collusion in which each firm controls one market. For the purpose of this example we assume that the entry cost c is arbitrarily close to zero. This is the worst case for an extensive margin collusive agreement as δ^{GS} is decreasing in c. From (3)

[12] Given that the set of MPEs to the multi-market game is a subset of the set of sub-game perfect equilibria (SPEs), the MPEs characterised in this paper remain equilibria of the game where players can employ history dependent strategies in equilibrium. Consequently, it is valid to compare the performance of extensive and intensive margin collusive agreements as competing equilibria within the set of potential SPEs. Moreover, hybrid collusive agreements with elements of both extensive and intensive margin collusion are potential SPEs.

[13] Characteristic smoothing can also be seen in the *predatory entry* example in section 5.

the critical discount rate is,

$$\delta^{GS} = \frac{\pi^*(2)}{\pi^*(1) - \pi^*(2)}.$$

Now consider an agreement in which firms collude at the intensive margins of both markets simultaneously. Suppose that each firm receives a profit of π^{coll} from each market in which they collude, while deviating nets a firm π^{dev} from each market in the period in which it deviates, followed by permanent reversion to the duopoly equilibrium. The critical discount rate δ^{IM} solves,

$$\frac{2}{1-\delta}\pi^{\text{coll}} \geq 2\pi^{\text{dev}} + \frac{2\delta}{1-\delta}\pi^*(2), \qquad \Longrightarrow \qquad \delta^{IM} = \frac{\pi^{\text{dev}} - \pi^{\text{coll}}}{\pi^{\text{dev}} - \pi^*(2)}.$$

Where firms compete by setting prices and the products for sale in the markets are close substitutes (implying $\pi^*(2) \to 0$) it follows that $\delta^{IM} > \delta^{GS} \to 0$. ■

Example 1 illustrates the importance that the nature of competition within a market plays in the overall stability of collusion at the extensive margin. As oligopolistically competitive profits fall relative to monopoly profits the return to initiating a deviation also falls while the relative magnitude of the subsequent punishment rises. The same is not true of collusion at the intensive margin. In highly competitive markets the return to undercutting a rival in a deviation can approach the monopoly return. It follows that extensive margin collusion may be more stable than intensive margin collusion in highly competitive markets, while the reverse would be true for markets in which either the nature of the strategic interaction or the degree of product differentiation leads to a softer competitive environment.

Moreover, example 1 clearly demonstrates that the two varieties of collusion may exist as substitutes. A cartel has the ability to select between the two collusive mechanisms but within each market collusion requires that firms either coordinate participation or strategic behaviour. The following example demonstrates one potential form of complementarity between intensive and extensive margin collusion.

Example 2. Consider the two firm game from example 1 augmented by the presence of a third identical, separable and symmetric market. Suppose that the only stable partition satisfies $n_1 = n_2 = n_\varnothing = 1$. The most robust collusive agreement at the extensive margin requires each firm to act as a monopolist in one market while competing as a duopolist in the third market. This agreement delivers each firm an instantaneous profit of $\pi^*(1)+\pi^*(2)$ each period and is stable where $\delta \geq \delta^{GS}$. Assuming that $3\pi^{\text{coll}} > \pi^*(1) + \pi^*(2)$ the cartel can increase its profitability by colluding at the intensive margin of all three markets, however this agreement will reduce cartel stability if $\delta^{IM} > \delta^{GS}$.

A third alternative is for the cartel to collude at the extensive margins of two markets and the intensive margin of the third market. If this agreement is enforced by the threat of permanent reversion to the competitive equilibrium then there are two ways in which a firm can cheat: A firm could deviate by entering its rival's market in the participation stage. Firms have the opportunity to react to the deviation in the market stage reverting to duopoly competition in both the target market and the third market, and reverting to the competitive equilibrium in all subsequent periods. This deviation is not profitable if,

$$\frac{1}{1-\delta}\left(\pi^*(1)+\pi^{\text{coll}}\right) \geq \left(\pi^*(1)+2\pi^*(2)\right)+\frac{\delta}{1-\delta}3\pi^*(2),$$

which in turn implies,

$$\delta \geq \frac{2\pi^*(2)-\pi^{\text{coll}}}{\pi^*(1)-\pi^*(2)} < \delta^{GS} < \delta^{IM},$$

where the second inequality follows from the assumptions $3\pi^{\text{coll}} > \pi^*(1)+\pi^*(2) > 3\pi^*(2)$. Alternatively, a firm could deviate in the market stage, claiming π^{dev} from the third market and triggering a reversion to the competitive equilibrium in the following period. A market stage deviation is not profitable if,

$$\frac{1}{1-\delta}\left(\pi^*(1)+\pi^{\text{coll}}\right) \geq \left(\pi^*(1)+\pi^{\text{dev}}\right)+\frac{\delta}{1-\delta}3\pi^*(2),$$

implying,

$$\delta \geq \frac{\pi^{\text{dev}}-\pi^{\text{coll}}}{\pi^*(1)+\pi^{\text{dev}}-3\pi^*(2)} < \delta^{IM}.$$

It follows that by combining the two collusive mechanism both cartel members receive instantaneous profits of $\pi^*(1)+\pi^{\text{coll}} > \pi^*(1)+\pi^*(2)$ each period from an agreement that is more stable than colluding at the intensive margin of all markets. ■

Finally, it is useful to emphasize that collusion at the extensive margin (as we have modelled it here using Markov perfect equilibrium) involves a different speed of reaction to a deviation than does collusion at the intensive margin (as it is usually modelled). The reason is that intensive margin collusion is coordinating on behavior while extensive margin collusion coordinates on participation. Therefore, a deviation from an intensive margin collusion equilibrium allows the deviator to earn instantaneous profits holding the behavior of rivals as fixed something that is not the case with intensive margin collusion where deviation profits merely hold participation (and not behavior) of rivals as fixed. To see this distinction, we provide a comparison using the international trade model of Bond and Syropoulos (2008).

Example 3. There are two firms and two identical markets. Firm 1 (resp. 2) has its home in market 1 (resp. 2). Let q denote a firm's home sales and x denote its exported sales. Price in a market is determined by $1-(q+x)$. Transporting goods between markets costs $t(<\frac{1}{2})$ per unit and there are no other production costs. Market entry costs are infinitesimally small. Industry profits are maximized if each firm has a monopoly in their respective home markets earning $\pi^*(1)=\pi^{\text{coll}}=1/4$. Bond and Syropoulos (2008) assume that, when they compete, firms are Cournot competitors (i.e., they can commit to quantities). Thus, $\pi^*_{1,1}(2)=\pi^*_{2,2}(2)=\frac{1}{9}(1+t)^2$ while $\pi^*_{1,2}(2)=\pi^*_{2,1}(2)=\frac{1}{9}(1-2t)^2$. Finally, when a firm deviates and enters its rival's market, its rival keeps its behavior constant under intensive margin collusion at the monopoly output. Playing a best response to this, earns the rival $\pi^{\text{dev}}_{1,2}(2)=\pi^{\text{dev}}_{2,1}(2)=\frac{1}{16}(1-2t)^2$.

For firm 1, the no deviation constraint for intensive margin collusion is,

$$\frac{1}{1-\delta}\pi^*(1)\geq\pi^*(1)+\pi^{\text{dev}}_{1,2}(2)+\frac{\delta}{1-\delta}(\pi^*_{1,1}(2)+\pi^*_{1,2}(2)),$$

while the no deviation constraint for extensive margin collusion is,

$$\frac{1}{1-\delta}\pi^*(1)\geq\pi^*(1)+\pi^*_{1,2}(2)+\frac{\delta}{1-\delta}(\pi^*_{1,1}(2)+\pi^*_{1,2}(2)).$$

Comparing these two expressions, it is clear that $\delta^{IM}<\delta^{GS}$ if and only if $\pi^{\text{dev}}_{1,2}(2)<\pi^*_{1,2}(2)$, which it is for the Cournot case. In a stronger, within market, competitive environment, it is possible that $\pi^{\text{dev}}_{1,2}(2)>\pi^*_{1,2}(2)$ making extensive margin collusion more stable than intensive margin collusion. It is instructive to note that both critical discount factors are decreasing in t. Thus, Bond and Syropoulos' main result hold regardless of the type of collusion analyzed.[14] ■

This example demonstrates that observed market separation can occur under intensive margin collusion as it necessarily does under extensive margin collusion. The reaction of rivals immediately upon deviation is what distinguishes them in this context. Specifically, Bond and Syropoulos have a trade model in mind that involves the imported goods appearing (say, with the speed and surprise of a Star Trek transporter) with rivals being unable to adjust their behavior. In contrast, here importation takes place via a 'slow boat' entry process whereby deviators expect to be greeted with equilibrium competitive behavior in rival markets but will, like intensive margin

[14] This comparison only considers the case where intensive margin collusion results in no cross hauling of goods between markets. Bond and Syropoulos (2008) demonstrate that when discount factors are low, such cross-hauling can support a more profitable cartel outcome. In this case, the comparative static on transportation costs can change.

collusion, have a period's grace before any responding competitive behavior in their home market. Depending upon the facts of international trade or operation across markets, each type of assumption may suit different contexts.

4 Targeted Enforcement and Uncertainty

The nature of collusion at the extensive margin creates the potential for firms to employ enforcement mechanisms that are temporary, targeted and scale with the size of a deviation. Firms may prefer to employ temporary punishment strategies where either uncertainty triggers punishments along the equilibrium path, or punishments are costly rendering the threat of permanent punishments not credible. The latter possibility is considered in section 5.

We begin by developing collusive equilibria for three tit-for-tat enforcement mechanisms: Untargeted response enforcement in which any deviation is punished by a game wide reversion to the competitive equilibrium; targeted response enforcement in which punishments are targeted at the offending firm; and proportional response enforcement in which punishments are both targeted and scaled. Under perfect information untargeted response enforcement is shown to support the most stable collusive agreements due to a *scorched earth* effect, while target and proportional response enforcement have identical participation constraints.

The picture becomes more complex once uncertainty is introduced. Loosely following Green and Porter (1984), we introduce the possibility that firms make errors, triggering punishments along the equilibrium path. Proportional response enforcement is shown to dominate target response enforcement both in terms of expected profits and the range of parameter values over which collusion is supported. This result is significant as — to the best of our knowledge — this is the first paper to provide a game theoretic justification for the use of proportional response in self-enforcing contracts.

One artefact of the MPE refinement is that the tit-for-tat punishments developed in this section last for a single period. This is an entirely artificial restriction which, nevertheless, allows us to isolate two key features of temporary punishments in our framework. First, we show that where collusion takes place at the extensive margin punishments of one period length may be sufficient to deter deviations. Second, by limiting the length of punishments we confine our focus to the scale and scope of punishment strategies within any given period. In the larger class of sub-game perfect equilibria all punishment strategies developed here can be enhanced by increasing the length of the punishment phase.

For the purposes of this section it is useful to strengthen assumption 2 to ensure the expansion incentive is strong enough to promote temporary entry into a market.

Assumption 3 (Short-Run Expansion Incentive): Consider a set of separable, identical and symmetric markets. For all $q \in \{1, 2, \ldots\}$ MPE instantaneous profits satisfy $\pi^*(q) > \pi^*(q+1) > c$.

4.1 Untargeted Response Enforcement

In *untargeted response enforcement* all firms respond to a transgression in period t by entering every market in period $t+1$. The punishment phase concludes once every firm is present in every market in the game. The punishment phase of untargeted response enforcement is the analogue of a price war in intensive margin collusion; once a deviation is observed the collusive agreement collapses for a single period.

Proposition 4 (Untargeted Response Enforcement): *Consider a separable, identical and symmetric multi-market game satisfying assumptions 1 and 3. Moreover, consider the partition* $P = \{N_\varnothing, \{N_i\}_{i \in I}\}$ *and the participation stage strategy* $a^{UR}(\cdot)$ *such that* $a_i^{UR}(a^{t-1}) = N$ *if there exists* $k \neq l$ *such that* $a_k^{t-1} \cap N_l \neq \varnothing$, *and* $q \neq r$ *such that* $N_q \nsubseteq a_r^{t-1}$; *and* $a_i^{UR}(a^{t-1}) = N_i \cup N_\varnothing$ *otherwise. The collusive agreement* (P, a^{UR}) *defines an MPE if and only if* δ *satisfies,*

$$\delta \geq \delta^{UR} = \max_{i \in I} \left[\frac{\sum_{j \neq i} n_j \pi^*(2)}{n_i \pi^*(1) - \sum_{j \in I} n_j \pi^*(I)} \right]. \tag{5}$$

Proof. Beginning with the case in which $a^{t-1} = a^P$, firm i may deviate by taking the action $a_i^t = N_i \cup N_\varnothing \cup Q$ where $\varnothing \neq Q \subseteq \bigcup_{j \neq i} N_j$. This action triggers a total collapse of the collusive agreement in period $t+1$ thus $a^{t+1} = N^I$. Given that $N_q \subset a_r^{t+1} = N$ for all $q \neq r$ in period $t+1$, the punishment phase is concluded and firms withdraw back to the collusive profile of participation in period $t+2$. That is to say, for $\tau \geq t+2$ all firms collectively revert to the strategy profile $a^\tau = a^P$.

The punishment that a firm receives is insensitive to the number of markets that it enters in a deviation. Therefore, the worst case deviation is where a firm deviates by entering the set of markets $Q = \bigcup_{j \neq i} N_j$ in period t. This deviation does not improve firm i's profit if,

$$\delta n_i \big(\pi^*(1) - \pi^*(I)\big) \geq \sum_{j \neq i} n_j \big(\pi^*(2) - c + \delta \pi^*(I)\big), \tag{6}$$

where the RHS represents the gains to firm i from participating in every market in $\bigcup_{j \neq i} N_j$ as a duopolist in period t, and as an I-opolist in period $t+2$, less the cost of entry; while the LHS represents the instantaneous profits that are lost as a result of all remaining firms establishing a presence in all markets in N_i in period $t+1$. It is straight forward to see that once a transgression has occurred all firms should respond by entering every market maximising the instantaneous profits of firms in the current period and minimising the length of the punishment phase. Failure by firm i to withdraw from all markets in $\bigcup_{j\neq i} N_j$ where $N_q \subseteq a_r^{t-1}$ for all $q \neq r$ triggers a new punishment phase in the same way as entry, however the deviating firm does not incur entry costs as it is already present in all markets. It follows that firm i will not deviate where the state is $a^{t-1} = N^I$ if,

$$\delta n_i\big(\pi^*(1) - \pi^*(I)\big) \geq \sum_{j\neq i} n_j\big(\pi^*(2) + \delta\pi^*(I)\big), \tag{7}$$

which is (6) with the entry cost removed. Given that $c > 0$ it is condition (7) that is critical for determining cartel stability. Rearranging (7) yields (5). □

Once a deviation triggers punishment, the dominant strategy is for all firms to enter all of their rivals' markets, thereby both maximizing instantaneous profits and minimizing the length of the punishment phase. Notice that even though one firm unilaterally instigates the punishment phase, all firms back down in the same period. This multilateral withdrawal creates an advantage for the firm who instigated the initial deviation as its actions are unchallenged for a period.

Entry costs are irrelevant for cartel stability in untargeted response enforcement as the worst case deviation occurs where a firm fails to withdraw from its rivals' markets at the conclusion of tit-for-tat punishment. This feature is shared by both targeted and proportional response enforcement.

4.2 Targeted Response Enforcement

In contrast to collusion at the intensive margin, the nature of collusion at the extensive margin permits firms to target punishments such that they only impact upon the offending firm. Under *targeted response enforcement* punishments are confined to the firm that instigated the transgression and are carried out exclusively by those firms who suffered from the transgression. This reduces each transgression within the game to a bilateral disagreement.

The enforcement mechanism employs a disproportionate response insofar as once the punishment phase begins both the transgressor and the aggrieved firm enter all of each other's markets. Bilateral withdrawal is instigated once both firms are present in all markets in one and other's components of the partition. Of course a deviation can target multiple rival firms simultaneously. In this case, the deviating firm enters into bilateral punishments with every target firm simultaneously.

Proposition 5 (Targeted Response Enforcement): *Consider a separable, identical and symmetric multi-market game satisfying assumptions 1 and 3. Moreover, consider the partition* $P = \big\{N_\varnothing, \{N_i\}_{i\in I}\big\}$ *and the participation stage strategy* $a^{TR}(\cdot)$ *such that,*

$$a_i^{TR}(a^{t-1}) = N_i \cup N_\varnothing \cup \bigcup_{j\in J_i(a^{t-1})} N_j,$$

where the (possibly empty) set,

$$J_i(a^{t-1}) = \{j \in I \setminus \{i\} : a_j^{t-i} \cap N_i \neq \varnothing \text{ and } N_j \nsubseteq a_i^{t-1}; \quad \text{or } a_i^{t-1} \cap N_j \neq \varnothing \text{ and } N_i \nsubseteq a_j^{t-1}\}.$$

The collusive agreement (P, a^{TR}) *defines an MPE if and only if* δ *satisfies,*

$$\delta \geq \delta^{TR} = \max_{i\in I}\left[\max_{K\subseteq I\setminus\{i\}}\left(\frac{\sum_{j\in K} n_j \pi^*(2)}{n_i\big(\pi^*(1) - \pi^*(\|K\|+1)\big) - \sum_{j\in K} n_j\pi^*(2)}\right)\right]. \tag{8}$$

Proof. This proof follows the proof of proposition 4. Beginning with the case in which $a^{t-1} \in a^P$, firm i may deviate by taking the action $a_i^t = N_i \cup N_\varnothing \cup Q$ where Q is a non-empty subset of $\bigcup_{j\in I\setminus\{i\}} N_j$. This action triggers targeted punishments in period $t+1$. Note that firm j is in the set $J_i(a^t)$ if and only if $Q \cap N_j \neq \varnothing$, while for all $j \in J_i(a^t)$ we have $J_j(a^t) = \{i\}$ thus firm i's punishment phase action is $a_i^{t+1} = N_i \cup N_\varnothing \cup \big(\bigcup_{j\in J_i(a^t)} N_j\big)$ while all firms $j \in J_i^t$ take the action $a_j^{t+1} = N_j \cup N_\varnothing \cup N_i$. Given that $N_i \subset a_j^{t+1}$ and $N_j \subset a_j^{t+1}$ for all $j \in J_i(a^t)$, the targeted punishment is concluded and the firms withdrawal from their rivals' markets in period $t+2$. That is to say, for $\tau \geq t+2$ all firms collectively revert to the strategy profile $a^\tau = a^P$. Throughout the punishment phase all firms $k \in I \setminus \big(\{i\} \cup J_i(a^t)\big)$ play the strategy $a_k^\tau = a_k^P = N_k \cup N_\varnothing$ and do not experience any change in profits as a result of the targeted punishments.

The magnitude of the punishment firm i experiences as a result of the deviation is sensitive to the number of rival firms targeted by the deviation, but not to the total number of markets entered. Therefore, the worst case deviation is where a firm deviates by entering all markets belonging to a subset of rival firms $K \in I \setminus \{i\}$ in period t. This deviation does not improve firm i's profit if,

$$\delta n_i\big(\pi^*(1) - \pi^*(\|K\|+1)\big) \geq (1+\delta)\sum_{k\in K} n_k\big(\pi^*(2) - c\big), \tag{9}$$

where the RHS represents the gains to firm i from participating in every market in $\bigcup_{k\in K} N_k$ as a duopolist in periods t and $t+1$; while the LHS represents the instantaneous profits that are lost as a result of all firms $k \in K$ establishing a presence in all markets $n \in N_i$ in period $t+2$.

It is straight forward to see that once a transgression has occurred every firm k with $J_k(a^t) \neq \varnothing$ should respond by entering every market in $\bigcup_{l\in J_k(a^t)} N_l$. Failing to do so extends the length of the punishment phase and reduces the firm's instantaneous profit in period $t+1$.

Entering more markets than is dictated by the enforcement mechanism in period $t+1$ cannot be profitable where (9) holds. To see this note the return to a firm k from selecting an action $a_k^{t+1} = a_t^{TR}(a^t) \cup Y$ is weakly less than the return to taking an identical action where the state is a^P.

Failure by a firm i to withdraw from all markets in $\bigcup_{j\in J_i(a^t)} N_j$ where $N_i \subseteq a_j^{t+1}$ and $N_j \subseteq a_i^{t+1}$ triggers a new punishment phase in the same way as entry, however the deviating firm does not incur entry costs as it is already present in all markets. It follows that firm i will not deviate in period $t+2$ if,

$$\delta n_i\big(\pi^*(1) - \pi^*(\|K\|+1)\big) \geq (1+\delta)\sum_{k\in K} n_k \pi^*(2), \tag{10}$$

which is (9) with the entry cost removed. Rearranging (10) yields (8). □

Comparing (5) and (8) it is clear that untargeted response enforcement supports more stable collusive agreements than targeted response enforcement. However, the two enforcement mechanisms are equivalent in a two-firm game.

The reason that a discrepancy may arise where three or more firms exist can be seen when comparing (6) and (9). Under targeted response enforcement a deviating firm receives duopoly profits from the markets it enters in both the period of the initial deviation and the subsequent period when punishments are implemented. Conversely, when enforcement is untargeted every firm enters every market reducing the instantaneous profits of every market in the game to $\pi^*(I)$ in the period following the deviation.

Intuitively, it is valuable for two firms to punish each other, even where neither firm was involved in the initial transgression, because in doing so they reduce the instantaneous profit of every market to the lowest level possible in an MPE. In turn, this *scorched earth* effect enhances the stability of a cartel as it reduces the payoff to any initial deviation.

4.3 Proportional Response Enforcement

The final form of tit-for-tat enforcement we consider is *proportional response enforcement* in which punishments are both targeted and scaled to match the size of the initial transgression. In proportional response enforcement firm j responds to entry by firm i, into a subset of markets in N_j, by entering an equal number of markets in N_i. However, if the number of markets entered is at least equal to the size of one firm's partition then firms respond as per the targeted response enforcement mechanism by entering every market belonging to the rival firm. Once both firms are entering or present in an equal number of markets that is strictly less than $\min\{n_i, n_j\}$; or are entering or present in every market; both firms simultaneously withdraw from all markets in their rival's component of the partition. As in targeted response enforcement, punishments are bilateral in nature. If multiple firms suffer as a consequence of a deviation, proportional response requires each aggrieved firm to employ a punishment proportional to its own loss.

The following proposition shows that under perfect information the targeted response enforcement mechanism supports precisely the same same set of partitions as targeted response enforcement. We show below that this parity does not extend to an environments with uncertainty.

Proposition 6 (Proportional Response Enforcement): *Consider a separable, identical and symmetric multi-market game satisfying assumptions 1 and 3. Moreover, consider the partition $P = \big\{N_\varnothing, \{N_i\}_{i\in I}\big\}$ and the participation stage strategy $a^{PR}(\cdot)$ such that,*

$$a_i^{PR}(a^{t-1}) = N_i \cup N_\varnothing \cup \Big(\bigcup_{j\in \bar{J}_i(a^{t-1})} (a^{t-1} \cap N_j) \cup Q_{i,j} \Big) \cup \bigcup_{j \in J_i(a^{t-1})} N_j,$$

where the (possibly empty) sets $\bar{J}_i(a^{t-1})$ and $J_i(a^{t-1})$ are defined,

$$\bar{J}_i(a^{t-1}) = \Big\{ j \in I \setminus \{i\} : 0 < \max\big\{ \|a_i^{t-1} \cap N_j\|, \|a_j^{t-1} \cap N_i\| \big\} < \min\{n_1, n_2\} \\ \text{and } \|a_i^{t-1} \cap N_j\| \neq \|a_j^{t-1} \cap N_i\| \Big\},$$

and,

$$J_i(a^{t-1}) = \big\{ j \in I \setminus \big(\bar{J}_i(a^{t-1}) \cup \{i\}\big) : a_j^{t-1} \cap N_i \neq \varnothing \text{ and } N_j \not\subseteq a_i^{t-1}; \\ \text{or } a_i^{t-1} \cap N_j \neq \varnothing \text{ and } N_i \not\subseteq a_j^{t-1} \big\},$$

and the set $Q_{i,j}$ is defined to be any (possibly empty) subset of $N_j \setminus a_i^{t-1}$ that contains a number of markets,

$$\|Q_{i,j}\| = \max\big\{0, \|a_j^{t-1} \cap N_i\| - \|a_i^{t-1} \cap N_j\|\big\}.$$

The collusive agreement (P, a^{PR}) *defines an MPE if and only if* $\delta \geq \delta^{TR}$ *as defined in* (8).

Proof. Beginning with the case in which $a^{t-1} = a^P$, once again the worst case deviation is where the firm with the smallest partition deviates by entering all markets belonging to a subset of rival firms $K \in I \setminus \{i\}$ in period t. From the proof proposition 5 this deviation is not profitable if (9) holds for all $K \subseteq I \setminus \{i\}$.

As in proposition 5, firms have no incentive to punish a transgression by entering fewer markets than the mechanism dictates. Firms i and j do not instigate a bilateral withdrawal until either $\|a_i^{t-1} \cap N_j\| = \|a_j^{t-1} \cap N_i\| < \min\{n_i, n_j\}$, or $N_i \subset a_j^{t-1}$ and $N_j \subset a_i^{t-1}$. Thus failure to enter the required number of markets both reduces a firm's instantaneous profit and prolongs the punishment phase.

It follows from the proof of proposition 5 that failing to withdraw from markets as required by the enforcement mechanism, or entering more markets than is dictated by the enforcement mechanism, cannot be profitable where (10) holds. □

Targeted and proportion response enforcement perform identically under perfect information due to the fact that in both cases the worst case deviation is for the firm with the smallest partition to enter all markets belonging to a subset of rival firms. The contrast between the three tit-for-tat equilibria emerges where uncertainty is introduced into the model.

4.4 Uncertainty

Introducing uncertainty into the framework provides a basis for comparing the performance of otherwise equivalent forms of tit-for-tat collusion. For the purposes of this section we assume that the source of uncertainty in the multi-market setting is the possibility that a firm makes an error in the participation stage, entering more markets than the firm intended. Specifically, with some probability the action a_i^t chosen by firm i is instead implemented as $\hat{a}_i^t \supset a_i^t$. Intuitively, such an error might occur if an overzealous manager — unaware of the existence of the cartel — oversteps their authority and initiates entry into a market without seeking permission from their superiors.

Define $D_K = \times_{k \in K}\{1, \ldots, n_k\}$ and let $d_K = \{d_k\}_{k \in K} \in D_K$ represent a profile of accidental entry such that $d_k = \|\hat{a}_i^t \cap N_k\|$ is the number of markets belonging to firm k that firm i enters. The probability that firm i erroneously enters a profile of markets d_K belonging to a set of firms K is written $\sigma(i, K, d_K)$ while the probability that no error occurs is $\sigma(\varnothing)$.

Assumption 4: (a) An error only occurs where firms collectively play the participation stage action profile $a^t = a^P$; (b) Every possible error occurs with strictly positive probability where the state satisfies (a) hence $\sigma(i, K, d_K) > 0$ for all $i \in I$, $K \subseteq I \backslash \{i\}$ and $d_K \in D_K$; (c) At most one error occurs in each period hence,

$$\sigma(\varnothing) + \sum_{\substack{i \in I \\ K \subseteq I \backslash \{i\} \\ d_K \in D_K}} \sigma(i, K, d_K) = 1;$$

(d) When an error occurs it is observed simultaneously by all firms including the firm that makes the error. However, only the firm that makes the error is aware that its observed action $\hat{a}_i^t$ is not equal to the action a_i^t selected by the firm.

An error has the effect of triggering punishments along the equilibrium path. In common with Green and Porter (1984) the error always implies overly aggressive play by a firm and only impacts the game in a period in which all firms have behaved collusively ($a_i^t = a^P$). Unlike Green and Porter (1984) the observed deviation is real, however it occurs despite the firm's intent to maintain the collusive equilibrium.

The probability that firm i is neither the instigator, nor the target, of a deviation is,

$$\sigma_i = \sigma(\varnothing) + \sum_{\substack{j \neq i \\ K \subseteq I \backslash \{i,j\} \\ d_K \in D_K}} \sigma\big(j, K, d_K\big).$$

The following proposition demonstrates that proportional response enforcement dominates targeted response enforcement both in terms of expected payoffs and the stability of the collusive agreement.

Proposition 7: *Consider a separable, identical and symmetric multi-market game satisfying assumptions 1, 3 and 4. Moreover, consider the partition $P = \big\{N_\varnothing, \{N_i\}_{i \in I}\big\}$ and suppose that the pair (P, a^{TR}) defines a collusive equilibrium under uncertainty for the probability function $\sigma(\cdot)$ and discount factor $\delta \in (0, 1)$. It follows that:*

1. *The pair (P, a^{PR}) defines a collusive equilibrium under uncertainty for the probability function $\sigma(\cdot)$ and discount factor δ;*
2. *The strategy profile a^{PR} delivers firms (weakly) higher expected profits than a^{TR}, with strict inequality for any pair of firms $\{i, j\}$ such that $n_i \geq 2$ and $n_j \geq 2$;*
3. *The pair (P, a^{PR}) defines a collusive equilibrium for a weakly wider range of probability functions $\sigma(\cdot)$ and discount factors $\delta \in (0, 1)$ than the pair (P, a^{TR}).*

Proof. We begin by characterising the continuation values under both forms of collusion where the state satisfies $a^{t-1} = a^P$. The continuation value to firm i of selecting the action $a_i^t = a^{XX}(a^P)$ is,

$$V_i^{XX} = n_\varnothing \pi^*(I) + \sigma_i\big(n_i\pi^*(1) + \delta V_i^{XX}\big) + W_i \\ + \delta Z_i^{XX} + \delta(1-\sigma_i)\big(n_\varnothing \pi^*(I) + \delta V_i^{XX}\big),$$

for $XX \in \{TR, PR\}$ where,

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

Solving for V_i^{XX} yields,

$$V_i^{XX} = \frac{1}{1-\delta}\left(\frac{\sigma_i n_i \pi^*(1) + W_i + \delta Z_i^{XX}}{1+\delta-\delta\sigma_i} + n_I\pi^*(I)\right). \tag{11}$$

The term Z_i^{XX} represents the probability weighted profits that a firm receives in period $t+1$ when a punishment phase is triggered by an error in period t. Under targeted response enforcement,

$$Z_i^{TR} = \sum_{\substack{K\subseteq I\setminus\{i\} \\ d_K \in D_K}} \sigma(i, K, d_K)\Big(n_i\pi^*(\|K\|+1) + \sum_{k\in K}\big(n_k\pi^*(2) - (n_k - d_k)c\big)\Big) \\ + \sum_{\substack{j\neq i \\ i\ni K\subseteq I\setminus\{j\} \\ d_K\in D_K}} \sigma(j, K, d_K)\big(n_i\pi^*(2) + n_j\pi^*(\|K\|+1) - n_j c\big) \leq Z_i^{PR},$$

with strict inequality if $n_i \geq 2$ and there exists $j \neq i$ such that $n_j \geq 2$. Intuitively, from assumption 4 it follows that firm i will err entering exactly one of firm j's markets with strictly positive probability. Under proportional response enforcement firm j responds by entering exactly one market in N_i inflicting a lighter punishment on i than would be the case under targeted response enforcement. It follows from (11) that $Z_i^{PR} \geq Z_i^{TR}$ implies $V_i^{PR} \geq V_i^{TR}$ proving 2.

The highest return to a once off deviation is,

$$V_i^{XX-} = \max_{K\subseteq I\setminus\{i\}}\Big[n_i\big(\pi^*(1) + \delta\pi^*(K+1)\big) + \sum_{j\in K} n_j\big((1+\delta)\pi^*(2) - c\big)\Big] \\ + (1+\delta)n_\varnothing\pi(I) + \delta^2 V_i^{XX},$$

for all $XX \in \{TR, PR\}$. The collusive agreement is stable if $V_i^{XX} - V_i^{XX-} \geq 0$ for all $i \in I$. This difference can be written,

$$V_i^{XX} - V_i^{XX-} = (1+\delta)\frac{\sigma_i n_i \pi^*(1) + W_i + \delta Z_i^{XX}}{1+\delta-\delta\sigma_i} \\ - \max_{K \subseteq I \setminus \{i\}} \Big[n_i\big(\pi^*(1) + \delta\pi^*(K+1)\big) + \sum_{j \in K} n_j\big((1+\delta)\pi^*(2) - c\big)\Big].$$

The difference $V_i^{XX} - V_i^{XX-}$ is continuous and for a given probability function $\sigma(\cdot)$ increasing in Z_i^{XX} proving 1.

Now suppose that $\sigma(\varnothing)$ is increased by reducing every $\sigma(i, K, d_K)$ proportionately. Increasing $\sigma(\varnothing)$ increases σ_i and reduces the weight of the Z_i^{XX} term in the weighted average,

$$\frac{\sigma_i n_i \pi^*(1) + W_i + \delta Z_i^{XX}}{1+\delta-\delta\sigma_i},$$

it follows that the difference $V_i^{XX} - V^{XX-}$ is increasing in both $\delta \in (0,1)$ and $\sigma_i \in (0,1)$ proving 3. □

Proposition 7 provides two compelling reasons why a cartel would prefer proportional response enforcement over targeted response enforcement. Namely, it delivers higher expected returns to the cartel for any given uncertainty profile, as well as supporting equilibria over a larger range of discount factors and uncertainty functions.

While the expected returns to employing targeted and proportional response enforcement obey proposition 7, the relationship between these profits and the expected returns to untargeted response enforcement are ambiguous. The ambiguity arises because while punishments affect each firm more often where punishments are untargeted[15] under some specifications the profits earned during an untargeted punishment may dominate the profits that a firm earns as a participant in either a targeted or a proportional response punishment. Of course an untargeted response is equivalent to a targeted response in a two-firm game.

5 Implications for Antitrust Policy

The framework developed in this paper has several implications for antitrust policy. The model admits forms of market sharing than would not usually be predicted by models of collusion including *oligopoly competition with a collusive fringe* and collusion enforced by the threat of *predatory entry*. The nature of collusion at the

[15]A firm is involved in a punishment phase with probability $1 - \sigma(\varnothing)$ under untargeted response enforcement and probability $1 - \sigma_i$ under both targeted and proportional response enforcement.

extensive margin may also make cartel detection more difficult. Moreover, a policy of punishing reversion to a collusive partition following tit-for-tat punishment may have the effect of increasing the stability of a collusive agreement.

5.1 Oligopolistic Competition with a Collusive Fringe

The presence of an asymmetrically valuable market may act as a barrier to forming a stable collusive partition. Consider the case of a separable and symmetric multi-market game with a set of markets $N' = N \cup \{L\}$. Suppose that the markets in the set N are identical and that the monopoly and I-opoly profits, and entry cost for market L satisfy,

$$\pi_L^*(1) > \pi_L^*(I) - (1-\delta)c_L \geq \|N\| \left(\frac{\pi^*(1) + (1-\delta)(\|I\| - 2)c}{\|I\| - 1} - \pi^*(I) \right) > 0. \quad (12)$$

We term L a *large market* and note that the magnitude of the duopoly profit is so large that a partition P cannot be supported in a collusive equilibrium if $L \in N_i$ for any $i \in I$.

Nevertheless, the presence of a large market in a multi-market game need not prevent a stable collusive outcome. To the contrary, as the following proposition demonstrates, adding a large market to a game has no affect on the range of collusive equilibria which may arise.

Proposition 8: *Consider a separable and symmetric multi-market game satisfying assumptions 1 and 2. Let $N' = N \cup \{L\}$ represent the set of markets in the game and suppose that the markets in N are identical while the MPE instantaneous profits and entry costs of the large market L satisfy (12). The pair (P'^*) defines a collusive equilibrium for the game (N', I) and discount factor $\delta \in (0, 1)$ if and only if the pair (P, a^*) defines a collusive equilibrium for the game (N, I) where $N_i' = N_i$ for all $i \in I$ and $N_\varnothing' = N_\varnothing \cup \{L\}$.*

Proof. From the definition of a collusive equilibrium $N_\varnothing \subseteq a_i^*(a^t)$ for all $i \in \{1, 2\}$ and $a^t \in 2^{N^I}$. Moreover, given the expansion incentive neither firm has an incentive to exit a market in $N_\varnothing$ either on or off the equilibrium path. It follows that markets in $N_\varnothing$ play at most a trivial role in the participation constraints of any collusive equilibrium, and therefore the composition of $N_\varnothing$ does not affect the existence or stability of a collusive equilibrium so long as the composition of the components $\{N_i\}_{i \in I}$ remain unchanged. □

Intuitively, proposition 8 holds because the large market can always be assigned to the the contested component of a collusive partition $N_\varnothing$. The remaining markets in N can then be divided between the monopolised components of the collusive partition in manner consistent with the participation constraints of the relevant collusive agreement. Because all firms maintain a presence in all markets in $N_{\varnothing}$ regardless of the prevailing state of the world, these markets produce the same MPE instantaneous profits both on and off the equilibrium path. A corollary of proposition 8 is that in any separable multi-market game the markets which constitute the component $N_{\varnothing}$ have no impact on the stability of a collusive partition.

One consequence of proposition 8 is that we cannot generally 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* in which firms compete fiercely in the large market while at the same time dividing up the small markets in a collusive partition.

There are a number of market structures that may display a collusive fringe. Consider, for example, the market for beer or sodas. All major firms in these markets tend to be in direct competition with one and other, selling their products through supermarkets and grocery stores. At the same time these same firms sign exclusive deals with restaurant chains, convenience stores, 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.

In each of these cases, the defining feature of the large market is that it is very profitable relative to the smaller peripheral markets, and that it cannot be effectively segmented into separable smaller markets. In contrast, the smaller peripheral markets can be partitioned between two or more firms. 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 in accordance with the long-run expansion incentive (assumption 2), the partitioning of these markets is not consistent with competitive behaviour and we can conclude that we are observing collusion at the extensive margin.

5.2 Predatory Entry

Throughout this paper firm behaviour has been driven by the expansion incentive (assumptions 2 and 3). The expansion incentive plays a critical role in our framework as it provides firms with both an incentive to deviate and the incentive to implement punishments. But what happens when entry can result in a market yielding negative MPE instantaneous profits to participating firms?

Here we consider the role that predatory entry may play in sustaining collusion at the extensive margin. We define *predatory entry* to be entry by a firm into a market with the purpose of reducing the instantaneous 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. The following example illustrates the concept.

Example 4. Consider a two-firm separable and symmetric multi-market game in which there are two markets $N = \{m, d\}$. Market m is a natural monopoly market which produces MPE instantaneous profits $\pi^*_m(1) > 0 > \pi^*_m(2)$ and has an associated entry cost c_m, while the market d is a natural duopoly with $\pi^*_d(1) > \pi^*_d(2) > c_d$. The purpose of the example is to identify the conditions under which the partition $P = \{N_1, N_2\}$ with $N_1 = \{m\}$ and $N_2 = \{d\}$ can be sustained as a collusive equilibria.

Under perfect information grim-strategy collusion produces the most robust cartel where assumption 2 holds. Conversely, in this example the presence of the natural monopoly market renders grim strategy collusion ineffective. To see this consider a deviation from the collusive agreement in which firm 1 enters market d. The grim-strategy requires firm 2 to respond by entering and remaining in market m indefinitely. But this response is not sub-game perfect as $\pi^*_m(2) < 0$ and therefore once the punishment begins either firm can increase its payoff by withdrawing from the natural monopoly market. It follows that the threat of grim-strategy punishment is not credible and therefore firm 1 can enter market d without threat of reprisal.

The presence of the natural monopoly market introduces asymmetric incentives into the multi-market game. Firm 1 has an incentive to enter market d in order to attain duopoly profits from that market. In contrast firm 2 has no interest in entering market m as doing so forces the MPE profits from 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 collusive partition.

All three of the tit-for-tat enforcement mechanism developed in section 4 are equivalent in a two-firm, two-market game. From (5) it follows that the threat of tit-for-tat punishment is sufficient to deter firm 1 from entering market d so long as,

$$\delta \geq \frac{\pi^*_d(2)}{\pi^*_m(1) - \pi^*_m(2) - \pi^*_d(2)}. \tag{13}$$

Here the fact that $\pi_m^*(2) < 0$ enhances cartel stability as it increases the cost of the punishment that follows entry. Moreover, this condition ensures that firm 1 will withdraw from market d following the single period of a punishment phase as failing to do so triggers an entirely new punishment phase in the subsequent period with the same payoffs as the initial deviation.

We do not have to establish an equivalent condition for firm 2. Firm 2 has no incentive to initiate a deviation along the equilibrium path as the return to entering m is negative. However, it is necessary to verify that firm 2 will be willing to carry out the punishment in the event that firm 1 deviates.

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 firm 2 will be willing to employ tit-for-tat punishments if and only if,

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

implying,

$$\delta \geq \frac{-\pi_m^*(2) + c_m}{\pi_d^*(1) - \pi_d^*(2) - \pi_m^*(2) + c_m}. \tag{14}$$

Given the assumptions on the MPE profits of the two markets, firm 2 will be willing to implement the tit-for-tat punishment strategies if it is sufficiently patient. ■

5.3 Concentration of Ownership and Mergers in a Cartel

Where collusion takes place at the intensive margin increasing the concentration of ownership within the cartel tends to increase the stability of the collusive agreement. For example consider grim-strategy intensive margin collusion between I identical firms in a Bertrand market. If the monopoly profit in the market is π^m, each firm receives an instantaneous profit of π^m/I in each period the collusive agreement holds. A deviation nets a firm the full monopoly profit in the period of the deviation, and a return of zero in all subsequent periods as the game reverts to the competitive equilibrium. The critical discount factor for this example is $\delta^{IM} = (I-1)/I$ which is unambiguously increasing in I.

The situation is more complex where collusion takes place at the extensive margin. Consolidation of ownership within a cartel must eliminate a component of the collusive partition and the way in which the markets in this component are redistributed has implications for cartel stability. For the purposes of this paper we distinguish between mergers that combine two components of a partition and a change in the concentration of ownership within the cartel that maintains the relative market shares of the participating firms. In each case the change of participation can lead to either an increase or decrease in the stability of the cartel so long as at least two firms are present following the consolidation.

Consider a merger between two members of a cartel. We assume that the merger has the effect of reducing the number of members of a cartel by one as well as combining the two components of the pre-merger collusive partition belonging to the merging firms. The following proposition shows that if at least one of the firms with the smallest component of the pre-merger partition is not involved in the merger then the merger reduces the stability of the cartel.

PROPOSITION 9: *Consider an $\|I\| \geq 3$ firm, separable, identical and symmetric multimarket game satisfying assumptions 1 and 3, and a collusive agreement (P, a^{GS}) where $P = \{N_\varnothing, \{N_i\}_{i\in I}\}$. Let $k \in I$ be the (possibly unique) argument that maximises δ^{GS} as defined in (3). Suppose that two firms $j, l \neq k$ merge giving rise to a collusive partition P' such that $N'_i = N_i$ for all $i \neq \{j, l\}$ and $N'_{\{j,l\}} = N_j \cup N_l$. The merger strictly increases the value of δ^{GS}.*

Proof. Let $n'_i = \|N'_i\|$. From assumption 3 it follows that $\pi^*(I-1) > \pi^*(I)$. For all firms not participating in the merger including firm k,

$$\frac{\sum_{j\neq i} n'_j\big(\pi^*(2) - c\big)}{n'_i\big(\pi^*(1) - \pi^*(I-1)\big) + \sum_{j\neq i} n'_j\big(\pi^*(2) - \pi^*(I-1) - c\big)} > \frac{\sum_{j\neq i} n_j\big(\pi^*(2) - c\big)}{n_i\big(\pi^*(1) - \pi^*(I)\big) + \sum_{j\neq i} n_j\big(\pi^*(2) - \pi^*(I) - c\big)},$$

and therefore δ^{GS} must increase as a consequence of the merger. □

For a firm i excluded from the merger, the union of two firms has no effect on either MPE instantaneous profits in the collusive equilibrium ($N_i = N'_i$) or the incentive to initiate a deviation ($\sum_{j\neq i} n_j = \sum_{j\neq i} n'_j$). The effect of the merger on an excluded firm i only becomes apparent once the grim-strategy punishment is initiated. With fewer firms in the cartel the reduction of profits that results from multilateral entry is reduced ($\pi^*(1) - \pi^*(I) > \pi^*(1) - \pi^*(I-1)$) which in turn increases the total returns to a deviation.

However, where a merger increases the size of the smallest component of the partition the merger may increase the stability of the cartel. In this case combining two components of a partition both reduces the merged firm's incentive to engage in an initial deviation ($\sum_{i\neq j} n_i > \sum_{i\notin\{j,l\}} n'_i$), and increases the merged firm's stake in the success of the collusive agreement ($(n_j + n_l)\pi^*(1) > n_j\pi^*(1)$). Nevertheless, the reduction in the severity of grim-strategy punishment may still dominate and as such the effect of a merger on stability is in general indeterminate.

A merger causes an asymmetric change in the relative market shares of the firms in a cartel. It is also interesting to consider the effect of a change in cartel participation which retains the relative market shares of the remaining firms. The following proposition supposes a collusive agreement in which each firm is a monopolist in a single market. Changing the number of firms is assumed not alter the composition of the remaining firms' partitions.[16]

PROPOSITION 10: *Consider a separable, identical and symmetric multi-market game satisfying assumptions 1 and 3. Let (P, a^{GS}) represent a grim-strategy collusive agreement and suppose that the partition P satisfies $n_i = 1$ for all i. A cartel with q members is more stable than a cartel with $q+1$ members if and only if,*

$$\pi^*(1) > q^2\pi^*(q) - (q^2-1)\pi^*(q+1). \tag{15}$$

Proof. Let $\delta^{GS}(q)$ be the critical discount factor for the collusive agreement (P, A^{GS}) with q firms. From (3),

$$\delta^{GS}(q) = \frac{(q-1)\big(\pi^*(2)-c\big)}{\pi^*(1)+(q-1)\big(\pi^*(2)-c\big)-q\pi^*(q)}.$$

The discount factor $\delta^{GS}(q)$ increases with the addition of one more firm if and only if $\delta^{GS}(q+1)-\delta^{GS}(q) > 0$ which in turn implies (15). □

As in proposition 9 reducing the number of members of a cartel also reduces the severity of the punishments that can be levelled against a firm. However, given that all firms have equal shares of the markets reducing the number of members of the cartel also reduces the incentive to initiate a deviation as the largest deviation available to a firm is to enter the remaining $I-1$ markets.

Condition (15) illustrates the balance of these two factors. Assumption 3 bounds the term $\pi^*(q+1)$ such that $\pi^*(q) > \pi^*(q+1) > 0$. Taking the limit of (15) as $\pi^*(q+1) \to 0$ yields,

$$\pi^*(1) > q^2\pi^*(q),$$

indicating that where the punishment that can be delivered by q firms is much greater than the punishment that can be delivered by $q-1$ firms, consolidation of ownership within the cartel will only increase stability if $\pi^*(q)$ is already very small relative to

[16] The number of markets in the game is innocuous as the critical discount factor depends only on the proportion of markets controlled by each firm.

the monopoly profit. Note that this condition is much stronger that the necessary condition (4) established in proposition 3. Contrast this with the limit of (15) as $\pi^*(q+1) \to \pi^*(q)$,

$$\pi^*(1) > \pi^*(q),$$

a condition which must be satisfied for collusion to be stable at any $\delta \in (0,1)$. Intuitively, if the punishment that can be delivered by q firms is approximately the same as that which can be delivered by $q-1$ firms then the sole effect of consolidation is to reduce the number of markets monopolized by rival firms thus reducing the incentive to deviate.

5.4 Cartel Detection

In a multi-market setting, collusion at the extensive margin can be implemented by a smaller group of managers than collusion at the intensive margin.[17] Consider the BW (1990) model of multi-market contact, each firm in the colluding cartel must move its actions away from its instantaneous best response in each of the markets subject to the collusive agreement. Consequently, if all firms are present in every market then $\|N \times I\|$ groups of market level managers have knowledge of, and are possibly active participants in, facilitating collusion at the intensive margin of some market.

Contrast this with an extensive margin collusive agreement across the same set of markets. Because each firm i confines its activities to the markets in $N_i \cup N_\varnothing$ the total number of market level management groups is equal to $\|N\| + \big(\|I\| - 1\big) \times \|N_\varnothing\|$ which is significantly less than $\|N \times I\|$. Moreover, these market level managers need not have any knowledge of the collusive arrangement. Firm i's management for a market $n \in N_i$ pursue monopoly strategies when no other firm is present in the market, and respond to entry by adopting the appropriate oligopoly strategy. Likewise, managers in a market $n \in N_\varnothing$ always adopt I-opoly strategies. It follows that knowledge of the cartel can be confined to the firm level management of the colluding firms; specifically, to those managers who are responsible for making the market participation decisions on behalf of their firms.

To the extent that restricting the number of people aware of an illegal activity reduces the risk of detection, a cartel operating in a multi-market environment has a strong incentive to confine collusion to the extensive margin.

[17]See Harrington (2005) for a review of tools and studies involved in detecting collusion at the intensive margin.

5.5 Challenges for Antitrust Enforcement

The nature of collusion at the extensive margin poses a number of challenges for antitrust authorities. Cartels engages in a pattern of mutual forbearance in which each participating firm refrains from entering any market allocated to one of its partner firms. In order to punish a cartel an antitrust authority must be able to demonstrate that the absence of a firm from a market has the purpose of supporting the system of mutual forbearance.

Additional complications arise where a collusive agreement crosses national boundaries. Consider the case in which the boundary of each market corresponds to the boundary of an antitrust authority. In order for an antitrust authority to prosecute members of the cartel the authority must have the power either to punish firms located in foreign countries for failing to establish a presence within the antitrust authority's jurisdiction, or to punish a local firm for failing to expand beyond the boundaries of the authority's mandate.

Aggressive punishment of cartel behaviour within a jurisdiction may be one factor that prompts cartels to confine their collusion to the extensive margin in the first place. By increasing the cost of coordinating actions within a market, an anti-trust authority creates an incentive for a cartel to structure its activities such that no two members are coordinating their activities within a given jurisdiction.[18]

Another perverse consequence of vigorous antitrust enforcement may arise where the markets that are subject to the collusive partition are all located within a single regulatory jurisdiction. While it may be difficult to identify collusive behaviour where firms are adhering to their collusive strategies, tit-for-tat entry and withdrawal involves a distinctive pattern of behaviour that may attract the attention of anti-trust authorities. The problem with punishing a firm for withdrawing from a market — as was done *Rural Press* case — is that it provides a cartel with an incentive to discard tit-for-tat enforcement mechanisms in favour of persistent punishments such as grim-strategy enforcement which in turn increases the stability of the cartel. An immediate corollary of this observation is that in aggressively punishing firms for reverting to collusive participation, an antitrust authority reduces the credibility of threats of predatory entry, destabilising cartels that rely on such threats.[19]

---

[18] Under intensive margin collusion, international anti-trust enforcement may lead to a free-riding effect whereby authorities in one jurisdiction leave it to others to engage in costly enforcement of a collusive arrangement as that will destabilize the international cartel (see Choi and Gerlach, 2009). Here that option is not available as compulsion of entry is not generally considered an anti-trust enforcement policy. Nonetheless, the returns to international coordination of anti-trust enforcement are likely to be high when there is extensive margin collusion across national boundaries.

[19] Similar counter-intuitive results arise in collusion on the intensive margin. See, for example,

6 Discussion

This paper develops a framework for studying collusion at the extensive margin. We show that collusion can be sustained by a range of enforcement mechanisms and that while the grim-strategy is the most stable, proportional response enforcement dominates targeted response enforcement in the presence of simple forms of uncertainty.

However, this is not a comprehensive treatment. The model developed in this paper utilizes an artificially simple state space in which a firm's participation in a market is binary; either in or out. It may be more realistic to model entry as taking a number of periods. Moreover, having entered a market it is reasonable to assume that a firm is committed to being present in a market for some minimum period of time. We employ the simple structure to illustrate the possibility of a collusive MPE in the simplest possible framework. Adding complexity to the state space will tend to increase the range of possible enforcement mechanisms, adding richness to the model without altering the qualitative nature of the results.

Enriching the model may also improve the stability of the tit-for-tat enforcement mechanisms. Suppose for example, that a firm can combine entry with a commitment to remain in the market for a number of periods. Such a commitment could take the form of a market specific contracts to supply a product for a number of periods, or to rent real estate or hire labour. Given that these commitments would increase the length of the punishment phase, relative to the length of uncontested entry, the ability to make the commitment must improve the stability of the collusive agreement.

The ability to commit to entry for more than one period will also be useful where entry costs exceed oligopolistically competitive profits (a violation of assumption 3). In this case firms can commit to entry for a period of time sufficient in length for the punishing firm to recoup the initial cost thereby facilitating punishment.

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.

McCutcheon (1997).

References

Bernheim, B. D. & M. D. Whinston (1990), Multimarket Contact and Collusive Behavior, *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.

Brandenburger, A. & H. Stuart (2007), Biform Games, *Management Science* 53, pp. 537–549.

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.

Fershtman, C. & A. Pakes (2000), A Dynamic Oligopoly with Collusion and Price Wars, *RAND Journal of Economics* 31 (2), pp. 207–236.

Feinberg, R. M. (1984), Mutual Forbearance as an Extension of Oligopoly Theory, *Journal of Economics and Business* 36 (2), pp. 243–249.

Feinberg, R. M. (1985), "Sales-at-Risk": A Test of the Mutual Forbearance Theory of Conglomerate Behavior, *Journal of Business* 58 (2), pp. 225–241.

Feuerstein, S. (2005), Collusion in Industrial Economics – A Survey, *Journal of Industry, Competition and Trade* 5 (3), pp. 163–198.

Friedman, J.W. (1971), A Noncooperative Equilibrium for Supergames, *Review of Economic Studies* 38 (1), pp. 1–12.

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

Harrington, J. E. (2005), Detecting Cartels, *Handbook of Antitrust Economics*, forthcoming.

Jacquemin, A. and M.E. Slade (1989), Cartels, Collusion and Horizontal Merger, *Handbook of Industrial Organization* Vol.1, North Holland: Amsterdam, Chapter 7.

Maskin, E. & J. Tirole (1988a), A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs, *Econometrica* 56 (3), pp. 549–569.

Maskin, E. & J. Tirole (1988b), A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles, *Econometrica* 56 (3), pp. 571–599.

McCutcheon, Barbara (1997), Do Meetings in Smoke-Filled Rooms Facilitate Collusion? *Journal of Political Economy*, 105, pp. 330- -350.

Porter, R.H. (2005), Detecting Collusion, *Review of Industrial Organization* 26 (2), pp. 147–167.

Porter, R.H. and J.D. Zona (1999) Ohio School Milk Markets: An Analysis of Bidding, *Rand Journal of Economics* 30 (2), pp. 263–288.

Shapiro, C. (1989), Theories of Oligopoly Behavior, *Handbook of Industrial Organization* Vol.1, North Holland: Amsterdam, Chapter 6.

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).