

CDE
October 2012

MULTI-MARKET COLLUSION WITH TERRITORIAL ALLOCATION

ADITYA BHATTACHARJEA

Email: aditya@econdse.org
Delhi School of Economics
University of Delhi

UDAY BHANU SINHA

Email: uday@econdse.org
Delhi School of Economics
University of Delhi

Working Paper No. 217

Centre for Development Economics
Department of Economics, Delhi School of Economics

Multi-market Collusion with Territorial Allocation^Ψ

Aditya Bhattacharjea ^a and Uday Bhanu Sinha ^b

Delhi School of Economics

Version: October 2012

Abstract: This paper develops a supergame model of collusion between price-setting oligopolists located in different markets separated by trade costs. The firms produce a homogenous good and sustain collusion based on territorial allocation of markets. We first show, in a more general framework than some earlier literature, that a reduction in trade costs can paradoxically increase the sustainability of collusion. Then we prove a new paradox where the scope for collusion may be enhanced by an increase in the number of firms. We discuss several implications for trade and antitrust policy in this context.

Keywords: Multimarket contact, collusion, trade costs, territorial allocation, cartels

Address for Correspondence: Department of Economics, Delhi School of Economics, University of Delhi, Delhi – 110007.

^Ψ We are grateful to Abhijit Banerji and Sabyasachi Kar for useful discussion at different stages of this paper.

^a Email: aditya@econdse.org

^b Email: uday@econdse.org

Multi-market Collusion with Territorial Allocation

Aditya Bhattacharjea and Uday Bhanu Sinha

1. Introduction

Since the 1990s, there has been a dramatic increase in the prosecution of cartels by American and European antitrust/competition agencies. Many of these cartels have involved firms from more than one country, and many of these have been charged with dividing up international markets on the basis of territorial allocation. Typically, firms have reciprocally agreed to stay out of each other's home markets ("mutual forbearance") as well as other markets traditionally served by their rivals, respecting each other's "spheres of influence". The first such major cartel case involved two European chemical giants, Britain's Imperial Chemical Industries (ICI) and Belgium's Solvay. They had maintained a long-standing agreement, known as "Page 1000", after the Second World War, whereby Solvay was to sell soda ash almost exclusively in continental Europe, and ICI in the British Commonwealth and the rest of Asia, Africa and South America. Other smaller European producers also respected the "home market principle" of not entering any national market where their rivals had production facilities. The Page 1000 agreement was supposedly terminated in 1972 when Britain entered the European Community, making such market sharing illegal under EC competition law. But the two firms continued to stay out of each other's markets. In 1990, this arrangement was found to be a 'concerted practice' by the European Commission.¹

Many instances of territorial allocation were discovered in the decades following the ICI-Solvay decision. In 1994, the Commission fined 42 cement producers for (among other infringements) agreeing not to enter each other's home markets. In subsequent cases involving steel tubes and methionine, it found that European producers had agreed not to sell to each other's national markets and to Japan, with Japanese producers reciprocating.² Similarly, in the choline chloride (vitamin B4) cartel, three manufacturers in North America reached an agreement (known as the Ludwigshafen protocol) with their three European rivals

¹ This account is taken from the text of Commission Decision 91/297/EEC, published in the *Official Journal of the European Communities*, L152 (1991). The Commission decision was reversed twice on appeal on procedural grounds, but the two firms did not dispute the fact that they did not sell into each other's markets. In fact, they introduced expert testimony to argue that this was a possible outcome of Cournot competition.

²The same principle was discovered in a case prosecuted on both sides of the Atlantic involving American, German, and Japanese producers of graphite electrodes.

to withdraw from each other's home markets, and to share the Latin American and Asian markets.³ Apart from these individual cases, a recent study of 81 international cartels detected by European and American competition agencies between 1980 and 2007 found that eighty per cent of them allocated territories or specific customers to their members (Levenstein and Suslow, 2011, p.475).

Historically, territorial allocation by international cartels was prevalent in the early twentieth century and was very common during the depressed 1930s. A notorious example was an arrangement covering synthetic fibres, whereby ICI agreed with its American rival du Pont to stay out of the US, Central American and Venezuelan markets, in exchange for du Pont staying out of most of the British Empire. The two firms shared Canada and the rest of Latin America through joint sales ventures (Scherer, 1994, pp.44-45, 58). A study of 71 international cartels of the interwar period found that 30 of these involved exclusive territories; along with export quotas, this was the most frequent form of cartel organization in the chemicals and minerals industries (Suslow, 2005, Table 4, p.717). Even earlier, according to Notz (1920), 114 international cartels were known to have existed before the First World War, most of them involving territorial allocation. He described in this context cartels producing steel rails, quinine, aluminium, and explosives.⁴

Some important features of the cartels covered in this historical review can be summarized as follows: (i) The cartels were based on *spheres of influence* (SOI) in territorial markets, respecting the home market principle. (ii) Most of them involved more than one firm in each territory. (iii) Apart from adhering to the home market principle, many of the agreements divided up third-country markets. (iv) These cartels were predominantly found in industries producing homogenous products. And (v) collusive prices were constrained by the

³These cases are drawn from Harrington (2006), Connor (2007), and De (2011), who also provide information on many other international cartels that functioned on the basis of global price fixing and sales quotas rather than territorial allocation. In many cases, territories were allocated by continent or country, while quotas were used to divide up the markets among firms within those territories.

⁴Most of the cases reviewed above concerned allocation of national markets by international cartels, but large domestic markets could also be carved up into exclusive territories. In an early American case interpreting the Sherman Act, six producers of cast iron pipe in the west and south of the country were found to have allocated specific cities among themselves, and had fixed prices just low enough to prevent eastern manufacturers from being competitive after allowing for freight charges (*Addyston Pipe & Steel Co. v. United States*, 175 U.S. 211 (1899), pages 215-16 and 236-237). For that matter, many of the recent EU cases involved territorial allocation within the integrated internal market of the European Union.

possibility of arbitrage in some cases. In this paper we provide a theory of cartels with all these features.

We start off with a standard supergame model of collusion between price-setting oligopolists located in different markets separated by trade costs. The firms produce homogenous goods and try to sustain collusion based on territorial allocation of markets. We first prove that a decrease in trade costs may promote collusion, a result that we call the *trade cost paradox*. Earlier authors have obtained this paradox, along with some other implications of collusion with SOI, using linear demand. We extend these results to general demand. Then we prove a new paradox with interesting policy implications. We show that the scope for collusion is enhanced by an increase in number of firms over some range (the *competition paradox*). Several of our results conform to the real-world cartel cases we reviewed above, but have not figured in the existing theoretical literature so far.

The common element of these two paradoxes is that pro-competitive changes in the economic environment might actually promote collusion. This runs counter to the conventional wisdom in economic thinking. Therefore, it is imperative to revisit some standard policies in the light of our findings. While the first paradox shows that trade liberalization is not necessarily a substitute for antitrust/competition policy, the competition paradox has some more subtle positive and normative applications. We show that tariff-induced entry is even more harmful than suggested by the earlier trade policy literature, and we also obtain some unsettling implications for standard antitrust practices. In particular, we show that the wrong industries may be investigated for cartelization, and that potentially harmless or even beneficial mergers may be disallowed.

We proceed as follows. In section 2, we set up a basic model of multimarket duopoly with three markets A, B and C, with one firm each in markets A and B. We establish that collusion with firms monopolizing their home markets and sharing the third market is an equilibrium if the firms do not discount future profits too heavily. In section 3 we derive some comparative static results on the effect of varying market sizes on the likelihood of collusion. From section 4 onwards we confine ourselves to a scenario with only markets A and B. We establish the trade cost paradox with symmetric trade costs in section 4. In section 5 we show that the paradox also holds for a range of asymmetric trade costs, and that tariff harmonization facilitates collusion. The competition paradox is analysed in section 6. In all cases, we investigate how far the paradoxes are sensitive to our assumptions. In the context of

the competition paradox, we endogenize firms' entry and merger decisions in section 7, deriving implications for antitrust policy. Conclusions and possible directions for future research are outlined in section 8.

Related Literature

The possibility of SOI emerging as a collusive equilibrium in an infinitely repeated game was first demonstrated by Pinto (1986). Section V of the classic paper by Bernheim and Whinston (1990; hereafter BW) formalized the analysis of multimarket contact between duopolists located in two spatially separated markets.⁵ Both papers assumed that deviation from the collusive arrangement would be punished by firms reverting to a more competitive outcome, but did not employ the standard grim-trigger punishment involving eternal reversion to the Nash equilibrium of the constituent stage game. We highlight the basic differences between the grim-trigger and BW punishments in the context of our model. The effects of varying market size on the likelihood of SOI, using both punishment strategies, were analyzed by Lommerud and Sørsgard (2001) with linear demand. They also obtained the inverse relationship between trade costs and the likelihood of collusion (which we call the trade cost paradox) with grim-trigger punishment,⁶ and showed that it did not hold with BW punishments. We prove all these results for general demand. Moreover, we go further to show that the relationship between the critical value of the discount factor that supports collusion and the level of trade costs depends on the convexity or concavity of the demand curve.

⁵Bernheim and Whinston (1990) noted that, apart from the possibility that cross-market retaliation against defection can reinforce collusion when firms compete in markets that differ in some respect, geographically distinct markets separated by trade costs are especially suited to collusion. For international cartels, this arrangement has several advantages as compared to sharing a unified world market at an agreed price backed by firm-specific sales quotas. First, each firm in the cartel can be given a monopoly in the countries where it has lower costs of supplying buyers because it has production or distribution facilities in that country or nearby. This allows for both monopoly pricing and rationalization of production, raising the pool of collusive profits and thus increasing the attractiveness of collusion relative to competition. Second, defection is discouraged because it must take the form of selling in another firm's allotted territory. This involves extra transportation costs, and is also easier to detect (by monitoring trade flows) than variations in output or price. But allocation of national markets also has more subtle and far-reaching implications, which we explore in this paper.

⁶This was independently proved by Gross and Holahan (2003), also using linear demand. In an appendix to their paper, Lommerud and Sørsgard (2001) showed that the result held for inverse demand characterised by $P = 1 - q^b$ with $b > 0$, using numerical simulations for 1000 values of b .

Several earlier authors have qualified the trade cost paradox. Lommerud and Sørgaard (2001) showed that it disappears if the duopolists compete in quantities rather than prices; Schröder (2007) showed that it disappears if trade costs are *ad valorem* or fixed rather than per unit. Both these papers *assumed* that collusion would take the form of exclusive SOI, which is not always optimal for the firms. Two-way trade takes place in a collusive equilibrium if trade costs are low enough,⁷ or if marginal costs of production are increasing strongly enough (Colombo and Labrecciosa, 2007). In these papers, all of which assumed linear demand, the trade cost paradox may or may not hold, depending on the parameters of the model, and it may hold only if trade costs vary in a restricted range.⁸ Our focus is exclusively on SOI, which is the only collusive equilibrium with constant marginal costs, price competition and homogeneous products. In this particular context, we show that the trade cost paradox holds for general price-elastic demand and any level of non-prohibitive trade costs. The paradoxical nature of this result is reinforced in our setting, because we show that a reduction in trade costs enhances the prospects for collusion even though no trade occurs in either the non-cooperative or the collusive equilibrium. However, supplementing the qualifications offered by earlier papers, we also show that the paradox does not hold with more than one firm in each country or with inelastic unit demand.

Our second paradox is entirely new, to the best of our knowledge. One of the corollaries of our competition paradox bears some resemblance to results derived in an older literature on inefficient entry induced by tariff protection when markets are oligopolistic (for example, Markusen and Venables, 1988; Bhattacharjea, 1995), but unlike in that literature the reciprocal relationship between incentives for entry and collusion plays a crucial role in our paper. Another early body of literature (e.g. Davidson, 1984; Rotemberg and Saloner, 1989) modelled the effects of various trade restrictions on incentives to collude, but with a given market structure and in a single country's market. In contrast, we examine the effect of entry

⁷ See Bond and Syropoulos (2008) and Bond and Syropoulos (2012) for quantity-setting firms, and Akinbasoye et al (2012) for price-setting firms with differentiated products. Such trade would be regarded as intra-industry trade, or cross-hauling in the case of homogenous products.

⁸ In other contributions to this literature, Belleflame and Bloch (2008) showed that collusion would take the form of SOI for high levels of fixed costs, but did not examine trade costs. Salvo (2010) models SOI with trade costs but assumes that consumers are located Hotelling-style on a line segment between the firms, unlike all the papers cited here (and ours), in which consumers and firms in each country are placed at the same dimensionless location. The effect of tariff harmonization is similar to a result obtained by Fung (1991), who showed that collusion would be strengthened by greater cost symmetry, where costs included both production and (possibly asymmetric) transport costs. However, like the other papers cited in footnote 7 above, Fung modeled collusive intra-industry trade with quantity-setting firms facing linear demand in each market.

and mergers on collusion in two national markets. Much more recently, Bond and Syropoulos (2012) have shown that an increase in the number of firms is pro-collusive (pro-competitive) at low (high) levels of trade costs. This result is the opposite of our competition paradox. However, their model is fundamentally different. They employ a Cournot stage game and analyze variation in the number of local firms while keeping the total number of firms in the world unchanged, treating this as preferential trade liberalization.⁹

2. The Model

Assume to begin with that there are three identical markets (A, B and C) and two identical firms (1 and 2) producing a homogeneous product. Firm 1 is located in market A, and firm 2 in market B. There are no domestic producers in market C. Each firm incurs a cost of c per unit to produce and sell within its own market, but must incur additional trade costs of t per unit to sell in another market, so its delivered cost there is $c^* = c+t$ per unit. Competitive arbitrageurs can exploit price differences between markets by buying where the price is lower and reselling elsewhere by incurring the same trade costs of t per unit. For most of the paper, we shall assume that t is the same in both directions between each market-pair. Transport costs are the obvious interpretation, making the model applicable to any cartel that allocates spatially separated markets amongst its members. In one section, we allow for asymmetric trade costs, which can be interpreted as import tariffs that differ between countries. This is more relevant to international cartels, and for the sake of consistency we shall henceforth refer to markets A and B as countries, with market C being treated as the rest of the world.

Demand in country j is given by $Q_j = q(P_j)$, with the following additional assumptions:

A1: There exists a finite choke price $\bar{P}_j = \bar{P} \forall j$, such that $q(P) > 0$ if $P < \bar{P}$, and $q(P) = 0$ if $P \geq \bar{P}$.

A2: $q'(P_j) \leq 0$, with equality in the special case of inelastic unit demand.

A3: For downward-sloping demand, $(P - c)q''(P) + q'(P) < 0$.

⁹ They have used a particular demand specification according to which an increase in the number of local firms expands the size of the domestic market.

A1 and A2 are standard assumptions. A3 imposes a restriction on the convexity of the demand curve, which we use below to ensure that arbitrage is not profitable when there is collusion in all three markets. A3 is necessary only as long as we retain market C. We drop this in section 4 onwards, and assume instead the following restriction on convexity of demand function that is implied by A3 :

$$A3': (P - c)q''(P) + 2q'(P) < 0$$

This is the familiar condition that ensures the concavity of the monopolist's profit function, and along with A1, guarantees the existence of a unique monopoly price $P^m(c)$.

Above a certain level of t , which we call \bar{t} , trade costs become prohibitive and the domestic firm can charge the monopoly price with no disciplining effect from competitively-supplied imports. Nor can it benefit from violating a cartel agreement that involves monopoly pricing in each market. Variations in trade costs above \bar{t} are redundant, as the markets are insulated from foreign competition. We therefore assume:

$$A4: c^* \leq P^m(c), \text{ which is the same as assuming } t \leq P^m(c) - c \equiv \bar{t}$$

In each period, firms simultaneously set their prices. They interact for an infinite number of periods, with a common per-period discount factor of δ . To begin with, we derive conditions under which they collude by treating their respective home markets as exclusive spheres of influence, but share market C.¹⁰ This collusive outcome will be supported by either of the two punishment strategies that are standard in the literature. We first examine grim-trigger punishments in which, after a deviation, firms revert to the Bertrand-Nash equilibrium of the single-period game. The results will then be contrasted with those supported by the more severe punishments used by Bernheim and Whinston (1990).

To economize on notation, define:

$$\pi_j^m(c) \equiv (P^m - c)q(P^m) \tag{1}$$

where $\pi_j^m(c)$ is the monopoly profit that a firm based in country j can earn in its own market. We allow for variations in market size by scalar multiplication of $q(P)$. This leaves a monopolist's first-order optimality condition unaffected and hence translates directly into

¹⁰ As demonstrated by Bernheim and Whinston (1990), it is optimal for a cartel to allocate home markets exclusively to the respective home firms. It can easily be shown that equal sharing of market C is also optimal if they have symmetric costs of supplying it.

multiplication of the relevant profit expressions by the same scalar, with no change in P^m . Monopoly price is thus independent of local demand and depends only on costs.¹¹ Therefore, in collusion with SOI, $P^m(c)$ prevails in both markets A and B, because the local monopolists have identical production costs. The collusive price in market C is of course higher at $P^m(c^*)$ because both firms also incur trade costs. But A3 guarantees that the monopoly price cannot rise by more than the increase in marginal cost, so $P^m(c^*) - P^m(c) < c^* - c = t$, and arbitrage from markets A and B cannot undermine monopoly pricing in market C. (With inelastic demand, regardless of the number of consumers, the monopoly price is infinitesimally less than the choke price \bar{P} , assumed identical across markets, so again arbitrage is ruled out.) Therefore, with collusion in all three markets, cross-market arbitrage cannot constrain monopoly pricing in any market.

Just as variation in market size translates proportionately into changes in monopoly profits, so does division of the market between the firms. Each firm thus earns $\pi_C^m(c^*)/2$ from market C. A firm can defect from the collusive arrangement in any market by slightly undercutting the monopoly price, allowing a firm to become the sole seller in that market for one period. Formally, we follow the technical convention of having a defector charge the monopoly price instead of undercutting it by an infinitesimal amount, and award it the entire market by resolving the tie in its favour. For each market $l \neq j$, the defector therefore obtains (in addition to the monopoly profit it retains in its home market):

$$\pi_l^d(c^*) \equiv (P^m - c^*)q(P^m) \quad (2)$$

Defection in any market triggers the punishment, which takes the form of reverting to the Nash equilibrium in each market forever. But if defection anywhere is followed by competition everywhere, a deviating firm may as well maximize its short-term gain by capturing all three markets for one period, which is assumed to be the length of time it takes for other firms to detect that there has been a deviation. Then the punishment begins. In the ensuing Nash equilibrium, we have the standard Bertrand duopoly results: each firm retains its home market by charging a limit price equal to the marginal (delivered) cost c^* of imports from the other firm, while in market C the firms share the market by pricing at c^* . Thus, $P =$

¹¹This kind of demand shift (which assumes that demand variation arises from replication of identical consumers) leaves the choke price \bar{P} unaffected. For linear demand, it means that the slope of the demand curve is allowed to vary, keeping the vertical intercept unchanged at \bar{P} . This is important for contrasting our results with those of an earlier paper.

c^* in each market, and again arbitrage is unprofitable. In every period after the deviation is detected, each firm therefore earns per-period profits of

$$\pi_j^p(c) \equiv (c^* - c)q(c^*), \quad j = A, B \quad (3)$$

in its home market. It earns no profit either in its rival's home market where it makes no sales due to the rival's limit pricing, or in market C which it shares but at price equal to cost. Expression (3) therefore gives a firm's per-period global profits in the punishment phase.

Because P^m and c^* are invariant to changes in market size, profit expressions (2) and (3) are also scaled by multiplicative changes in local demand, with no change in the local price. Our subsequent analysis will repeatedly use the values of the three profit expressions (1) to (3) at the bounds of the permissible range of trade costs, as well as their derivatives with respect to trade costs, so we summarize all these results in Table 1. [at the end]

We begin by assuming that the markets are of equal size. The following result will be useful below. It is derived for firm 1, but by the symmetry of the firms and markets A and B, a similar condition can be obtained for firm 2 by interchanging subscripts A and B.

Lemma 1: $\pi_A^m(c) - \pi_B^d(c^*) - \pi_A^p(c) \leq 0$.

Proof: Suppose $\pi_A^m(c) - \pi_B^d(c^*) - \pi_A^p(c) > 0$. From the definition of the profit expressions above, this can be rewritten as $(P^m - c)q(P^m) - (P^m - c^*)q(P^m) > (c^* - c)q(c^*)$. For $t > 0$, that is, for $c^* > c$, this inequality can be rewritten as $q(P^m) > q(c^*)$. However, for price-elastic demand and $t < \bar{t}$, $P^m > c^*$ and therefore $q(P^m) > q(c^*)$. Thus, the supposition must be false, and Lemma 1 holds. Further, it must hold with equality at the extremes of the permissible range of trade costs: for $t = 0$, $c^* = c$, so $\pi_A^m = \pi_B^d$ and $\pi_A^p = 0$; for $t = \bar{t}$, $c^* = P^m$, so $\pi_B^d = 0$ and $\pi_A^p = \pi_A^m$, and Lemma 1 holds with equality in both cases. If each country instead has a unit mass of identical consumers, each buying exactly one unit of the good at any price up to \bar{P} , then $q(P) = 1$ for all $P < \bar{P}$. Then, for a monopoly price infinitesimally below \bar{P} , $\pi_A^m(c) = \bar{P} - c = (\bar{P} - c^*) + (c^* - c) = \pi_B^d(c^*) + \pi_A^p(c)$, and again Lemma 1 holds with equality.

We are now in a position to derive the incentive-compatibility conditions (ICCs) for collusion with SOI. For firm 1, based in country A, collusion is sustainable if the following ICC holds:

$$\frac{\pi_A^m(c) + \pi_C^m(c^*)/2}{1-\delta} \geq \pi_A^m(c) + \pi_B^d(c^*) + \pi_C^d(c^*) + \frac{\delta\pi_A^p(c)}{1-\delta} \quad (4)$$

The left side of the inequality gives the present value of the firm's monopoly profit in market A while staying out of market B and sharing market C. The right side gives the payoff from defection: for one period, the firm retains its profit in market A and also captures markets B and C at the monopoly price, but incurs higher unit costs on these export sales. Thereafter, the defector can expect only the Nash level of profits in its home market. Recalling that $\pi_C^d(c^*) = \pi_C^m(c^*)$, because defection in market C garners the entire monopoly profit in that market, we solve (4) as an equation to obtain the minimum discount factor required to sustain collusion across all three markets:

$$\delta_{ABC}^* = \frac{\pi_B^d(c^*) + \pi_C^m(c^*)/2}{\pi_A^m(c) + \pi_B^d(c^*) + \pi_C^m(c^*) - \pi_A^p(c)} \quad (5)$$

Exploiting symmetry, we get the same value of δ_{ABC}^* for firm 2 by interchanging subscripts A and B. Collusion with exclusive spheres of influence is thus sustainable for $\delta \geq \delta_{ABC}^*$.

In what follows, we shall follow the standard practice of interpreting anything that reduces (increases) this critical minimum discount factor—and its variants to be derived below—as facilitating (inhibiting) collusion, given δ . Our first such results follow directly from (5). First, as we would expect, δ_{ABC}^* is increasing in $\pi_A^p(c)$: the more the home market is sheltered from foreign competition, the weaker the punishment, and thus the more patient the firms will have to be to resist the temptation to deviate. Second, $\delta_{ABC}^* \geq 1/2$, with equality in the special case of inelastic unit demand.¹²

3. Variation in market size

We can also examine the consequences of varying the relative sizes of the markets, beginning with a situation where they are assumed to be identical so as to allow (5) to apply to both

¹²*Proof:* Suppose not, so $\delta_{ABC}^* < 1/2$. Substituting definitions (1) to (3) and rearranging, this implies $\pi_A^m(c) - \pi_B^d(c^*) - \pi_A^p(c) > 0$, which contradicts Lemma 1, so the supposition is false. In the special case of inelastic unit demand, Lemma 1 implies that $\pi_A^m(c) = \pi_B^d(c^*) + \pi_A^p(c)$. Substituting this into (5) and simplifying gives $\delta_{ABC}^* = 1/2$.

firms.¹³ Now allow market B to become larger. Recalling that multiplicative variations in demand $q(P)$ translate into identical variations in the relevant profit expressions, we can obtain (for firm 1):

$$\frac{\partial \delta_{ABC}^*}{\partial \pi_B^d} = \frac{[\pi_A^m(c) - \pi_A^p(c) + \pi_C^m(c^*)/2]}{[\pi_A^m(c) + \pi_B^d(c^*) + \pi_C^m(c^*) - \pi_A^p(c)]^2} > 0 \quad (6)$$

where the numerator can be unambiguously signed because $\pi_A^m(c) > \pi_A^p(c)$, regardless of the presence or size of market C. Inequality (6) tells us that if the market in country B is larger than in country A, the scope for collusion with exclusive spheres of influence is reduced. Unless the firm values future profits sufficiently, it is unable to resist the temptation of violating the agreement and invading its rival's large domestic market. On the other hand, from the perspective of firm 2, the subscripts A and B on the right side of the equalities in (5) and (6) are interchanged, so that a larger market B (now representing home profits for firm 2) *reduces* its critical discount factor. But collusion requires the ICCs of *both* firms to be satisfied, and hence the higher δ_{ABC}^* of firm 1 is the binding constraint. Hence, generalizing the result obtained by Lommerud and Sørsgard (2001) for two countries with linear demand, we have

PROPOSITION 1: Asymmetry in the size of the firms' home markets increases the critical discount factor, and hence reduces the scope for collusion.

What about changes in the relative size of market C, keeping markets A and B identical? (5) gives us:

$$\frac{\partial \delta_{ABC}^*}{\partial \pi_C^m} = \frac{[\pi_A^m(c) - \pi_B^d(c^*) - \pi_A^p(c)]}{[\pi_A^m(c) + \pi_B^d(c^*) + \pi_C^m(c^*) - \pi_A^p(c)]^2} < 0 \quad (7)$$

where we can unambiguously sign the numerator on the basis of Lemma 1 for downward-sloping demand with $0 < t < \bar{t}$.¹⁴ The scope for collusion is therefore increasing in the profitability of market C.¹⁵ Collusive sales to market C provide an additional profit opportunity, the loss of which reinforces the punishment.

¹³ Notice that δ_{ABC}^* is homogeneous of degree zero in the profit expressions, and thus the prospects for collusion are unaffected by proportionate expansion of demand in all three markets.

¹⁴ For inelastic demand, the numerator is identically zero, as we already know from the unchanging value of $\delta_{ABC}^* = 1/2$.

¹⁵ Variation in the profitability of market C could be on account of its size or the firms' costs of supplying it, which (within some range) need not be the same as the cost of supplying each other's

Suppose now that market C does not exist, or that it is too small relative to fixed costs of exporting to make it worthwhile for firms based in countries A and B. Eliminating $\pi_C^m(c^*)$ from the above expressions gives us the standard two-country scenario used by earlier authors, but without their assumption of linear demand. We get

$$\delta_{AB}^* = \frac{\pi_B^d(c^*)}{\pi_A^m(c) + \pi_B^d(c^*) - \pi_A^p(c)} \quad (8)$$

As the firms and markets are mirror images of each other, firm 2 will have a similar ICC, simply interchanging A and B, and thus an identical δ_{AB}^* . It is straightforward to show, using the same approach as for the three-market case, that $\delta_{AB}^* \geq \frac{1}{2}$, with equality in the case of inelastic demand.¹⁶ For downward-sloping demand, we can show that $\frac{1}{2} < \delta_{ABC}^* < \delta_{AB}^*$ which is also evident from (4) as we reduce $\pi_C^m(c^*)$ towards zero. Further, setting $\pi_C^m(c^*) = 0$ in (6) establishes that $\partial \delta_{AB}^* / \partial \pi_B^d > 0$, so Proposition 1 also holds in the two-market case.

Bernheim-Whinston punishments

These results change significantly if, instead of grim trigger punishments, we use the maximal punishments used by Bernheim and Whinston (1990) in their two-market SOI model, with prices on the punishment path calibrated so as to ensure that the present value of profits is zero for all firms. Although earlier authors who developed SOI models refer to this as ‘optimal’ punishment, we avoid this terminology and describe it as Bernheim-Whinston (BW) punishment, because it can take one of two forms: (a) Both firms charge prices equal to c in markets A and B and c^* in market C, with firms making all the sales in their respective home markets while sharing market C for all periods after a deviation. This is an extension to the three-market scenario of the formulation suggested by Bernheim and Whinston, which as they acknowledged involves a Nash equilibrium with firms playing weakly dominated strategies. Alternatively, they proposed (b), the ‘stick and carrot’ punishment strategy of Abreu (1986), with firms charging below-cost prices for one period, and then reverting to the collusive price if all firms have participated in the punishment. It is the latter approach that is

home markets. It is also easy to show that departures from equal sharing of market C raises the critical discount factor of the firm with the smaller share, and therefore the binding δ_{ABC}^* .

¹⁶*Proof:* Suppose not, i.e. $\delta_{AB}^* < \frac{1}{2}$. Then (8) implies $\pi_A^m(c) - \pi_B^d(c^*) > \pi_A^p(c)$, which contradicts Lemma 1. Thus the supposition is false, and $\delta_{AB}^* \geq \frac{1}{2}$. If each country has a unit mass of identical consumers, each buying exactly one unit of the good at any price up to \bar{P} , then $q(P) = 1 \forall P < \bar{P}$. Then $\pi_A^m(c) = \pi_B^c(c^*) + \pi_A^p(c)$. Substituting from (1), (2) and (3) gives $\delta_{AB}^* = \frac{1}{2}$.

usually described as ‘optimal’ punishment, although there is some doubt about its optimality when firms have asymmetric costs (Miklos-Thal, 2009). BW themselves stated that “our basic points regarding the effect of multimarket contact do not rely in any way on the use of optimal punishments” (Bernheim and Whinston, 1990, p.11, n. 20).

Both approaches (a) and (b) can be captured in the ICC by setting the discounted value of profits in the punishment phase equal to zero. This amounts to setting the last term in (4) to zero, and thus suppressing $\pi_A^p(c)$ in all subsequent expressions. Using the same procedure used for grim-trigger punishment, we get (from the perspective of firm 1) the following results for the minimum discount factor $\hat{\delta}$ with BW punishment:

$$\hat{\delta}_{ABC} = \frac{\pi_B^d(c^*) + \pi_C^m(c^*)/2}{\pi_A^m(c) + \pi_B^d(c^*) + \pi_C^m(c^*)} < \frac{1}{2} \quad (9)$$

$$\frac{\partial \hat{\delta}_{ABC}}{\partial \pi_B^d(c)} = \frac{\pi_A^m(c) + \pi_C^m(c^*)/2}{[\pi_A^m(c) + \pi_B^d(c^*) + \pi_C^m(c^*)]^2} > 0 \quad (10)$$

$$\frac{\partial \hat{\delta}_{ABC}}{\partial \pi_C^m(c)} = \frac{[\pi_A^m(c) - \pi_B^d(c^*)]/2}{[\pi_A^m(c) + \pi_B^d(c^*) + \pi_C^m(c^*)]^2} > 0 \quad (11)$$

$$\hat{\delta}_{AB} = \frac{\pi_B^d(c^*)}{\pi_A^m(c) + \pi_B^d(c^*)} < \frac{1}{2} \quad (12)$$

In order to sign these inequalities, we used definitions (1) and (2) to get $\pi_A^m(c) > \pi_B^d(c^*)$, which is just the obvious fact that if markets A and B are of equal size, then for the same level of monopoly price, profits in the export market will be lower because of the higher costs of supplying the same quantity. Comparison with our results for grim-trigger punishment are instructive. Inequality (10), and the corresponding inequality for the two-market case obtained by setting $\pi_C^m(c^*) = 0$, show that market size asymmetry makes collusion less likely, so Proposition 1 continues to hold. But (11) shows that greater profitability of market C also makes collusion less likely, unlike with grim-trigger strategies. (12) is essentially the result derived by another route by Bernheim and Whinston (1990, p.13). Collectively, the preceding results allow us to rank the critical discount factors as follows. For downward-sloping demand curves:

$$\hat{\delta}_{AB} < \hat{\delta}_{ABC} < \frac{1}{2} < \delta_{ABC}^* < \delta_{AB}^* \quad (13)$$

while for inelastic unit demand:

$$\hat{\delta}_{AB} < \hat{\delta}_{ABC} < \frac{1}{2} = \delta_{ABC}^* = \delta_{AB}^* \quad (14)$$

It is not surprising that the greater severity of BW punishments allows collusion to be sustained for lower discount factors. Comparing (11) with (7), the reversal of the sign of the derivative with respect to the profitability of market C is surprising, but can be explained as follows. For BW punishments, the intuition is clear: enlargement of the third market increases the reward to defection, which must be neutralized by a higher discount factor so as to increase the present value of future collusive profits if collusion is to be sustained. With grim trigger punishments, however, the extra term on the RHS of (4) shows that a higher δ also increases the present value of the stream of positive Nash equilibrium profits that the defector can expect, weakening the punishment and tilting the balance still further in favour of defection. Inequality (7) shows that on balance, the latter effect is stronger: a lower discount factor must reinforce the punishment more than enough to offset the increased temptation to defect that is provided by the enlargement of market C. Another way of understanding the result is to consider market C in isolation as a standard model of collusion between identical Bertrand duopolists producing a homogeneous product, for which the familiar textbook result is $\delta^* = 1/2$. When the ICC for market C is pooled with those of A and B, then for both punishment mechanisms the resulting minimum discount factor for three markets lies somewhere between $1/2$ and that for two markets, depending on the profitability of market C.

These results suggest that if firms use BW punishments, then higher collusive profits in market C (the rest of the world, including the developing countries) may destabilize the cartel. In fact, $\hat{\delta}_{ABC}$ is minimized when $\pi_C^m(c^*) = 0$, which suggests that the firms should set $P = c^*$ in market C. However, if trade costs are not too high (specifically, if $t < (P^m - c)/2$), a higher price $P^m(c) - t > c^*$ will be required in market C to prevent a backflow of their exports that would undermine monopoly pricing in their home markets.¹⁷

¹⁷Punishment strategies are of course unobservable, so we have to infer them from collusive pricing behavior. A study of 156 price observations obtained from international cartel cases showed that (controlling for duration, market structure, and sector of operation) overcharges were slightly lower in the European and North American markets—the home bases of the firms in most cases—than in the rest of the world (Bolotova, 2009). However, lower overcharges could have been due to the greater threat of antitrust prosecution in those markets. Also, if the market in the rest of the world is relatively small, this could depress $\pi_C^m(c^*)$ sufficiently without requiring the cartel to charge less than monopoly prices in order to maintain incentive compatibility.

4. First paradox: lower trade costs facilitate collusion

4.1: Benchmark case: linear demand

For the rest of the paper, we confine ourselves to the case of two countries A and B. We proceed to establish two paradoxes, according to which ostensibly pro-competitive changes in trade costs or market structure may actually promote collusion. The first paradox, a positive relationship between the minimum discount factor and the level of trade costs, has been derived by several earlier authors with linear demand, while we shall do so for general demand. Nonetheless, we shall begin with the linear case for three reasons. First, earlier authors normalized either the slope of the demand curve to unity or marginal cost to zero. However, we need to retain positive marginal costs to allow for the possibility of BW punishments, which can require below-cost pricing in the first period after a defection. And we need flexibility in regard to the slope in order to establish a Lemma that we shall use to prove a new paradox. Second, we shall show that many of the crucial qualitative features of the relationship between the critical discount factor and the level of trade costs remain valid for general demand. Finally, we need to work out the linear demand case for numerical simulations that will be used to illustrate the paradoxes, along with analytical results for general demand.

We develop the linear case allowing for one or more firms in each market. Assume linear inverse demand $P = a - bQ$ in each market, with $a, b > 0$ and $a > c$.¹⁸ Without loss of generality, from the perspective of the country A firm(s), we can obtain $\bar{t} = (a - c)/2$ and

$$\pi_A^m = (a - c)^2/4b$$

$$\pi_B^d = (a - c - 2t)(a - c)/4b$$

$$\pi_A^p = t(a - c - t)/b$$

We now denote the critical discount factor for the case of one firm in a country as $\delta_{AB}^*(1)$ in order to facilitate comparison with results for the multi-firm case to be obtained in section 6

¹⁸Using direct demand functions of the form $q = A - BP$, but normalizing $c = 0$, Gross and Holahan (2003) derived $\delta_{AB}^*(1) = A/2(A - Bt)$. Their A is a/b and their B is $1/b$ in our notation, so this expression is the same as the one we derive below. Variation in b in our indirect demand functions cannot be replicated in their framework because it would involve simultaneous changes in both A and B in their direct demand specification.

below, with $\delta_{AB}^*(n)$ reserved for $n \geq 2$. With grim trigger punishments, we substitute the relevant profit expressions into (8), and after some simplification we get

$$\delta_{AB}^*(1) = \frac{(a-c)[a-c-2t]}{2(a-c-t)[a-c-2t]} \quad (15)$$

For $t < \bar{t} = (a - c)/2$, we can eliminate the expressions in square brackets, giving

$$\delta_{AB}^*(1) = \frac{(a-c)}{2(a-c-t)} \quad (16)$$

which increases monotonically from $1/2$ as t increases from zero. At $t = \bar{t}$, $\delta_{AB}^*(1)$ in (15) is undefined, because the ratio becomes $0/0$. However, we can establish that $\lim_{t \rightarrow \bar{t}} \delta_{AB}^*(1) = 1$ by applying L'Hôpital's rule.¹⁹ The relationship between $\delta_{AB}^*(1)$ and t is depicted by the curve in Figure 1, using a simulation with $a = 10$, $b = 1$, and $c = 2$, giving a permissible range of t between 0 and $\bar{t} = 4$.

With Bernheim-Whinston punishments, substitution of the relevant profit expressions into (12) gives

$$\hat{\delta}_{AB}(1) = \frac{a-c-2t}{2(a-c-t)} \quad (17)$$

which *decreases* monotonically from $1/2$ to zero as t goes from zero to \bar{t} , as depicted in Figure 1 for the simulation with the same parameter values. (The remaining curves in the Figure should be ignored for the present.)

4.2: The trade cost paradox with symmetric trade costs

We now examine the effect of varying levels of trade costs, still assuming them to be symmetric in both directions, but allowing for general demand. We shall show that collusion is facilitated by a reduction in symmetric trade costs, which can be interpreted temporarily either as a fall in transport costs between countries or bilateral tariff liberalization, both of which can be regarded as 'globalization'. An alternative, cross-

¹⁹The requirement for this limiting argument seems to have escaped earlier authors who derived the critical discount factor with linear demand. In Lemma 2 below we show that the same argument applies for general demand. $\delta_{AB}^*(1) = 1$ implies that collusion is unsustainable even with minimal discounting of future profits, but of course $t = \bar{t}$ means that the markets are completely segmented, and the firms can charge monopoly prices without colluding.

sectional, interpretation is that lower trade costs represent country-pairs separated by smaller distances.

With grim-trigger punishment, a decrease in symmetric trade costs has opposing effects on the stability of a SOI cartel. While increasing the attractiveness of invading the foreign market ($\pi_B^d(c^*)$ from the perspective of firm 1), it also reduces the Nash level of punishment profits $\pi_A^p(c)$ that the firm can expect thereafter in its home market, dampening its incentive to invade. Earlier authors were unable to resolve this tension except in the special case of linear demand in each market. Dispensing with this restrictive assumption, we first show that the paradox holds for grim-trigger punishments, general price-elastic demand, symmetric trade costs, and one firm in each country. We then show that it does *not* hold, or is modified, in the presence of Berheim-Whinston punishments, asymmetric trade costs, or multiple firms in each country. The last two of these modifications give rise to two new paradoxes with interesting policy implications. In all other respects, we retain the framework in which earlier authors established the trade cost paradox: two identical markets separated by constant per unit trade costs; Bertrand competition in homogenous products; constant marginal costs and no capacity constraints.

PROPOSITION 2 (Trade cost paradox): For all price-elastic, twice-continuously differentiable demand functions satisfying assumptions A1-A3', a reduction in symmetric trade costs in the relevant range given by A4 facilitates collusion in the two-country SOI model with one firm in each country and grim-trigger punishments, that is, $\partial \delta_{AB}^(1)/\partial c^* > 0$.*

Proof: From profit expressions (1) to (3), $\partial \pi_A^m(c)/\partial c^* = 0$, $\frac{\partial \pi_B^d(c^*)}{\partial c^*} = -q(P^m)$, and $\frac{\partial \pi_A^p(c^*)}{\partial c^*} = q(c^*) + q'(c^*)(c^* - c)$.²⁰ Differentiating (8) with respect to c^* , substituting the relevant derivatives, and simplifying yields

$$\frac{\partial \delta_{AB}^*}{\partial c^*} = \frac{-q(P^m)[\pi_A^m - \pi_A^p(c)] + [q(c^*) + q'(c^*)(c^* - c)]\pi_B^d(c^*)}{[\pi_A^m(c) + \pi_B^d(c^*) - \pi_A^p(c)]^2} \quad (18)$$

Note that the numerator is positive if $\{q(P^m)[\pi_A^m - \pi_A^p(c)]/\pi_B^d(c^*)\} - q(c^*) < q'(c^*)(c^* - c)$. Plugging in the values of profit expressions, this can be written as

²⁰By the strict concavity of the profit function, we know that the sign of this last derivative must be positive because the constrained (limit) price c^* is below the monopoly profit maximizing level.

$$\left\{ \frac{(P^m - c)q(P^m) - (c^* - c)q(c^*)}{P^m - c^*} \right\} - q(c^*) < q'(c^*)(c^* - c)$$

Then by rearranging terms we get

$$(c^* - c)(P^m - c^*) q'(c^*) + (P^m - c)q(c^*) > (P^m - c)q(P^m) \quad (19)$$

First note that as $t \rightarrow 0$, $(c^* - c) \rightarrow 0$ and $q(c^*) \rightarrow q(c)$. Since $q(c) > q(P^m)$, the left hand side of (19) is necessarily greater than the right hand side for low values of t . Now we show that (19) holds for all $t < \bar{t}$. Note that the RHS of (19) is constant. Differentiating the LHS with respect to c^* , we get

$$\begin{aligned} (P^m - c^*) q'(c^*) - (c^* - c)q'(c^*) + (c^* - c)(P^m - c^*) q''(c^*) + (P^m - c)q'(c^*) \\ = 2(P^m - c^*) q'(c^*) + (c^* - c)(P^m - c^*) q''(c^*) \\ = (P^m - c^*) [2q'(c^*) + (c^* - c)q''(c^*)] \end{aligned}$$

This expression is negative, because $(P^m - c^*) > 0$ for all t in the relevant range, while the concavity of the profit function ensures that $2q'(c^*) + (c^* - c)q''(c^*) < 0$. Therefore, the LHS of (19) is always a decreasing function of t in the relevant range. As $t \rightarrow \bar{t}$, $c^* \rightarrow P^m$ and the LHS coincides with the RHS. Thus (19) is satisfied for all values of t in the relevant range. Hence, $\partial \delta_{AB}^* / \partial c^* > 0$. \square

Thus, using nothing more than the concavity of the profit function, we have shown that the tension between the two effects of reduced trade costs on the critical discount factor can be unambiguously resolved for general price-elastic demand: the increase in the severity of Bertrand-Nash punishments outweighs the greater profitability of defection in the foreign market, facilitating collusion despite greater competition through market integration. The following lemma gives us more insight into the behavior of $\delta_{AB}^*(1)$, and also help us to obtain a further result below.

Lemma 2: For $t = 0$, $\delta_{AB}^*(1) = 1/2$, and as $t \rightarrow \bar{t}$, $\delta_{AB}^*(1) \rightarrow 1$.

Proof: The first result follows directly from substituting the limiting values of the constituent profit expressions at $t = 0$ (see Table 1) into (8). For the second result, notice that at $t = \bar{t}$, $\pi_A^m(c) = \pi_A^p(c)$ and $\pi_B^d(c^*) = 0$, so $\delta_{AB}^*(1)$ in (8) is undefined. But $q(c^*) = q(P^m)$ at $t = \bar{t}$, so from (20), regardless of the curvature of the demand curve, $\delta_{AB}^*(1)$ must tend to the same

value at \bar{t} as for a hypothetical linear demand curve. But we have already shown by applying L'Hôpital's rule to (15) that with linear demand, $\delta_{AB}^*(1) \rightarrow 1$ as $t \rightarrow \bar{t}$. \square

Lemma 3: For $t \in (0, \bar{t})$, the value of $\delta_{AB}^*(1)$ for a strictly convex (concave) demand curve lies above (below) that of a linear demand curve that is tangential to it at the monopoly output level.

Proof: Substituting the values of the profit expressions (1) to (3) into (8), we get

$$\delta_{AB}^*(1) = \frac{\pi_B^d(c^*)}{\pi_A^m(c) + \pi_B^d(c^*) - \pi_A^p(c)} = \frac{(P^m - c^*)q(P^m)}{P^m - c)q(P^m) + (P^m - c^*)q(P^m) - (c^* - c)q(c^*)} \quad (20)$$

Now consider any strictly convex or concave demand curve, and imagine a hypothetical linear demand curve that is tangential to it at $q(P^m)$. Given the constant marginal cost c , a monopolist's first order condition is

$$(P^m - c)q'(P^m) + q(P^m) = 0$$

By construction, the hypothetical linear demand curve would have the same slope, monopoly output and monopoly price as the given convex or concave demand curve. For any given t in the relevant range, $q(c^*)$ is higher (lower) for a convex (concave) demand function as compared to the hypothetical linear demand function, while all other terms in the expression for $\delta_{AB}^*(1)$ given by (20) are the same. Thus, for any $t \in (0, \bar{t})$ the value of $\delta_{AB}^*(1)$ for the convex (concave) demand function must be greater (less) than that of the hypothetical linear demand function. \square

Lemmas 2 and 3 show that for any price-elastic demand curve, the relationship between $\delta_{AB}^*(1)$ and t is very similar to that depicted for linear demand in Figure 1. The $\delta_{AB}^*(1)$ curve will always be anchored at $\frac{1}{2}$ at the lower limit of the permissible range of t and approach 1 at the upper limit, while for intermediate values it will lie above (below) the benchmark curve if demand is convex (concave).

4.3: Qualifications and extensions

The trade cost paradox holds for general price elastic demand. We have already shown that for the special case of inelastic demand, $\delta_{AB}^*(1) = \frac{1}{2}$ regardless of the level of symmetric

trade costs (provided that they are not high enough to deter trade altogether). Now we turn to some substantive extensions. We relax, first individually and then cumulatively, the assumptions of grim-trigger punishment, symmetric trade costs and one firm per country, providing additional insights into firm behavior in each case while qualifying the trade cost paradox and deriving the competition paradox. These are developed at length in separate sections below, but the consequences of altering the punishment strategy can be discussed very briefly here.

Using linear demand, Lommerud and Sørsgard (2001) showed that the trade cost paradox does not hold with optimal punishments. With discounted payoffs on the punishment path held to zero, reduced trade costs only increase $\pi_B^d(c^*)$, requiring firms to be more patient in order to sustain collusion. This is shown in Figure 1, based on our simulation using linear demand. But it is easy to show that the qualitative features of this relationship remain the same with general demand. From (12), the fact that $\partial\pi_B^d/\partial c^* = -q(P^m) < 0$, and the limiting value of $\pi_B^d(c^*) = 0$ at $t = \bar{t}$, it follows directly that the minimum discount factor with BW punishments decreases monotonically from a value of $1/2$ at $t = 0$ to zero at $t = \bar{t}$, and the paradox does not hold.

5: The trade cost paradox with asymmetric trade costs

Thus far, we have assumed that trade costs are symmetric in both directions. Allowing for asymmetric trade costs permits us to analyze the effects of differences in tariff rates between the two countries and unilateral tariff changes in either country. This introduces an asymmetry in the tension we referred to while motivating the trade cost paradox. From the perspective of country A, an increase in its tariff rate t_A enhances protection of its home firm, assuring a higher Nash profit $\pi_A^p(c)$ if collusion breaks down, making it more likely to defect. Its monopoly profits from abiding by the cartel arrangement, as well as its gain from invading the foreign market, remain unchanged. But the increase in t_A leaves the Nash profits of the foreign firm unchanged while reducing the profit that it can expect by invading country A's market, making it *less* likely to defect. The incentives are reversed for increases in t_B . The consequences for collusion can be easily derived, using the same approach that we used for variations in symmetric trade costs, but applying it separately to the ICCs of the two firms.

For the firm located in country A, an increase in t_A (given t_B) raises $\pi_A^p(c)$ and thereby raises the critical discount factor obtained in (8). As t_A goes from zero to \bar{t} , it can be shown that this critical discount factor, now denoted by $\delta_A^*(1)$, rises from a value of less than $\frac{1}{2}$ to unity. For the firm in country B, interchanging subscripts gives

$$\delta_B^*(1) = \frac{\pi_A^d(c^*)}{\pi_B^m(c) + \pi_A^d(c^*) - \pi_B^p(c)}$$

The increase in t_A reduces $\pi_A^d(c^*)$ and therefore $\delta_B^*(1)$, which falls from a value greater than $\frac{1}{2}$ to zero as t_A rises from zero to \bar{t} . For $t_A = t_B$, symmetry is restored and $\delta_A^* = \delta_B^*$. Figure 2 depicts the relationship between t_A , $\delta_A^*(1)$ and $\delta_B^*(1)$ in a simulation with a linear demand function, setting $t_B = 2$ and other parameter values as in Figure 1. (The $\delta_A^*(2)$ and $\delta_B^*(2)$ curves, depicting the relationship for two firms in each country, should be ignored for the time being.) The curves are drawn for given level of t_B , variations in which can be represented as parametric shifts. When $t_B = 0$ the intercept of both curves on the vertical axis is at $\frac{1}{2}$; increases in t_B lower (raise) the intercept of $\delta_A^*(1)$ ($\delta_B^*(1)$) to zero (unity) as $t_B \rightarrow \bar{t}$. Collusion is sustainable only if the firms discount the future by a factor $\delta \geq \max(\delta_A^*, \delta_B^*)$, represented by the upper envelope of the two curves. This qualifies our trade cost paradox, as unilateral reductions in t_A facilitate collusion only as long as $t_A > t_B$, but further liberalization discourages it.²¹ Another way of looking at this result is that the binding critical discount factor that supports collusion reaches a minimum when $t_A = t_B$. This is worth stating as another proposition:

PROPOSITION 3. Tariff harmonization at any level facilitates SOI cartels.

Tariff harmonization restores the symmetry in the ICCs of the two firms, so this result is related to Proposition 1, which established that symmetry in market size facilitates collusion. However, with BW punishments, a rise in t_A leaves $\delta_A^*(1)$ unaffected, but reduces $\delta_B^*(1)$ as before. This situation can be visualized as a horizontal $\delta_A^*(1)$ curve combined with a downward sloping $\delta_B^*(1)$ curve. The binding minimum discount factor is given by the upper envelope of the two, so we can see that with BW punishments the trade cost paradox does not hold with asymmetric trade costs.

²¹Of course, when $t_B = 0$, both curves originate at $\delta_A^* = \delta_B^* = \frac{1}{2}$ on the vertical axis, and the paradox holds for *any* reduction in $t_A > 0$.

We can conclude from our first paradox that trade liberalization, whether bilateral or unilateral, is not always a substitute for antitrust enforcement. It would be wrong, however, to infer that a multilateral or unilateral increase in tariff rates, by destabilizing the cartel (or deterring its formation) is a substitute for antitrust. First, commitment to a high tariff to destabilize the cartel amounts to replacing a private trade restriction with an official one that increases the home firm's monopoly power, with the usual deadweight loss. Second, there is no rent-shifting from the foreign firm, so the 1980s literature on strategic trade policy is not applicable, although the criticism of that literature in terms of the sensitivity of the optimal tariff to parameters of the model applies here as well. Third, we have shown that the positive relationship between the tariff and the critical value of the discount factor upon which the first paradox is based does not survive with BW punishments. Now we turn to our second paradox where we allow for endogenous entry.

6. Second Paradox: A more competitive market structure facilitates collusion

6.1: The minimum discount factor with multiple firms

To introduce our competition paradox, let us now allow for more firms. Suppose there are $n_j \geq 2$ identical firms in country j who collude by equally sharing monopoly profits in their home market and staying out of the foreign market. A deviating firm can undercut the collusive price slightly and snatch the entire demand in both markets for one period, regardless of the number of firms in the foreign market. Bertrand-Nash reversion to punish defection now results in zero profits in the home market as well as the foreign market: transport costs shelter domestic firms from retaliatory competition from abroad, but not from each other.²² Grim trigger and Bernheim-Whinston punishments (which ensure that discounted profits are zero on the punishment path) therefore give the same ICC. For either punishment strategy, the ICC for a representative firm in country A is now:²³

$$\frac{\pi_A^m(c)/n_A}{1-\delta} \geq \pi_A^m(c) + \pi_B^d(c^*) \quad \text{for } t > 0 \quad (21)$$

which yields

²² Choi and Gerlach (2012) adopt this punishment strategy with two firms, each operating plants in both countries. They, however, do not consider changes in the number of firms.

²³Recall that we are using n_j to mean that there are $n_j \geq 2$ firms in country j ; critical discount factors for the single-firm case continue to be explicitly denoted by $\delta_{AB}^*(1)$.

$$\delta_{AB}^*(n_A) = \hat{\delta}_{AB}(n_A) = \frac{(n_A-1)\pi_A^m(c) + n_A\pi_B^d(c^*)}{n_A[\pi_A^m(c) + \pi_B^d(c^*)]} \geq \frac{1}{2} \quad (22a)$$

with equality only in the case of $n_A = 2$ and $c^* = P^m$ (i.e., $t = \bar{t}$). Substituting the values of the profit expressions from Table 1 into (22), for $t = \bar{t}$ we get the standard result for single-market collusion, $\delta_{AB}^*(n_A) = (n_A - 1)/n_A \geq \frac{1}{2}$. However, a discontinuity arises at the opposite extremity. From (22), $\lim_{t \rightarrow 0} \delta_{AB}^*(n_A) = \frac{2n_A - 1}{2n_A}$. But at $t = 0$ we have an integrated single market with $n_A + n_B$ firms, for which

$$\delta_{AB}^* = \frac{n_A + n_B - 1}{n_A + n_B} \quad \text{for } t = 0 \quad (22b)$$

(22a) and (22b) two coincide for $n_A = n_B = 1$, so this problem did not arise in the international duopoly model we have used so far.²⁴ But with more than one firm in each country, we have to be careful in investigating the effects of changes in the number of firms and trade costs in the vicinity of $t = 0$.

Treating n as a continuous variable,

$$\begin{aligned} \frac{\partial \delta_{AB}^*(n_A)}{\partial n_A} &= \frac{1}{(n_A + n_B)^2} > 0 \text{ for } t = 0 \\ &= \frac{\pi_A^m(c)}{n_A^2[\pi_A^m(c) + \pi_B^d(c^*)]} > 0 \text{ for } t > 0 \end{aligned} \quad (23)$$

Regardless of the level of trade costs, scope for collusion are thus reduced as the number of firms is increased beyond two. As in single-country models, collusion is less attractive when collusive profits have to be shared between more firms while the gains from defection remain unchanged. The ICC for firms in country B is once again obtained by interchanging subscripts A and B. The minimum δ^* that sustains collusion involving firms from both countries must be $\max\{\delta_{AB}^*(n_A), \delta_{AB}^*(n_B)\}$. Therefore, we can state

PROPOSITION 4: With two or more firms in each country, then (a) if the number of firms is unequal to begin with, the range of discount factors that can support collusion is decreasing in the number of firms in the country with more firms; or (b) if the number of firms is equal to begin with, the range of discount factors that can support collusion is decreasing in the

²⁴ In an integrated market, regardless of the number of firms, the allocation of sales between the markets is irrelevant as long as the shares of each country's firms are symmetric in the two markets and the total output is at the monopoly level, shared equally between the two countries.

number of firms in either country alone, or with equal increases in the number of firms in both countries.

The new element introduced by multimarket contact is that the scope for collusion may now depend not just on the number of firms with which one is sharing collusive profits in one's own country, but also on the number who are sharing it in the other country.

Before proceeding to the analysis of competition paradox, let us examine whether our first paradox survives in a scenario with more than one firm in each country. Suppose that $n_A \geq n_B \geq 2$, so that $\delta_{AB}^*(n_A)$ is the binding minimum discount factor. For $t > 0$ we can differentiate (22a) with respect to c^* and substitute $\partial\pi_A^m(c)/\partial c^* = 0$ and $\partial\pi_B^d(c)/\partial c^* = -q(P^m)$ to obtain

$$\frac{\partial\delta_{AB}^*(n_A)}{\partial c^*} = \frac{\partial\hat{\delta}_{AB}(n_A)}{\partial c^*} = \frac{-q(P^m)\pi_A^m(c)}{n_A[\pi_A^m(c) + \pi_B^d(c^*)]} < 0$$

$\delta_{AB}^*(n_B)$ will be similarly inversely related to the level of symmetric trade costs, but by Proposition 4, $\delta_{AB}^*(n_A)$ will continue to bind. Thus, for $t > 0$ the trade cost paradox disappears with $n_j \geq 2$ firms in each country. The intuition is the same as in the case of BW punishments, where discounted profits are zero on the punishment path, and thus insensitive to the level of trade costs. The same is the case here, with more than one domestic firm ensuring a Nash reversion payoff of zero in the home market with grim trigger punishment. The only effect of a reduction in (symmetric) trade costs is to increase the reward to defection in the foreign market, requiring a higher discount factor to restore the viability of collusion.

However, an anomaly now arises at $t = 0$. From 22(a), $\lim_{t \rightarrow 0} \delta_{AB}^*(n_A) = \frac{2n_A - 1}{2n_A}$, which can be shown to be greater than (equal to) $\delta_{AB}^*(t = 0) = \frac{n_A + n_B - 1}{n_A + n_B}$ as n_A is greater than (equal to) n_B . Thus, if $n_A > n_B$, a reduction of trade costs from a very low level to zero *reduces* the minimum discount factor required to sustain collusion, so the trade cost paradox holds locally. The reason is that a firm's share of collusive profits π_A^m/n_A under SOI is lower than $(\pi_A^m + \pi_B^m)/(n_A + n_B)$ in an integrated market for $n_A \geq n_B$, while deviation payoffs are

$\pi_A^m + \pi_B^m$ in both scenarios. Higher collusive profits under market integration allow collusion to be sustained at lower discount factors.²⁵

We now consider unilateral changes in country A's tariff with more than one firm in each country. The various profit expressions (1) to (3) that figure in a firm's ICC, and therefore the critical discount factor, are independent of the number of firms in the other country. We can therefore work with (22a) alternately for country A and B firms. Variations in t_A have no effect on the critical discount factor of country A firms, as once again protection of the domestic market has no effect on Nash profits $\pi_A^p(c)$, which remain zero. The critical discount factor for country B firms is decreasing in t_A , as in the one-firm case. Figure 2 illustrates the relevant curves for two firms in each country, based on our simulation with linear demand.

In order to avoid cluttering the diagram, we do not draw the curves for more than two firms, but the direction of the parametric shift is still given by (23), as in the case of symmetric trade costs, with the same intuition. Interchanging subscripts gives a similar expression for firms in country B. In Figure 2, increasing the number of firms would therefore result in successive vertical upward shifts of the curves representing $\delta_A^*(n_A)$ and $\delta_B^*(n_B)$ as the case may be, while they remain respectively horizontal and downward sloping. The binding discount factor that sustains collusion is the upper envelope of the two, as before.

The forgoing analysis implies that with two or more firms in country A and one or more in country B, we again get a horizontal $\delta_A^*(n_A)$ curve and a downward sloping $\delta_B^*(n_B)$ curve. Thus, the trade cost paradox does not hold with asymmetric trade costs and more firms. Once again, the parallel with BW punishments is obvious. However, with multiple firms the trade cost paradox is further limited to a situation in which n_A is not too large relative to n_B . Otherwise, the horizontal $\delta_A^*(n_A)$ curve may lie everywhere above the $\delta_B^*(n_B)$ curve, making $\delta_A^*(n_A)$ the binding critical discount factor.²⁶ Variations in t_A then have no effect on the likelihood of collusion, and trade cost paradox no longer holds.

²⁵ See Bond and Syropoulos (2008, 2012) for a similar discontinuity result with quantity setting firms, but with a different underlying logic based on differences in deviation and punishment payoffs.

²⁶ Recalling our earlier results, a reduction in t_B relative to t_A can also bring about this outcome.

6.2: The competition paradox with symmetric trade costs

Returning to the case of symmetric trade costs, we now derive our second paradox, again beginning with the expressions for the benchmark case of linear demand. For $n \geq 2$, substitution of the linear demand parameters into (22) gives

$$\delta_{AB}^*(n) = \frac{(2n-1)(a-c)-2nt}{2n(a-c-t)} \quad \text{for } t > 0 \quad (24)$$

which decreases monotonically from a limiting value of $(2n-1)/2n$ as $t \rightarrow 0$ to $(n-1)/n$ at $t = \bar{t}$, as in the general case. The same simulation values give us the family of curves in Figure 1 for various values of n , applicable to both countries. The qualitative features of these curves hold for general demand as well. We have already proved above that with $n_j \geq 2$, $\partial \delta_{AB}^*(n)/\partial t < 0$. We therefore get a family of downward-sloping curves in (t, δ) space, parameterized by n , as in Figure 1. Recalling that $\partial \delta_{AB}^*(n_A)/\partial n_A > 0$, the lowest such curve corresponds to $n = 2$, for which the limiting values of $\delta_{AB}^*(2)$ are $3/4$ and $1/2$ respectively. The discontinuity at $t = 0$ makes no difference for $n_A = n_B = 2$, for which $\delta_{AB}^*(2) = \frac{n_A+n_B-1}{n_A+n_B} = 3/4$.

Thus, as t varies between zero and \bar{t} , $\delta_{AB}^*(2)$ decreases monotonically from $3/4$ to a limiting value of $1/2$, while we already showed in Lemma 2 that $\delta_{AB}^*(1)$ increases from $1/2$ to a limiting value of unity. Based only on their monotonicity and their values at the boundaries of t , the curves for $\delta_{AB}^*(1)$ and $\delta_{AB}^*(2)$ must intersect at some level of trade costs \tilde{t} , so for a range of trade costs $t > \tilde{t}$, we must have $\delta_{AB}^*(1) > \delta_{AB}^*(2)$. As we have assumed that the countries are identical, exactly the same expressions can be derived from the perspective of country B firms. Therefore, for all $t > \tilde{t}$, an increase in the number of firms from one to two in each country reduces the critical value of the discount factor required to sustain collusion with SOI, making it *more* likely. Note that this result applies only to symmetric increases in the number of firms, because if we add a firm in only one country, the higher $\delta_{AB}^*(1)$ of the other country remains the binding constraint.

What if we increase the number of firms beyond two in each country? We already know from (23) that $\delta_{AB}^*(n)$ is increasing in n beyond $n = 2$. After substituting the constituent profit expressions from Table 1 into (8) and (22)—and ignoring the obvious integer constraint on n —we can solve $\delta_{AB}^*(1) = \delta_{AB}^*(n)$ for \hat{n} , the number of firms required for $\delta_{AB}^*(n)$ to return to the level where $n = 1$:

$$\frac{\pi_A^m[\pi_A^m(c) + \pi_B^d(c^*) - \pi_A^p(c)]}{\pi_A^m[\pi_A^m(c) + \pi_B^d(c^*) - \pi_A^p(c)] - \pi_B^d \pi_A^p} \equiv \hat{n} > 1 \quad (25)$$

We have thus established that $\delta_{AB}^*(n)$ is non-monotonic, falling as n goes from 1 to 2 and then rising again, exceeding its original level if $n > \hat{n}$.

The linear demand case provides further insights, which we will then extend to general demand. As the number of firms increases beyond two, each successive $\delta_{AB}^*(n)$ curve in Figure 1 will intersect the upward sloping $\delta_{AB}^*(1)$ curve at some higher level of t , which we call $\tilde{t}(n)$. For all $t > \tilde{t}(n)$, $\delta_{AB}^*(1) > \delta_{AB}^*(n)$. (For example, $\delta_{AB}^*(1) > \delta_{AB}^*(2)$ for $t > 2$, as illustrated.) An explicit expression for $\tilde{t}(n)$ can be obtained in the linear case: from (16) and (24), for given n , $\delta_{AB}^*(1) > \delta_{AB}^*(n)$ if and only if $t > \frac{(n-1)(a-c)}{2n} = \frac{(n-1)}{n} \bar{t} \equiv \tilde{t}(n)$. Clearly, $\tilde{t}(n)$ is increasing in n . So given any n , we can have $\delta_{AB}^*(1) > \delta_{AB}^*(n)$ for high enough trade costs $t > \tilde{t}(n)$.

Although Figure 1 is drawn for linear demand, these results can be generalized. Proposition 2 established that $\delta_{AB}^*(1)$ is strictly increasing in t for all price-elastic demand functions satisfying our minimal assumptions. We also showed in Lemma 2 that $\delta_{AB}^*(1) = \frac{1}{2}$ at $t = 0$, and $\delta_{AB}^*(1) \rightarrow 1$ as $t \rightarrow \bar{t}$. The upward-sloping curve representing $\delta_{AB}^*(1)$ will therefore necessarily intersect any downward-sloping $\delta_{AB}^*(n)$ curve at some $\tilde{t}(n) \in (0, \bar{t})$.²⁷ So for *any* $n > 1$, we must have $\delta_{AB}^*(n) < \delta_{AB}^*(1)$ for some range of $t > \tilde{t}(n)$. The same applies to the other country, so if $\max(n_A, n_B) < \hat{n}$, then collusion will be sustainable for firms' discount factors $\delta_{AB}^*(1) > \delta \geq \max[\delta_{AB}^*(n_A), \delta_{AB}^*(n_B)]$. For example, for $t = 3$ in the simulation, Figure 1 shows that $\hat{n} = 4$ and $\delta_{AB}^*(2) < \delta_{AB}^*(3) < \delta_{AB}^*(4) = \delta_{AB}^*(1)$, so collusion with two or three firms in each country is sustainable with lower discount factors than with one firm in each.

Applying Lemma 3, $\tilde{t}(n)$ is lower (higher) for strictly convex (strictly concave) as compared to linear demand. However from (22a) it is clear that $\delta_{AB}^*(n)$ is unaffected by the demand curvature since the expression does not contain any quantity other than the monopoly one. We can now state our paradoxical proposition:

PROPOSITION 5 (Competition paradox): There exists a level of symmetric trade costs $\tilde{t}(n) < \bar{t}$ above which, compared to a situation with one firm in each country, any symmetric

²⁷ As we have shown that $\delta_{AB}^*(n) \geq \frac{3}{4}$ at $t = 0$, the discontinuity that we noted above makes no difference to this result.

or asymmetric increase up to n firms in both countries reduces the critical discount factor and makes collusion with SOI more likely. The range of trade costs for which this holds is higher (lower) for strictly convex (concave) as compared to a linear demand.

The intuition for this paradox is as follows. Any increase in the number of firms in a country requires collusive profits in the home market to be shared among a larger number of firms, while leaving the gains from invading the foreign market unchanged. By itself, this would continuously increase the critical $\delta_{AB}^*(n)$, making collusion less likely. But the increase from one firm to two discontinuously reduces Nash reversion profits from $(c^* - c)q(c^*)$ for a single home firm, to zero for two home firms. This causes a discontinuous reduction in δ_{AB}^* , which is larger the higher is c^* . Essentially, the severity of the Nash reversion punishment increases discontinuously while going from one to two firms in a country, and at higher (lower) levels of c^* this more than offsets (does not offset) the disincentive to collude that arises from having to share monopoly profits. Thus, as n increases from one to two, δ_{AB}^* falls (rises) at higher (lower) levels of t , and the $\delta_{AB}^*(1)$ curve in Figure 1 rotates clockwise to become the $\delta_{AB}^*(2)$ curve. As n increases further, the profit-sharing effect continues to operate, while Nash profits remain zero, and the successive $\delta_{AB}^*(n)$ curves shift parametrically upwards, shrinking the parameter space for which the paradox holds.

We have demonstrated that the competition paradox holds for general demand only if $\delta_{AB}^*(1)$ is increasing in t , that is, only if our earlier trade cost paradox holds. Therefore, the same conditions that cause the latter to fail will also cause the former to fail. In particular, as shown above, the trade cost paradox does not hold with Bernheim-Whinston punishments, for which $\hat{\delta}_{AB}(1)$ decreases monotonically from a value of $\frac{1}{2}$ at $t = 0$ to zero at $t = \bar{t}$. But with two or more firms in each country, we showed that $\hat{\delta}_{AB}(n)$ decreases monotonically to a value greater than $\frac{1}{2}$ at \bar{t} . Thus $\hat{\delta}_{AB}(1) < \frac{1}{2} < \hat{\delta}_{AB}(n)$ for all t , and the relevant curves never cross. (This is illustrated in Figure 1 for the simulation with linear demand, but we have just shown that the qualitative features of the curves are unchanged with general demand.) So the competition paradox does not hold with Bernheim-Whinston punishments. Nor does it hold with inelastic unit demand, for which $\delta_{AB}^*(1) = \frac{1}{2}$ regardless of trade costs. With this demand specification, it is straightforward to show that $\delta_{AB}^*(n)$ has the same general characteristics as

in the case of price-elastic demand.²⁸ Thus, $\delta_{AB}^*(1) < \delta_{AB}^*(n)$ for all t within the permissible range; again in terms of Figure 1, the curves never cross.

6.3: The competition paradox with asymmetric trade costs

In the case of asymmetric trade costs, the tipping point for the relative strengths of the profit-sharing and punishment-enhancing effects of an increase in the number of firms occurs where trade costs are symmetric. As we can see from Figure 2, for $t_A > (<) t_B$, $\delta_A^*(2) < (>) \delta_A^*(1)$. The same logic applies to the country B firm(s): at high levels of t_B the punishment-enhancing effect of adding a firm in country B, reducing δ_B^* , outweighs its profit-sharing effect. In Figure 2, this corresponds to relatively low levels of t_A to the left of the symmetry point. So for $t_A > (<) t_B$, $\delta_B^*(2) > (<) \delta_B^*(1)$, and thus raising n_B from one to two causes the δ_B^* curve to rotate anti-clockwise around the symmetry point. Further increases in the number of firms in either country shift the corresponding curve parametrically upwards, as in the case of symmetric trade costs.

These results imply that the competition paradox is reinforced if we consider asymmetric trade costs. From Figure 2, going from one firm to two in country A (B) now reduces the critical discount factor (given by the upper envelope of the relevant curves for country A and B firms) when $t_A > (<) t_B$. (Although Figure 2 was drawn for linear demand, we showed above that the qualitative features of these curves would be the same for general demand.) Thus, compared to Proposition 5, the paradox now holds for an increase from one to two firms in only one country.

Adding more firms in country A raises the horizontal $\delta_A^*(n_A)$ curve vertically, so it must intersect the upward-sloping curve for $\delta_A^*(1)$ at some $\tilde{t}_A(n_A)$. Consequently, the analysis that we conducted in the case of symmetric trade costs can be applied here as well. For any given n_A , this intersection gives the $\tilde{t}_A(n_A)$ above which $\delta_A^*(n_A) < \delta_A^*(1)$, as in the case of symmetric trade costs. And for any given t_A , the highest integer n_A such that $\tilde{t}_A(n_A) \leq t_A$ gives \hat{n}_A , the highest number of firms for which $\delta_A^*(n_A) \leq \delta_A^*(1)$. We can now state

²⁸ With choke price \bar{P} , substituting the profit expressions for inelastic demand into the formula for $\delta_{AB}^*(n)$ gives $\delta_{AB}^*(n) = \frac{(2n-1)(\bar{P}-c)-nt}{2n(\bar{P}-c)-nt}$, which decreases from $\frac{(2n-1)}{2n}$ to $(n-1)/n$ as t rises from zero to $\bar{t} = \bar{P} - c$.

PROPOSITION 5A: If import tariffs are asymmetric, then an increase in the number of firms from one in the country with the higher tariff makes collusion with SOI more likely.

7. Endogenous entry, mergers and implications for antitrust policy

Strictly speaking, Proposition 5 and 5A are comparative-static results that apply to exogenous variations in the number of firms. If we endogenize firms' entry decisions, we get some interesting extensions with both positive and normative implications. We work these out for the case of symmetric trade costs. We have shown that considering each country separately, the critical discount factor falls as we go from one firm to two, then rises to the original level with \hat{n} firms. Compactly stated, $\delta_{AB}^*(\hat{n}) = \delta_{AB}^*(1) > \delta_{AB}^*(n) \geq \delta_{AB}^*(2)$. Now let us assume that $n_B = 1 < n_A < \hat{n}$ and that given the level of symmetric trade costs t , the firms' discount factor δ is such that

$$\text{A5: } \delta_{AB}^*(1) > \delta > \delta_{AB}^*(n_A) \geq \delta_{AB}^*(2) > 1/2.$$

No SOI cartel with monopoly pricing would be sustainable under A5, because the country B monopolist would not be willing to participate. However, it would then be constrained to charge a price $c^* = c + t$ by the threat of imports from its rivals in country A.

Under A5, an entrant in market B will anticipate that its entry will reduce the critical discount factor to $\max\{\delta_{AB}^*(2), \delta_{AB}^*(n_A)\}$ and permit a cartel with SOI, in which it will share monopoly profits in its home market and cover its entry costs F , provided that $F < \pi_B^m(c)/2(1 - \delta)$. More entrants will raise $\delta_{AB}^*(n_B)$, but they can be accommodated in the cartel as long as $\delta \geq \max\{\delta_{AB}^*(n_A), \delta_{AB}^*(n_B)\}$, and consequently they will enter as long as $F < \pi_B^m(c)/n_B(1 - \delta)$. The prospect of a stable cartel can therefore encourage one or more firms to enter a market in which a domestic monopolist is competing with two or more foreign firms. Recalling that our result $\delta_{AB}^*(n) < \delta_{AB}^*(1)$ holds for $t > \tilde{t}(n)$, where $\tilde{t}(n)$ is increasing in n , we obtain the following:

Corollary to Proposition 5: When trade costs are relatively high and entry costs relatively low, SOI cartels are likely to involve more than one firm in each country.

Can entry benefit an incumbent monopolist, forced to limit price by the threat of imports, by facilitating a stable SOI cartel? Unlikely: at high levels of trade costs its profit

$\pi_A^p(c)$ as a monopolist constrained to limit price at c^* approaches the unconstrained monopoly level $\pi_A^M(c)$, and will therefore exceed the shared collusive profit $\pi_A^M(c)/n_A$ that it can anticipate from a cartel.

From a normative point of view, entry in order to facilitate collusion inflicts a double blow on social welfare: there is wastage of setup costs in addition to the rise in price that occurs despite (actually, because of) entry. Rational entrants will not enter if their entry raises $\delta_{AB}^*(n)$ above their discount factor, destabilizing the cartel, so entry will never be pro-competitive. This result illustrates in a new context two old themes in the literature on trade policy under imperfect competition: tariff-induced excessive entry, and trade restrictions as facilitating practices.²⁹

From an antitrust perspective, our competition paradox provides one more example of the breakdown of the simplistic relationship between market structure and firm conduct that motivated traditional industrial organization and antitrust. But even after that paradigm has fallen into disuse, competition agencies continue to use the number of firms in a market as a screen for selecting industries on which to focus their investigations to uncover cartel behaviour. The paradox shows that, at least in one context, this could be misleading: having more firms could make collusion *more* likely.

We can also apply our findings to a *reduction* in the number of firms resulting from a merger, to get the following:

COROLLARY 2: Starting with $1 < n < \hat{n}$ firms in each country, a merger to monopoly in either country raises the critical value of the discount factor to $\delta_{AB}^(1)$, making collusion with SOI less likely.*

Again, endogenizing firms' merger decisions yields additional complications for antitrust policy. A merger to monopoly is usually viewed with suspicion, but our result provides a different perspective. If an SOI cartel with monopoly pricing was already in place, then a merger of all the firms within a country will not affect the equilibrium price. Absent any efficiency gains, it will at best give the firms the same payoff (shared monopoly profits in the home market) and at worst destabilize the cartel by raising the critical δ_{AB}^* , forcing the newly-created monopolist to charge the limit price c^* rather than the monopoly price. If such a merger takes place, therefore, there must be efficiency gains; we assume that they take the

²⁹We distinguished our results from those of the earlier literature in the Introduction to this paper.

form of elimination of duplicated fixed costs.³⁰ With no change in the equilibrium price, if collusion is sustained post-merger, this cost saving must result in an increase in social welfare. This complicates antitrust authorities' analysis of the so-called "unilateral effects" resulting directly from the increased concentration caused by a merger. The firms would not want to disclose their cartel activities to the authorities reviewing the merger, but their efficiency defense may not convince an agency that quite reasonably—but in this case incorrectly—anticipates an increase in prices resulting from a merger to monopoly. Potentially welfare-improving mergers may therefore be disallowed.

On the other hand, if no pre-merger cartel existed because the firms' discount factors were too low relative to at least one country's $\delta_{AB}^*(n)$, a merger to monopoly in either country cannot change the situation. This result complicates the task of antitrust agencies in assessing the "coordinated effects" (known in Europe as "collective dominance effects") of a merger: the greater likelihood of collusion arising from a reduction in the number of firms, especially when the merger reduces market share asymmetries. In this case, the agencies may needlessly fear that the domestic monopolist that will be created by the merger will be more likely to collude with a foreign rival.³¹

8. Conclusions

Several international cartel cases have been investigated and prosecuted by different competition agencies around the world. Firms from different countries form cartels and stay away from each other's market in order to make higher profits. International cartels based on such territorial allocation have been a concern for policy makers for long. In this paper we have provided a theoretical analysis of such cartels.

³⁰A reduction in the variable costs of only country A firms would reduce the price and increase consumer surplus, but introducing cost asymmetry might upset the cartel, and will take us too far from our current focus.

³¹ Davies et al (2011) econometrically recovered the decisional practice of the European Commission in all merger cases since 1990 (when European merger regulations came into effect) in which coordinated effects were considered while other conditions inhibiting collusion (such as ease of entry, countervailing buyer power, non-transparent pricing) were not present. They found that the Commission almost always associated coordinated effects with mergers resulting in near-symmetric duopolies—the outcome analyzed in our model when there is one foreign firm competing with the merging home country duopolists.

We showed, based on actual cases, that such cartels not only involve more than one firm in each territory but are also largely found in industries producing homogenous products. We have examined the sustainability of collusion in an environment where price-setting oligopolists located in different markets separated by trade costs produce homogenous goods and sustain collusion based on territorial allocation of markets.

Using a standard supergame model and a general demand function, we have generalized the existing result that a reduction in trade costs can paradoxically increase the scope for collusion. We have shown that this trade cost paradox holds for symmetric as well as asymmetric trade costs. Our second paradox showed how an increase in number of firms might lead to greater scope for collusion. These paradoxes pose serious challenges to policy makers as they are contrary to conventional wisdom and thus require an understanding of the economic environment before implementing the standard policies of promoting competition. In future work, we intend to explore the effect of entry in the form of foreign direct investment on cartel stability, and whether national and international antitrust enforcement can play any role in inhibiting such cartel behavior.

TABLE 1

Constituent profit terms in the incentive-compatibility conditions for collusion by firms in country j

Profit term	Definition	Value at $c^* = c$ (or $t = 0$)	Derivative w.r.t. c^* (or t)	Value at $c^* = P^m$ (or $t = \bar{t}$)
$\pi_j^m(c)$	$(P^m - c)q(P^m)$	$\pi_j^m(c)$	0	$\pi_j^m(c)$
$\pi_l^d(c^*)$	$(P^m - c^*)q(P^m)$	$\pi_j^m(c)$	$-q(P^m)$	0
$\pi_j^p(c)$	$(c^* - c)q(c^*)$	0	$q(c^*)$ $+ q'(c^*)(c^* - c)$	$\pi_j^m(c)$

References

- Abreu, D. (1986), "Extremal Equilibria of Oligopolistic Supergames", *Journal of Economic Theory* 39: 191-225.
- Akinbosoye, O., E.W. Bond and C. Syropoulos (2012), "On the Stability of Multimarket Collusion in Price-setting Supergames", *International Journal of Industrial Organization* 30: 253-264.
- Bernheim, D. and M. Whinston (1990), "Multimarket Contact and Collusive Behavior", *Rand Journal of Economics* 21(1): 1-26.
- Bhattacharjea, A. (1995), "Strategic Tariffs and Endogenous Market Structures: Trade and Industrial Policies under Imperfect Competition", *Journal of Development Economics*, 47: 287-312.
- Bolotova, Y.V. (2009), "Cartel Overcharges: An Empirical Analysis", *Journal of Economic Behavior & Organization* 70: 321-341.
- Bond, E.W. and C. Syropoulos (2008), "Trade Costs and Multimarket Collusion", *RAND Journal of Economics* 39(4): 1080-1104.
- Bond, E.W. and C. Syropoulos (2012), "Economic Integration and the Sustainability of Mutimarket Collusion", *Economics Letters* 117: 42-44.
- Choi, J.P. and H. Gerlach (2012), "International Antitrust Enforcement and Multimarket Contact", *International Economic Review* 53(2): 635-657.
- Colombo, Luca and Paola Labrecciosa (2007), "Sustaining Collusion under Economic Integration", *Review of International Economics* 15(5): 905-915.
- Connor, J.M. (2007), *Global Price Fixing*. Berlin and Heidelberg: Springer-Verlag.
- Davidson (1984), "Cartel Stability and Tariff Policy", *Journal of International Economics* 17: 219-37.
- Davies, S., M. Olczak and H. Coles (2011), "Tacit Collusion, Firm Asymmetries and Numbers: Evidence from EC merger cases", *International Journal of Industrial Organization* 29: 221-31.
- De, O. (2011), "The Internal Structures and Organisation of EC Prosecuted Cartels and the Impact on their Performance", PhD dissertation, University of East Anglia.
- Fung, K.C. (1991), "Collusive Intra-Industry Trade", *Canadian Journal of Economics* 24(2): 391-404.
- Gross, J. and W.L. Holahan (2003), "Credible Collusion in Spatially Separated Markets", *International Economic Review* 44(1): 299-311.

- Harrington, J.E. (2006), "How Do Cartels Operate?", *Foundations and Trends in Microeconomics*, 2(1): 1–105.
- Levenstein, M. and V.Y. Suslow (2011), "Breaking Up is Hard to Do: Determinants of Cartel Duration", *Journal of Law and Economics*, 54(2): 455-492
- Lommerud, K.E. and L. Sørgaard (2001), "Trade Liberalization and Cartel Stability", *Review of International Economics* 9(2): 343-55.
- Markusen, J.R. and A.J Venables (1988), "Trade Policy with Increasing Returns and Imperfect Competition: Contradictory Results from Competing Assumptions", *Journal of International Economics* 24(3-4): 299-316.
- Miklos-Thal, J. (2011), "Optimal Collusion under Cost Asymmetry", *Economic Theory* 46: 99-125.
- Notz, William (1920), "International Private Agreements in the Form of Cartels, Syndicates, and other Combinations", *Journal of Political Economy*, 28(8): 658-679.
- Pinto, B. (1986), "Repeated Games and the 'Reciprocal Dumping' Model of Trade", *Journal of International Economics* 20: 357-66.
- Rotemberg, J.J. and G. Saloner (1989), "Tariffs vs Quotas with Implicit Collusion", *Canadian Journal of Economics*, 22(2): 237-44.
- Scherer, F.M. (1994), *Competition Policies for an Integrated World Economy*. Washington, D.C.: Brookings Institution.
- Schröder, Philipp J. H. (2007), "Cartel Stability and Economic Integration", *Review of International Economics*, 15(2), 313–320.
- Suslow, V.Y. (2005), "Cartel Contract Duration: Empirical Evidence from Inter-war International Cartels", *Industrial and Corporate Change* 14: 705-44.



