International Coordination through Trigger Punishment:
Fighting Corruption through the OECD Anti-Bribery Convention
Gaye B. Muderrisoglu ¹

The model is a variation on Morrow (1994) examining the incentive structures arising in international coordination questions which involve distribution and information problems. The main dilemma is the inability of any international organization of solving both problems – where either the distribution or the information problem prevents the other from being resolved completely. The application of this model of collusion to the OECD Anti-Bribery Convention shows how the threat of punishment by all signatories through defection in the second round leads to greater coordination on the part of all states in the first round. This is a model involving the two period repetition of the stage game, with the second being a punishment stage conditional on the public signal at the end of the first stage. The Convention meant to address the supply-side of corruption in developing countries through a formalized anti-bribery cartel. The model is a refinement of the Folk Theorem on Trigger Punishment, where the Anti-Bribery Convention does lead to honest revelation of information when communication occurs and to more cooperation via leadership.

International organizations provide formal platforms for inter-state cooperation, but states also collaborate on many issues without having to resort to the creation of a formal regime. The collusion of OECD countries with the Anti-Bribery Convention is an example of such cooperation. The convention provides an example of international coordination in the face of information and distribution problems. How the public signal works in the framework of the Anti-Bribery Convention is as a mechanism which allows for the honest revelation of information, and in this way allowing optimal cooperation among states that make up 68% of world’s foreign investment.² Several aspects of the anti-bribery convention then resemble a cartel, where collusion by member states of the

¹ Department of Political Science, University of Michigan, Ann Arbor, gayem@umich.edu. Prepared for presentation at the MPSA Conference in Chicago, April 20-23, 2006.
Convention is thought to control the market for bribes in international business interactions.

In addition to the red tape and bureaucracy encountered by corporations doing business in developing countries, the use of bribery to secure contracts or as a side-payment to ensure processing of applications for licenses, is a common occurrence. The idea for a cartel arose with increasing demands for bribery in the 80s that followed major corruption scandals, where bribes had reached millions of dollars in some infrastructure or military contracts abroad.³ The US government’s took the initiative after passing the US Foreign Corrupt Practices Act and then introduced it to the OECD states which passed several legislative measures against corruption by 1994. The study examines the nature of this cartel against bribery.

The members of the Convention are on the supply-side of bribes and the cartel works to decrease the decline in benefits of doing business where corruption is rampant. Morrow (1994) on the abstract explanation of the twin problems of distributive and information problems which international institutions are meant to remedy lays the foundation for this paper. The paper also follows the ideas of series of papers in economics that look at collusive solutions to non-cooperative nature of cartels. The main aspects of the model are trigger punishment strategies that operate in two time-periods with decline in demand for doing business without bribery as indicator of cheating.

³ http://www.cipe.org/programs/corruption/corruptiontoolkit.htm
On the Economics of Cartels

The idea of punishment in alternating periods by players in a game was first examined by Friedman (1971) who determined that by limiting options available to all actors at once, a non-cooperative equilibrium can be reached. As such, by simultaneously punishing all individual deviations from NE, his folk theorem notes that there is a set of non-cooperative agreements that can be had through repetition where everyone is essentially minimaxing from fear of having to minimize in a number of future periods. Porter (1983) and Green and Porter (1984) take this line of research one step further to show that (1) there is a switch from collusive to non-cooperative equilibria between periods via empirical data from an actual cartel (JEC, the freight railroad cartel that existed in the US between 1880 and 1886), and (2) that allotment of supply did essentially lead to collusive behavior.

Porter (1983) finds that members will have decreased incentives to cheat the longer the punishment period or lower the trigger price. What Abreu, Pearce and Stachetti (1985) – hereafter ABS - do is relax the assumptions by Porter (1983) for requirements of evidence of cheating in previous periods, and reduce the repeated game to an infinitely repeated game with symmetric sequential payoffs. The study uses Porter’s use of simultaneous revelation of information, and the concept of repeated game from the ABS model. Later, Abreu, Pearce and Stachetti (1991) follow up paper looking at these within the framework of imperfect monitoring in a repeated game set up.
The relevance of this literature to the model is that cartels often operate using punishment by all members if violations have been revealed through the public signal of the cartel’s product, namely the price. Radner (1985) also puts forth some review strategies that go into issues of disclosure and monitoring, which included Friedman’s random review but accounting for errors in monitoring. We also know that when distribution problems overwhelm coordination problems, there will be a need for a separate monitoring system. This model does not look at that part of the organization, but highlights a need for monitoring of the distributional issues not covered by communication through the convention.4

International Cooperation and Coordination Problems

Snidal (1985) highlights the richness of using 2 by 2 games to analyze international cooperation as ‘collective action problems applied to the international system,’ as it related to the international regime debates at the time.5 Koremenos, Lipset and Snidal (2001) later launched the Rational Design project to highlight the benefits to studying international organizations as solutions to particular problems between states. The research agenda has spun a series of design proposals that relate international organizations to various cooperation problems.6


5 Snidal points out the existence of 72 possible combinations of 2 x 2 games which vary by the preference ordering of the actors as demonstrated by Rapoport and Guyer in 1966.

Martin (1992) suggests that crises may arise in coordination problems that did not exist in the long run via changes in the external environment, while challenges may be costly in the long run. Yoffie and Gomes-Casseres (1994) note that coordination problems can often be overwhelmed by distribution problems, causing a breakdown in the institutional solution. This paper examines how the self-enforcing nature of game-theory may explain how a punishment strategy used by most cartels can lead to changes in both the two-way and one-way communicative equilibriums outlined in the original set up. The question posed by this model is to ask exactly where the trigger punishment strategy may lead to break down in the collusive agreement. Ayres (1992) emphasizes that in cartels, firms must be able (1) to reach an agreement, (2) to detect breaches of the agreement, and (3) to punish firms that breach to be able to qualify as successful explicit agreements to collude. Roberts (1985) separates the issues of moral hazard and adverse selection as two separate sides of unsuccessful cartels, where moral hazard applies more to the cheating and detection of cheating via trigger strategies and adverse selection assumes there is no error in detection of cheating instead focuses on incentive structures that most efficiently get individual actors to reveal their private information. My model precludes both adverse selection and moral hazard issues, but the focus of this paper is showing that the additional distributional concerns arising with the existence of the Anti-Bribery Convention could prevent coordination where it does not facilitate more honest communication. Honest communication is provided for in cases of communication between sides and communicative signaling from the leader. In the latter case, the leader makes the follower reveal more private information.
History of the Agreement

Collusion was occurring in the bribery market prior to the Convention, because of the widespread belief that providing certain gifts or bribes to certain public officials to be able to obtain contracts was unethical. This idea was catching on even though gift giving for political favors is thought to be standard practice in most developing nations. So although certain stronger investors knew there should be curbs put on bribery, there was no serious collusion against bribery prior to this agreement because it was not thought that abstention through collusion would be enough to prevent the cheaters from continuing to bribe. It wasn’t until the market for bribes got out of hand and reputations were at stake that collusion was carried to a formal agreement.

After the FCPA determined that many companies were in violation of the code in their dealings abroad that the U.S. moved the issue more firmly into the OECD arena. A secondary objective was to improve the existing accounting standards because in the 1970s many companies falsified their books to create ‘slush funds’ used to bribe foreign officials. This concern by U.S. officials led Congress to empower the Executive to begin negotiations in the OECD (Organization of Economic Cooperation and Development), especially with the major trading partners of the U.S. to get them to pass legislation that made bribery of foreign officials illegal. In 1976, the United States and the other members of the OECD signed the non-binding Declaration on International Investment and Multinational Enterprises, which denounced companies that gave bribes to public officials and officials that sought bribes from companies or graft. In 1977 the ICC
(International Chamber of Commerce) adopted a code to fight extortion and bribery in business.\(^7\)

The U.S. Unlawful Corporate Payments Act of 1977 notes that it was thought rather than have the firm disclose bribery (due to the adverse selection effects of the tradeoff between using bribes to get multimillion dollar contracts) that the procedure of bribery of foreign officials be criminalized. The object was to regain the confidence of the public in American business, political parties and public officials in the post-Watergate era. This way, The Subcommittee on Consumer Protection and Finance of the 94\(^{th}\) and 95\(^{th}\) House determined that the disclosure of domestic bribery which was illegal at the time would be complemented by the criminalization of bribery by US firms abroad. The later Senate version or the Foreign Corrupt Practices Act (FCPA) of 1977 made it illegal to give bribes to foreign officials ‘for the purposes of obtaining and keeping business.’ Until the 1970s, it is thought that up to a quarter of Fortune 500 companies operating abroad used bribery to do business.

This law was later amended in 1988 to make exceptions that were legal under foreign law and to ignore unintentional accounting mistakes that were thought to be disadvantaging the US companies against their competitors around the world and especially disproportionately hurting US investments due to the sheer quantity of US business abroad, relative to firms from states where bribery of foreign officials was not criminalized.\(^8\) Prior to this amendment the territorial identification of the definition of

\(^7\) http://www.oecd.org/department/
\(^8\) http://www.usdoj.gov/criminal/fraud/fcpa/1988hist.htm
foreign companies and nationals as the subjects of bribery was not recognized, only
foreign companies that were authorized to act in a business transaction were faulted. In
addition, the amendment made subsidiaries of U.S. companies, U.S. citizens or residents
liable and brought them under the jurisdiction of this law.\textsuperscript{9}

The UN Anti-Bribery Convention (in addition to the OAS talks) is currently under
negotiation to make the measures more comprehensive to include states where business
transactions take place. Rose-Ackerman (2002) puts forth an ethnical argument for why
bribery abroad is not in the interests of the MNCs, saying that it can lead to inefficiencies
in the market. Moreover, she says MNCs have an interest in maintaining stability and
legitimacy of governments in states they operate. Her explanations regarding
diseconomies of bribing in business transactions abroad makes a convincing case for
preventing bribery in international business. The Anti-Bribery Convention is a step in the
same direction, aiming to bring under control the supply-side measures against bribery. It
is also the case that placing ethical and legal obstacles to suppliers of bribes is key to
fighting corruption in the developing world. The legal dimension is especially important
since supplier country cooperation with international criminal investigations of bribery in
the contracting state is conditional on whether or not the supplier state has made bribery
abroad a criminal offense.\textsuperscript{10}

\textsuperscript{9} Abbott, Kenneth and Duncan Snidal provide a history of the agreement in their APSA 2002 presentation.
\textsuperscript{10} Rose-Ackerman, Susan, ‘Grand Corruption and the Ethics of Global Business,’ Journal of Banking and Finance,
Vol. 26 (9), (Sept 2002), 1889-1918.
Combination of distribution and information problems

The distributive and information problems are not going to be solved entirely by the institutionalization of collusion. The distributive conflict as modeled through the BotS (Battle of the Sexes) set up and the BPMA (Both Prefer Move A) and BPMB (Both Prefer Move B) alternatives are a close encounter with the distributive conflict that arise with the existence of the Convention itself. According to the original model, it is the relative relevance of the distribution vs. the information problem that determines the appropriate equilibrium.\textsuperscript{11}

The combination of distribution and coordination problems within the BotS game also poses a hurdle for the 1997 Anti-Bribery Convention members, since the distribution problem is not entirely resolved through coordination and the information problems via signaling. Nonetheless, the institution does pose a formal alternative to the straight-up leadership/follower equilibriums that are maintained within the collusive framework. Given that all members have the same discount rate in the game, I show that the amount of information revelation increases at two of the equilibria with communication, now that the Convention has come into play.

In line with the assumptions of non-cooperative games I agree with Morrow (1994) that uncertainty over what is best for the entire group makes it more likely that the sides will cooperate if they share their ‘private information honestly’ though not necessarily all the time. This dilemma carries over then into the form of communication between the players, since followers may not always choose to follow leaders, knowing that it is best

\textsuperscript{11} Morrow 1994, pp. 389.
for the whole of the OECD.\textsuperscript{12} By choosing to include only certain issues in the Convention at the exclusion of others, the Convention is realistic in its assumption of revelation of private information. Significant issues such as the definition of a public official have been left out of the Convention at the behest of important signatories like Germany and Japan, making the case for collusion more convincing.

The model seeks to introduce the following refinements of the conditions for the \textit{folk theorem} identified in the earlier paper. In that paper we find that, ‘[i]f problems of both distribution and information exist, an actor may be motivated to misrepresent its information about the value of solutions in the hope of producing an outcome it believes is better for itself.’ \textsuperscript{13} It is through simplified models that we can understand why cooperation continues, given these incentives against cooperating. This paper is most specifically concerned with two observations made by Morrow, (1) players will not be dishonest about their preferences when they have enough of a common interest, and (2) players values for leadership vary along the same measure of common interest.\textsuperscript{14}

By including a trigger punishment strategy on a variation of the forms of communication among players, the model outlined here shows that:

\textbf{(1)} trigger punishment strategies will support honest communicative equilibria over lower probabilities of common interest among actors if the discount factor is sufficiently high,\textsuperscript{15}

\textbf{(2)} trigger punishment strategies will support leadership equilibria over higher probabilities of common interest among actors, if the discount factor is sufficiently high.\textsuperscript{16}

\textsuperscript{12} Morrow, pp. 4.
\textsuperscript{13} Ibid., pp. 395-396.
\textsuperscript{14} Ibid., pp. 393-394.
In addition to demonstrating here how a cartel uses the public signal of the price for bribes in foreign countries, I find that by showing how the benefits to cheating are less than the one-period discounted costs of punishment in the formal set up of the convention. The subject of trigger punishment has been the subject of folk theorems in various forms, often regarding cooperation over a long period of time or regarding the identification of certain focal point of interest to actors that aligns their expectations regardless of the distributive consequences. On the other hand, the existence and success of cartels has generally been associated with the market structure of the industry and/or the enterprise, while their demise has been attributed to market entry, elasticity of supply, elasticity of demand, larger than optimal membership and other times their degree of explicit collusion on price vs. quantities of the good. Porter (1983) provided empirical evidence that the freight-rail industry lost its edge once competition was allowed. This study seeks to make headway as to whether the degree of institutionalization of the collusion is a good enough reason for colluding states to embrace price wars.

The repeated games framework is often used for achieving cooperation where there are no incentives to cooperate in the short-run. On the other hand, it is often the case that the full costs and benefits of an international agreement will not be realized until most of the details have been ironed out after a first run. This makes it in fact more likely that states will be coordinating without communicating on issues they know has distributive

---


17 Axelrod (1984) discusses the desirability of collective stability as long as the future of interaction is important enough as it applies to trench warfare, pp. 61-63.

consequences, but will gain more by ignoring its own signal about the state of the world when leadership equilibria pose advantages. The punishment mechanism is that of defection (refraining of punishment of individual investors for bribing foreign officials) by all member states in the second period, regardless of the messages received in the second stage. The difference between this model and Morrow (1994) is that (1) there is an equilibrium where reciprocal punishment takes place if an actor is caught lying about the signal that it has received in the first stage, and (2) there are in fact more gains to be had by following the leader when the alternative is two-way communication. The trigger mechanism then works to punish when the publicly observable price for bribery increases.

ASSUMPTIONS

The following assumptions underlie the model:

(1) States play a repeated stage game
(2) There is a finite set of players for the 2x2 game that can be generalized to n.
(3) There is a finite set of moves that a player can choose during the stage game where \(|M_i| = \{A, B\}\).
(4) Let \(S \equiv S_1 \times \ldots \times S_n\) denote the strategy (signaling-move) profile of the players.
(5) States care about the future of collusion with varying degrees.

THE MODEL

The model is a variation on Morrow (1994) which demonstrates the incentive structures and equilibrium involved in international coordination problems which also involve distribution and information problems. The main dilemma is the inability of any international organization of solving both problems – where either the distribution or the information problem prevents the other from being solved completely. The application of
this model to collusion among the members of shows how the threat of punishment by the other members leads to greater cooperation on the part of some of the members. The following is a model involving the two period repetition of the stage game set up by Morrow (1994), with the second being a punishment stage conditional on the public signal at the end of the first stage. The punishment mechanism is that of defection by all members in the second period, regardless of the messages received from Nature in the second stage. The difference between this model and Morrow (1994) is that there is an equilibrium where reciprocal punishment occurs if an actor is caught lying about the signal that it has received. This trigger punishment mechanism allows for the honest revelation of information.

The payoffs to coordination (international cooperation) are depicted in three possible renditions where the Battle of the Sexes version involves distribution problem along with the coordination problem and the other two possible stage games are pure coordination games – where players have to decide which move to coordinate on (either A or B). This paper maintains the same set up with respective probabilities of appearing of (1-p), (1/2p) and (1/2p). I take communication to mean communicative signaling between the parties, except when either one or both of these signals are ignored in the leadership/follower or pure cooperative results. The game is played in every investment state separately, such that, investor states that are members of the OECD will receive a new signal from the state they enter to invest. The defections constitute bribing of officials at that state to which a member has been determined to be lying about the signal it has received.
Punishment in OPEC takes the form of defection by all parties by not punishing their own individuals, which increases the costs of investment of all investing states. The punishment implies millions of dollar of lost revenue for OECD members. This paper shows what can be labeled as Trigger Punishment (Folk Theorem) where trigger mechanism of punishment by the other members leads to more cooperation between the states that are involved. The main finding of the paper is that trigger punishment can lead to more honest revelation of information in the first stage. This is a separate equilibrium than that found in Morrow (1994). The model has two states of the world: normal and punishment states. In the punishment state, the players will all play BB, regardless of what signals were received, only if Player 1 is revealed to be lying. If Player 2 is revealed to be lying, then they play AA in the next round.

**First Round of the Stage Game:**
1- Players get private information about the game being played  
Choose set: (can bribe or not)
2- Players exchange messages (1 or 2) about the game being played  
Choose set: (will bribe or not)
3- Players choose their strategy (A or B)  
Choose set: (punish or not)
4- The game is revealed (as having been either: Bots, BPMA or BPMB) and the payoffs for the first stage are received.  
(the one that punishes loses)

**Second Round:** I assume there is no discounting since we are looking for solutions to the extreme. This same punishment can be enforced for longer periods if necessary.
1- If the other player’s signal in the first stage is proven to be dishonest, then both play their set strategies of B in this second round. Otherwise, the game is exactly like the first round.

No one punishes in the second round if one player 1 is revealed to be lying (this is a symmetric game)
2- Players choose their strategy.
Both don’t punish
3- Players receive their payoff.
Both have increased costs
The game has the same set up as Morrow (1994) except for the order of play, the once repetition of the first stage and the revelation of the game at the end of each stage through the public signal which I assume to be the market price for bribery.

**Payoffs of the Stage Game:**

**BOTS**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a , 1</td>
<td>0 , 0</td>
</tr>
<tr>
<td>B</td>
<td>0 , 0</td>
<td>1 , a</td>
</tr>
</tbody>
</table>

**BPMA**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a , a</td>
<td>0 , 0</td>
</tr>
<tr>
<td>B</td>
<td>0 , 0</td>
<td>1 , 1</td>
</tr>
</tbody>
</table>

**BPMB**

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1 , 1</td>
<td>0 , 0</td>
</tr>
<tr>
<td>B</td>
<td>0 , 0</td>
<td>a , a</td>
</tr>
</tbody>
</table>

**Normal state:**

Messages: 1 if your signal is 1; 2 if your signal is 2 (ie. honest signals)
Moves: if the messages are both 1, play A; if they are both 2, play B. If one message is 1 and the other is 2, use public coin flip procedure to select which move to play.

If Player 1 sends message 1 when the game being played is BMPB, then move to Punish 1 state. If Player 2 sends message 2 when the game being played is BMPA, then move to Punish 2 state.

**Punish 1 state:**

Messages: Player 1 always sends 1, Player 2 always sends 2 (or any other form of babbling)
Moves: Play B
Return to Normal state after one round of this state.

**Punish 2 state:**

Messages: Player 1 always sends 1, Player 2 always sends 2 (or any other form of babbling)
Moves: Play A
Return to Normal state after one round of this state.
CONCLUSION

The communicative and leadership equilibria are enhanced since trigger punishment makes it in the interests of the players to cooperate. Furthermore, the members of the Anti-Bribery Convention are have plans to discuss issues which have not yet come under scrutiny, which may account for the coordination without communication aspects of the agreement. These include (1) the question of who may be included within the definition of a foreign official, (2) the tax deductibility of bribes (as gifts), and (3) the obligation of identification of ECA (Export Credit Agency) officials when doing business abroad. 19

This may be one reason why communication and leadership equilibria have been successful within the confines of the Convention.

The above model demonstrates that by instituting a trigger punishment mechanism for any cheating, the members of OECD ensure that their members are more honest in their compliance with the standards set by the cartel. The punishment for dishonest communication by either player leads to a round of punishment by both players in the second round. The above demonstrates the effectiveness of such a punishment strategy once it is revealed that one or the other has been dishonest at the end of the first round. The increase in the probability of honest communication is promising for the future effectiveness of the collusive efforts of most of the investor states and the threat of punishment with the Convention.

19 http://www.eca-watch.org/problems/corruption/ECAW_bribery_consensus_fails_22nov05.htm
Proof for Communicative Equilibria:

Proof for Communicative Equilibrium HERE

The moves of this equilibrium are self-enforcing at the stage of playing moves, since these are coordination games, only the signaling stage for the conditions which the signals will be followed honestly is checked here. Thus, for $p$ in range and $\delta = 1$, the strategies of the honest communicative equilibrium form an equilibrium. The degree of distributional conflict then determines how far the players will waver from the probability of honestly following their signals. Only message pairs (1,1) and (2,2) are supported in this equilibrium, while (1,2) and (2,1) automatically enter the punishment phase in the second round (play BB).

The two cases where the signals don’t match the messages will result in automatic BB game: (1) if the game is revealed to be BPMB and player 1 sends message 1 and (2) if the game is revealed to be BPMA and player 2 sends message 2. As a result, the players are both better off then if the game is revealed to be a BotS game. Both players lose in the case that either is revealed to have sent a dishonest message in the first round leading to an automatic play of BB in the second round. Symmetrically BPMA would have the worst outcome for Player 2. In comparison to the Morrow (1994) conditions for honest signaling, the probability for honest communication increases from $p \geq 1/2$.

Proof for Leadership Equilibrium HERE
References


