

The Honest Broker: Mediation and Mistrust



Andrew Kydd
Harvard University

Presentation at the Mershon Center,
Ohio State University, March 5, 2004

Mistrust can cause Sub-optimal Outcomes

■ The Security Dilemma

- Mistrust leads to conflict (Hobbes)



Problem of Exchange

- Mistrust can lead to market failure
- Milgrom, North and Weingast (1990) on the medieval revival of trade and *Lex Mercatoria*





Mediation is supposed to facilitate cooperation

- Can mediators promote trust?
- Under what conditions?
- Should they be biased or unbiased?

Standard model of these issues is the Prisoner's Dilemma

Player 1

Cooperate

Defect

Player 2

Cooperate

Defect

3, 3	1, 4
4, 1	2, 2

But PD is not appropriate

- Trust is a belief that the other side is likely to prefer cooperation to exploitation
 - Presupposes uncertainty on this score
- In one shot PD, neither side does
- In repeated PD, they either do or do not, depending on the discount factor, δ
 - Either way, there's no uncertainty

Mediation and Bias

- Theory of how mediation works is underdeveloped
- So not clear when mediation can promote trust
- One question concerns mediator bias
 - Is bias a good thing (Touval) or a bad thing (Young)?



Which Side Are You On?

- In my previous work I argued bias was essential (Kydd 2003)
- If the mediator tries to persuade one side to make a concession because the other side will fight, the mediator needs to be biased towards the side it is speaking to



However . . .

- Things may be different for different tasks that the mediator is trying to perform
- Perhaps for the trustbuilding role, being unbiased is better

A Model of Mediation and Mistrust

Player 2

Cooperate

Defect

Player 1

Cooperate

$1, 1, \rho$

$-a_1, b_2, -\beta$

Defect

$b_1, -a_2, \beta$

$0, 0, 0$

	Cooperate	$1, 1, \rho$	$-a_1, b_2, -\beta$
	Defect	$b_1, -a_2, \beta$	$0, 0, 0$

The Players' Types

- Each player may be trustworthy

$b_i < 1$, (Assurance preferences)

likelihood t_i

- Or untrustworthy

$b_i > 1$, (Prisoner's Dilemma preferences)

likelihood $1-t_i$

In equilibrium:

- Untrustworthy types have a dominant strategy to defect
- Trustworthy types can cooperate if they think the other side is likely enough to be trustworthy

The Mediator's Payoffs

- Reward ρ for successful cooperation
- Payoff β in case player 1 exploits player 2
- Payoff $-\beta$ in case player 2 exploits player 1
- So β is a measure of how biased the mediator is towards player 1
- If $\beta = 0$, the mediator is unbiased
 - If $\beta > 0$ the mediator favors player 1
 - If $\beta < 0$ the mediator favors player 2

The Mediator's Beliefs

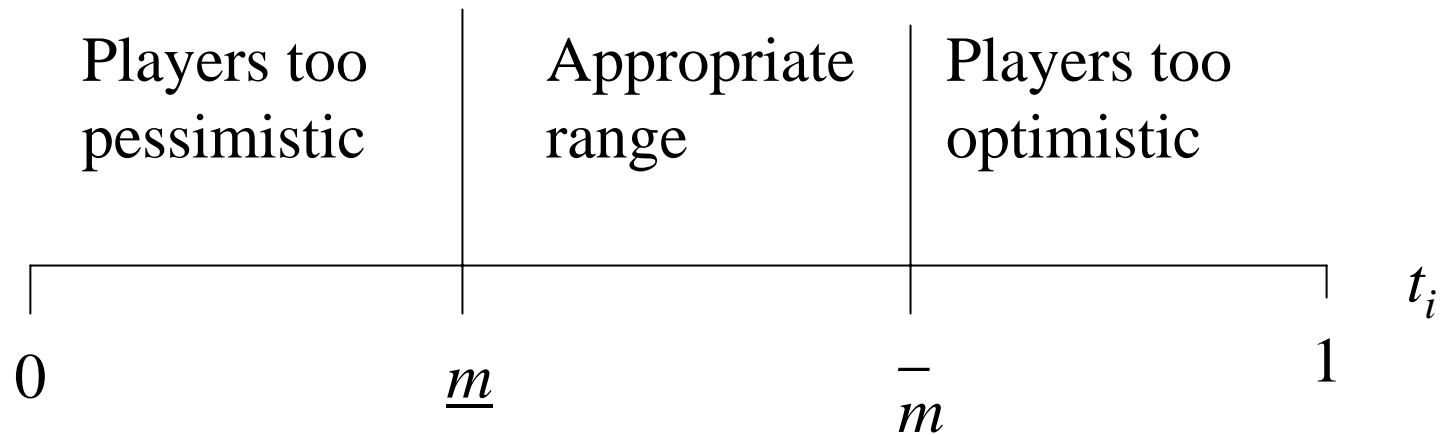
- Mediator gets signals from Nature about the players' types
 - likelihood of error is ε where $\varepsilon < 0.5$
- If the mediator gets the T signal, belief is
 - $P(iT|T) = t_i(1-\varepsilon)/[t_i(1-\varepsilon) + (1-t_i)\varepsilon]$
- If the mediator gets the U signal, belief is
 - $P(iT|U) = t_i \varepsilon/[t_i \varepsilon + (1-t_i)(1-\varepsilon)]$
- Note $P(iT|T) > t_i > P(it|U)$

Order of Play

- Nature chooses each player's type, signals mediator
- Mediator makes announcement about each player's type, T or U
- Players play game (simultaneous choice)

When Mediation can Help

Mediation is only useful for middling levels of uncertainty



Truth-telling equilibria

- We are interested in truth-telling equilibria
 - Where the mediator faithfully communicates to the players what it has learned from Nature
- If the players are trustworthy, the mediator says so, encouraging cooperation
- If the players are untrustworthy, the mediator says so, preventing cooperation

One Round Game

- In the one round game, there is no truthtelling equilibrium
- The mediator will not tell the truth about the players' types
 - Regardless of how biased or unbiased the mediator is
- Hence mediator is incapable of building trust

Why no truthtelling equilibrium?

- The mediator gets a positive payoff, ρ , if the players cooperate
- But gets zero if they both defect
- Hence the mediator has an incentive to encourage cooperation

Incentives for dishonesty

- If the mediator thinks the parties are likely to be trustworthy, the mediator will wish to say they are trustworthy, to encourage cooperation
- If the mediator thinks the parties are likely to be untrustworthy, the mediator still wants to vouch for them, to encourage cooperation
- Hence the mediator has an incentive to lie

The Repeated Game

- Consider a repeated version of the game
- Same mediator, new players each round
- Same information structure
- Mediator is “fired” if ever caught in a lie
 - That is, if she vouches for a player who subsequently defects
- Otherwise, mediator passes on to the next round

Repeated Game Payoffs

- Fee ϕ for mediating in each round
 - Unconditional on the outcome
- Discount factor δ
- Likelihood of getting to the next round γ
- Stage game payoff π

Equilibria in Repeated Game

- Here there are truthtelling equilibria
 - If the level of bias β is not too great
 - And the reward ρ is not too great
 - And the fee ϕ and discount rate δ are not too small

What's the incentive for honesty?

- If the mediator gets good news about the players, passing it on maximizes likelihood of reward, ρ , just as in one shot game
- If the mediator gets bad news about the players, passing it on maximizes the likelihood of getting to the next round, and earning the fee, ϕ
 - Can't be caught in a lie if you don't vouch for the parties

When is there a truth-telling equilibrium?

- Reward ρ cannot be too great
 - Otherwise it becomes similar to the one shot game
 - Mediator has incentive to encourage cooperation even if she has bad information about the parties



And . . .

- The fee ϕ and discount rate δ cannot be too small
 - Or again it will be like the one shot game in which there is little incentive to try to get to the next round



And . . .

- The mediator cannot be too biased (holding the other payoffs fixed), β cannot be too large or too small
 - Or the mediator will have an incentive to vouch for the party towards whom they are biased

For instance

- if β is big, the mediator will be happy when player 1 exploits player 2, and will have an incentive to encourage this outcome by vouching for player 1 after getting bad information about her

Conclusion

- Mediators can build trust provided that
 - They have a reputational incentive for honesty
 - They are not excessively biased towards one side or the other
- So in the trustbuilding context, unbiased mediators are best, unlike in the bargaining context