Bonus Maths 5: GTO, Multiplayer Games and the Three Player [0,1] Game

In this article, I'm going to be exploring some multiplayer games. I'll start by explaining the really rather large differences between two player, zero sum games and *any* other games, and illustrate this with three very simple poker-themed examples. I'll then discuss the three player [0,1] game. A lot of my interest in these topics has been sparked by the release of PokerSnowie – a poker bot that many people refer to as trying to 'play Game Theory Optimal (GTO) poker'. I will have more to say about this later, but let's start with some background.

Two Player Zero Sum Games v Other Games

A zero sum game is one where anything that the first player wins, the second player loses, and vice versa. Heads Up poker is, in principle, a two player zero sum game, although strictly speaking the rake is a third player who always wins. John von Neumann showed that two player zero sum games always have an equilibrium solution, i.e. a strategy (possibly mixed[1]) upon which neither player can improve by changing to a different strategy. This equilibrium solution may not be unique, but if there is more than one equilibrium, each player has the same payoff at each equilibrium, so that they are functionally equivalent. It's my belief that this is what advanced poker players are talking about when they say things like 'I'm trying to play close to GTO', and I'm pretty sure that the world's best HU FLHE players do exactly that. For multiplayer games, the picture is rather different.

Although John Nash's famous theorem proves that at least one equilibrium solution always exists for multiplayer games and two player nonzero sum games, it doesn't guarantee that the equilibrium is unique. In addition, if there are multiple equilibrium solutions, they don't all necessarily have the same payoffs. This means that to say you're trying to 'play GTO poker' doesn't make any sense in a multiplayer game such as SH NLHE. For multiplayer games, I object to the phrase 'play GTO poker' on principle, because it implies that there's a unique equilibrium strategy. Maybe there is, maybe there isn't, but I'd be very surprised to find that SH NLHE has a unique equilibrium solution.

The idea of an equilibrium solution is based on a local analysis of a strategy, i.e. no player can improve by deviating from that strategy. Finding the equilibria of a game tells you nothing about how the game will play out in practice, even between players who know what the equilibrium strategies are, as we'll see from looking at some toy examples in the next section.

I realize that I'm sounding like someone who is not a fan of PokerSnowie, but in fact I am. It's a very thoroughly trained neural network algorithm that has played a gazillion hands against both itself and opponents created to try to exploit it. All it does is try to maximise the EV of its strategy against the strategies that it has seen. That's got to be a very useful tool, and is manifestly a very strong poker bot. It may well be that it plays close to an equilibrium strategy, but, to be honest, I don't care. I'm happy to learn from a strong player, either human or artificial. I just wish people wouldn't keep talking about it 'playing GTO poker', because I don't think it's a meaningful concept in multiplayer games.

[1] A mixed strategy is one that involves potentially playing more than one strategy, with each strategy assigned a probability. For example, bluffing with a Queen 1/3 of the time and checking with it 2/3 of the time in the AKQ game.

Some Simple Examples of Multiplayer Games

The motivation for the following games should be taken with a pinch of salt. I came up with them after it was suggested to me that it's possible to win against PokerSnowie by playing super aggressively. I don't know whether this is correct, but it inspired me to create three toy poker games.

Let's consider a SH table where there are four players playing some rubbishy strategies and two other players who have to decide between two possible strategies: Aggro and Snowie. If they both play the strategy Snowie, they each win at 8bb/100 hands, but if they both play Aggro, they each win at 4bb/100 hands. If one plays Aggro and one Snowie, the player who plays Aggro wins more bb/100 than the player who plays Snowie. There are three qualitatively different possibilities with these restrictions, all of which are interesting.

Game 1

	Aggro	Snowie
Aggro	(4,4)	←(6,2)↓
Snowie	↑(2,6)→	(8,8)

Aggro wins at 6bb/100 hands against Snowie's 2bb/100 hands. In the table above, the first number is the winrate of the player who chooses a row and the second the winrate of the player who chooses a column. What happens if both players play Aggro? If either player switches to playing Snowie unilaterally, he will reduce his winrate to 2bb/100 hands. (Aggro, Aggro) is therefore an equilibrium strategy pair. Similarly, if both players play Snowie, they have no incentive to change strategy, so (Snowie, Snowie) is also an equilibrium strategy pair. As you can see, it's not difficult to have multiple equilibria, even if there are just two players and only two possible strategies for each player. In this example, the payoffs are different at the two equilibria. What does 'play GTO poker' mean in this game? If they play different strategies, the Snowie player has an incentive to turn Aggro, and vice versa. Which one will change strategy first, and how will the other player react? Who knows! That's the tricky point. You need to know about the dynamics of this game when it's played repeatedly to find out what happens. How should we model the dynamics? That's a good question, to which I shall return below, but I first want to make an observation about rationality.

All analysis of toy poker games, and indeed other games, that I've seen is based on the assumption that the participants are rational and try to maximise their payoffs. We assume that if you play Snowie and I play Aggro, I look at the table of payoffs, see that I'm better off playing Snowie and change strategy. In more complicated games, the players don't really know the payoffs from each strategy. You may say that nosebleed poker players know how their strategies fare against other strategies, and that the game is a complex quest to maximally exploit the other players. Maybe it is, maybe it isn't, but I know that low stakes poker players just play how they want to play. Let's say that you're playing Snowie and I'm playing Aggro. You say to me, 'If you were to switch to playing Snowie, we'd both increase our win rates.' Since I'm a pretty rational kind of a guy, I might well say

‘That's splendid! Your argument is persuasive. I'm going to play Snowie too'. However, I would not be surprised to find that some of the players at the stakes I play would tell you to piss off, and carry on playing Aggro. What would you do then? Sigh and play Aggro too? Maybe using player rationality as a modelling assumption isn’t always such a good idea.

Game 2

	Aggro	Snowie
Aggro	(4,4)	←(10,2)
Snowie	↑(2,10)	←(8,8)↑

Let's change the win rates of the players when they use different strategies to 10 and 2bb/100 hands for Aggro and Snowie respectively. If both players start using Snowie, they both have an incentive to switch to playing Aggro. Once they're playing different strategies, the Snowie player has an incentive to start playing Aggro too so that he's not getting pwned any more. In this case (Aggro, Aggro) is the only equilibrium. However, notice that they'd both make more money if they played (Snowie, Snowie). This game is identical to the classical Prisoner's Dilemma game. Both players end up worse off than if they’d settled on (Snowie, Snowie) because our modelling framework tells us that one of the players will start to play Aggro and claim their prize of 10bb/100 hands, even though this will result in an eventual win rate of only 4bb/100 hands when the other player adjusts too.

Game 3

	Aggro	Snowie
Aggro	↓(4,4)→	(10,6)
Snowie	(6,10)	←(8,8)↑

Finally, let's look at what happens when the win rates are 10 and 6bb/100 hands for the two strategies. If the players start off using the same strategy, they each have an incentive to change. Both (Aggro, Snowie) and (Snowie, Aggro) are equilibria. This game is identical to the classical Game of Chicken. There's also a mixed equilibrium strategy for this game where each player could use either strategy with some probability.

So, what might we see in practice if these games are played repeatedly? There is a large academic literature on repeated games, about which I do not claim to know much, but I'd expect a player's choice at each iteration of the game (or ‘hand’ as us poker players call them) to depend on his opponent's play in the previous few hands, and you can take ‘few’ to mean whatever you like here. In these simple games you know what strategy your opponent played in each of the previous iterations; in real poker, it takes a fair few hands to build up a picture of how your opponent is playing, whilst he, if he's thinking along the same lines, is trying to work out how you're playing. There's a complicated dynamic that depends on the timescales on which each of you is gathering and acting on information. This sort of thinking is what people who attempt to 'play GTO poker' say that they're aiming to avoid because they are trying to play 'optimally'. Once you realize that this simplistic view of poker is wrong (and yes I realize that a 'GTO strategy' would be complicated and almost impossible to find, but still it's a simple intellectual framework), you get back to the sorts of ideas that the more traditional player tries to implement (e.g. hand reading, estimating frequencies and ranges, exploiting your opponent). This sort of thinking is not some sort of opposite to 'GTO'; it's a valid way of thinking about the game within the context of game theory.

Even if a multiplayer game has a unique equilibrium solution, there's another possibility that isn't a feature of two player zero sum games – collusion! Let's look at an example.

The Three Player [0,1] Game

Now that we've looked at some simple multiplayer games, I'd like to consider a more complex example – the three player [0,1] game. I discussed the two player version of this game in the book, but here's a quick recap for you.

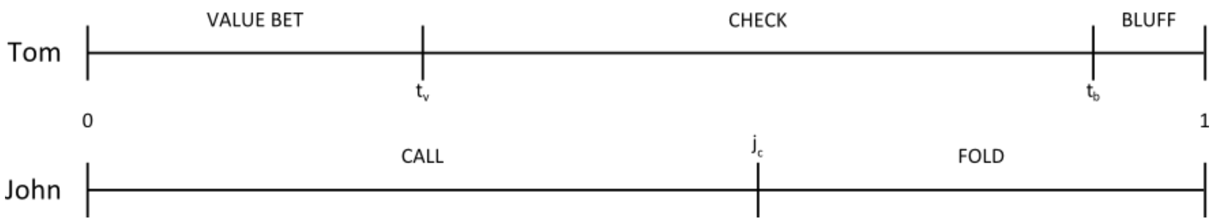
The two player [0,1] game

In this game, the pot starts out with $P in it, and each player is dealt a number between zero and one, with zero representing the best hand and one the worst. John is forced to check, and then Tom can either check behind and take a showdown, or bet $1. If Tom bets, John can either call, after which they show down their hands, or fold, which gives Tom the $P in the pot.

If Tom holds a hand t and John a hand j, the unique equilibrium strategy for this two player, zero sum game is:

Tom: Bet with his best hands, $0 \leq t \leq t_v$, bluff with his worst hands, $t_b \leq t \leq 1$, and check back his midstrength hands, $t_v \leq t \leq t_b$.

John: Call with his best hands, $0 \leq j \leq j_c$, fold his worst hands, $j_c \leq j \leq 1$.



You can find expressions for the value of these ranges as a function of pot size, P, in the book.

Note that $j_c > t_v$. John has to call wider than Tom's value range so that he catches bluffs at the right frequency. As you can see in Figure 1, the larger the pot, the wider Tom can value bet and the wider John has to call. Tom's ex-showdown expectation, $E(T)$, is also shown, and is positive, which shows that Tom's position allows him to extract value from John.

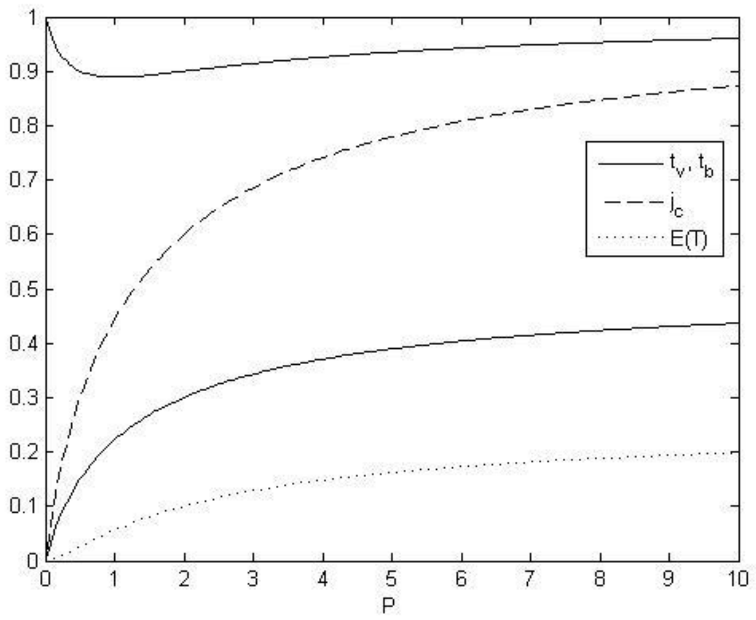


Figure 1: Equilibrium ranges and Tom's ex-showdown expectation in the two player [0,1] game.

The solution of the three player [0,1] game

The three player version of the [0,1] game involves John, Tom and Manu. John and Manu are forced to check, and Tom can either bet $1 or check behind. If Tom bets, John acts next and can either call, after which Manu can overcall or fold, or fold, after which Manu can call or fold. The structure of any equilibrium solution must be similar to that of the two player game, with Tom value betting his best hands and bluffing his worst hands, and the other two players calling and/or overcalling with their best hands.

We can write this as:

If Tom holds a hand t, John a hand j and Manu a hand m:

Tom: Bet with his best hands, $0 \leq t \leq t_v$, bluff with his worst hands, $t_b \leq t \leq 1$, and check back his midstrength hands, $t_v \leq t \leq t_b$.

John: Call with his best hands, $0 \leq j \leq j_c$, fold his worst hands, $j_c \leq j \leq 1$.

Manu: If John folds, call with his best hands, $0 \leq m \leq m_{c1}$.

If John calls, overcall with his best hands, $0 \leq m \leq m_{c2} \leq j_c$.

Otherwise, fold.

I'm not going to give any details of the calculations that lead to the following conclusions, but I find that there is an equilibrium solution with this structure, and I have good reason to believe that it is unique.

Before I show you what the equilibrium solution looks like, let's see whether your intuition is better than mine, because I got this completely wrong. Where do you think John and Manu's calling ranges lie relative to each other and Tom's value range? Will they both call wider than Tom's value range to catch bluffs? Which one will call wider? Take a moment to give it some thought.

Figure 2 shows the equilibrium solution. As you can see, John calls slightly tighter than Tom's value range ($j_c < t_v$) and leaves the bluff catching to Manu, who calls wider than Tom's value range ($m_{c1} > t_v$). John has the worst position initially and then, after Tom bets, he has the worst relative position. He is being squeezed, and can't bluffcatch without allowing Manu to exploit him by calling tighter. We can also see that, in comparison to the two player version of the game, Tom must value bet and bluff a little tighter when he has two opponents, as we might have expected, and that Manu's calling range in the three player game is very similar to John's calling range in the two player game.

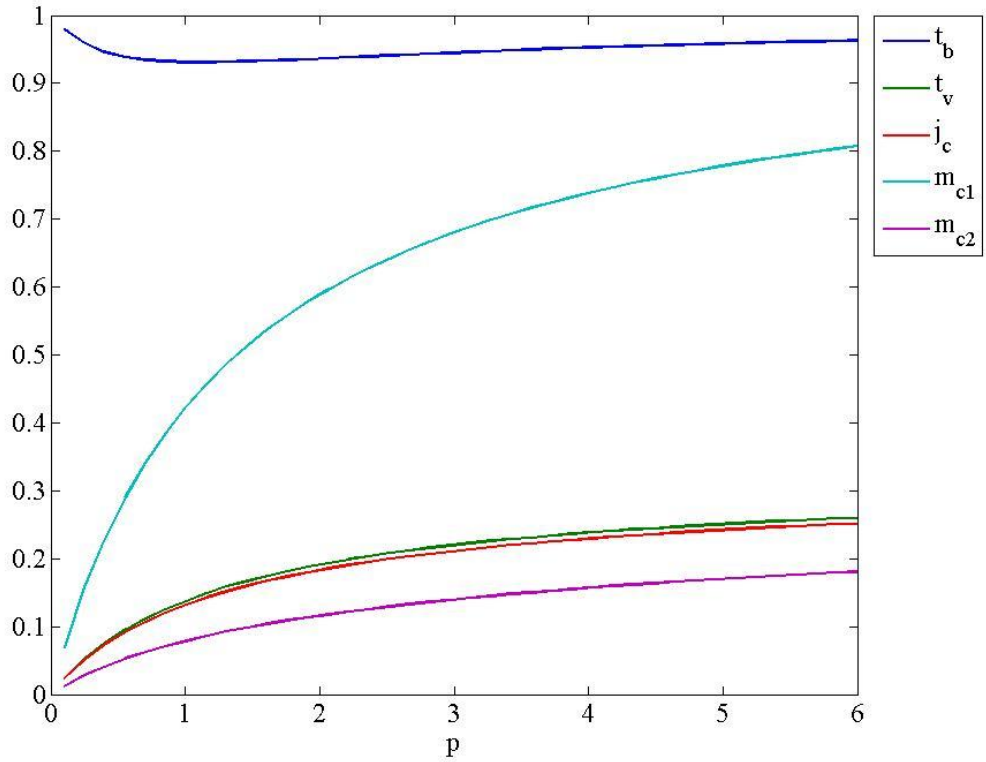


Figure 2: Equilibrium ranges in the three player [0,1] game.

Figure 3, which shows the ex-showdown expectations, demonstrates that the game is very favourable to Tom. His position allows him to exploit John and Manu roughly equally at equilibrium.

Collusion

An aspect of multiplayer games that isn't a feature of two player games is collusion, both implicit and explicit. If two of the players decide that they'll pool and split their winnings, the game becomes two player and zero sum. If the victim doesn't realize that he's being colluded against, he's playing a different game, with a different equilibrium solution and will usually lose more money. In explicit collusion, the colluders agree beforehand to collude and split up their winnings away from the table (this is cheating). In implicit collusion, the colluders have no pre-game agreement, but both recognize that they can improve their payoffs by playing a strategy that is not an equilibrium solution of the three player game, because they could exploit each other if they chose to, but which leads to them both making extra money from their victim. Implicit collusion is part of the game and should not, in my opinion, be considered cheating. The standard example is on the bubble of a SnG, where two or more players may decide to check a hand down when the likely bubble boy is all in, in order to increase the chances of their victim's elimination, after which all the colluders cash.

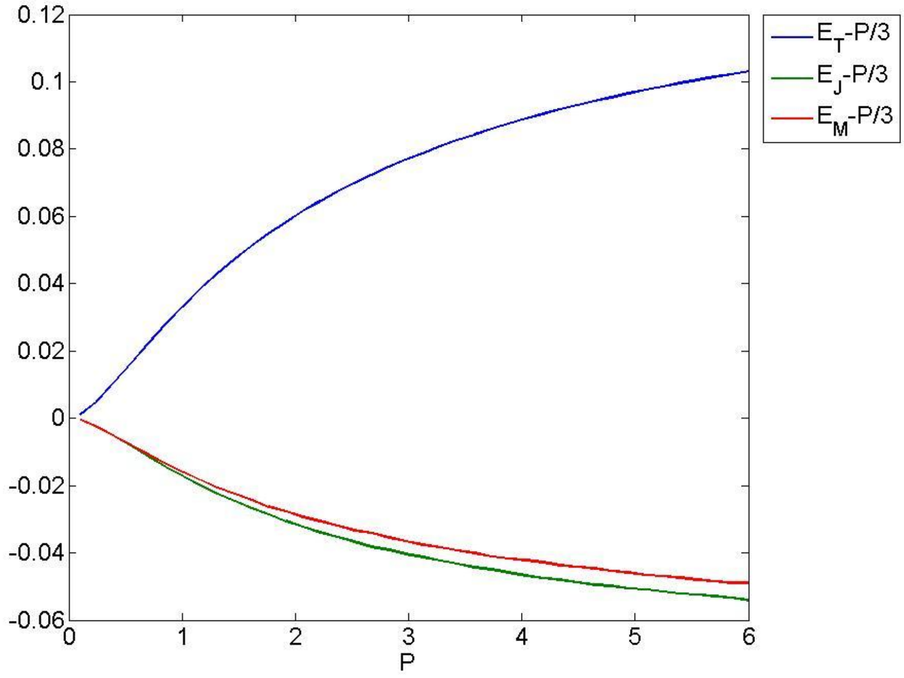


Figure 3: Equilibrium ex-showdown expectations in the three player [0,1] game.

Let's consider each possible alliance in turn.

1) *John and Manu v Tom*

Collusion between John and Manu is pointless. At equilibrium, Tom is protects himself by bluffing and value betting in the appropriate ratio, and there's nothing that Manu and John can do to exploit him.

2) *John and Tom v Manu*

The alliance between John and Tom is the one that I thought would work well, and in theory it does. Figure 4 shows that Tom bluffs extremely wide, and John calls every time. This sucks a huge extra profit out of Manu, as shown in Figure 5. However, there are two problems with this. Firstly, Tom loses heavily compared to his expectation playing the three player equilibrium strategy, because all his bluffs are called by John, who gets all of the additional profit and some of Tom's profit. This means that the alliance can only work as an explicit one, with John handing Tom his share of the spoils after the game in some dim alley behind the poker club. Secondly, this strategy will be blindingly obvious to Manu. Tom bets. John call. Tom bets. John calls…… and Manu will see the midstrength hands that Tom is turning into bluffs and the garbage with which John is calling. This is not a practical strategy.

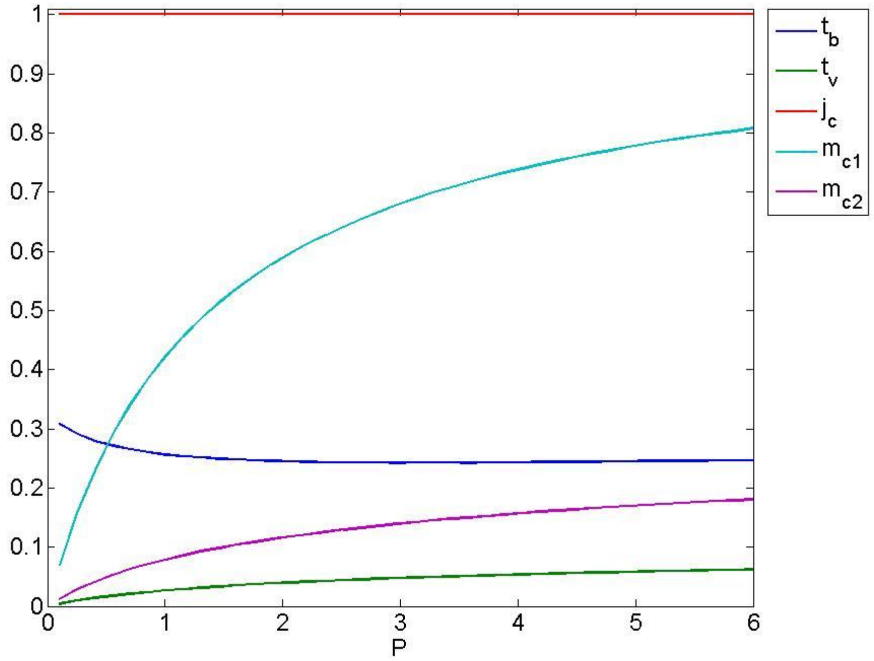


Figure 4: Manu's equilibrium ranges. John and Tom's maximally exploitative collusive ranges in the three player [0,1] game.

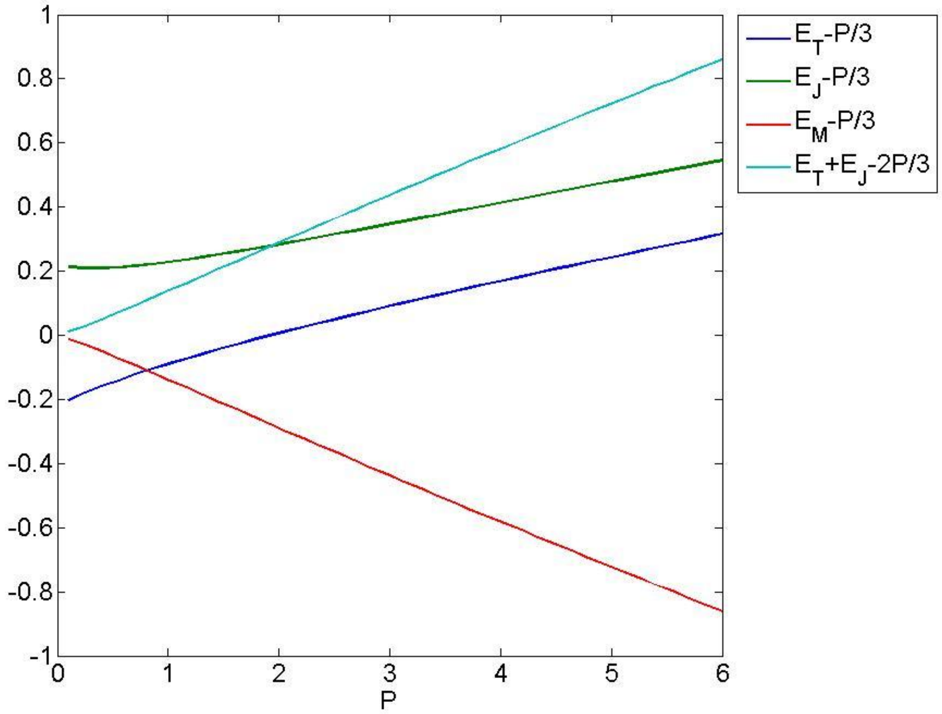


Figure 5: Ex-showdown expectations when John and Tom collude to exploit Manu.

3) *Tom and Manu v John*

This is an alliance where a subtle possibility for implicit collusion exists. We've seen that the three player equilibrium has Manu doing the bluff catching. If Manu and Tom are colluding, they can exploit John by having Tom bluff wider than in the three player equilibrium. Manu can exploit this by calling wider, but if he's in cahoots with Tom, all he cares about is making John fold. If Manu isn't bluff catching properly, John is folding too often, as shown in Figures 6 and 7.

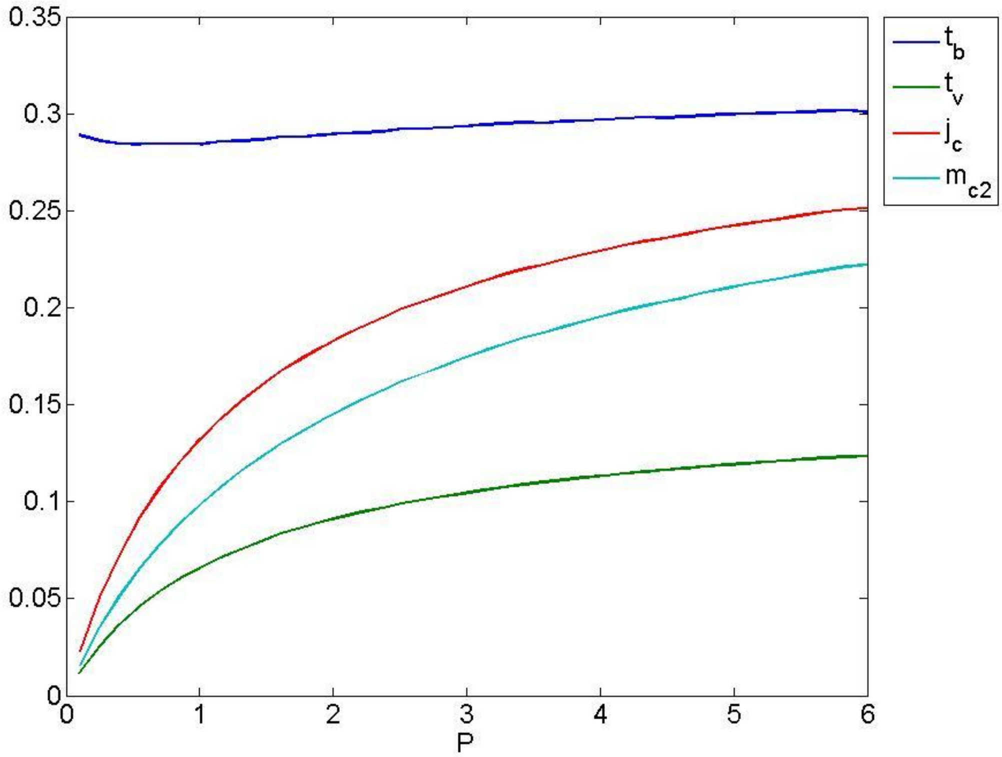


Figure 6: John's equilibrium range. Tom and Manu's maximally exploitative collusive ranges in the three player [0,1] game. Manu's calling range is not shown, as it doesn't affect the expectation of the colluders (Tom and Manu), only its distribution between them.

The subtle point is that, once John folds, Manu and Tom have won the pot. Manu can then adjust his calling frequency to distribute the profit fairly between himself and Tom. For example, if the pot is $1:

Three player equilibrium:

$t_v = 0.138, t_b = 0.931, m_{c1} = 0.421$

Profit: (T, J, M) = (0.366, 0.316, 0.317)

Tom and Manu v John's three player equilibrium strategy:

$t_v = 0.066, t_b = 0.285, m_{c1} = 0.268$

Profit: (T, J, M) = (0.421, 0.207, 0.372)

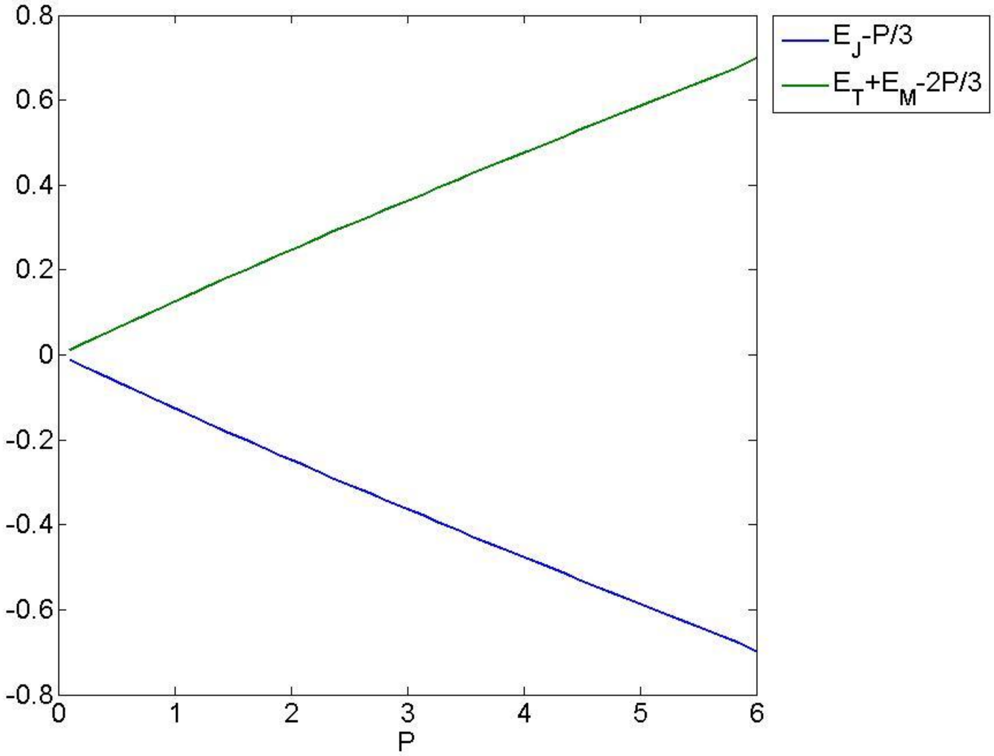


Figure 7: Ex-showdown expectations when Tom and Manu collude against John.

This is a plausible implicitly collusive strategy. Firstly, the profits are redistributed at the table. No trips to the alley behind the poker club are required. Secondly, this would not be completely obvious to John. Tom would be betting rather more than he is at equilibrium, but he could tone down the bluffing and still make this sort of strategy work. Manu is not calling much less than he would be at the three player equilibrium. Thirdly, there would actually be fewer showdowns than at equilibrium, so John would have less information than he would at equilibrium.

The difficulty with this is that, since Tom's frequencies are far from the three player equilibrium, there's an opportunity for Manu to defect from the alliance and exploit Tom instead. If Tom plays the collusive strategy and Manu plays the three player equilibrium strategy, Manu is a big winner, because he's calling wider than Tom wants him to and picking off all his bluffs. We can explore this by creating a metagame out of the situation.

A metagame

Let's fix the pot size at $1 and give each player two possible strategies. One strategy for each player is to play the three player equilibrium strategy, which we'll label E. Tom and Manu will also have the possibility of playing the collusive strategy that I described above, which we will label C. We will give John the possibility of playing a calling range that maximally exploits Tom and Manu if they collude against him, in this case $j_c = 0.296$, an anti-collusive strategy that we'll label AC. Tables 1 and 2 give the expectations, (T, J, M), of each of the players for each of the eight possible combinations of strategies.

John - **E**	Tom - **E**	Tom - **C**
Manu - **E**	(0.366,0.316,0.317)	(0.264,0.207,0.529)
Manu- **C**	(0.366,0.316,0.317)	(0.421,0.207,0.372)

Table 1: Expectations, (Tom, John, Manu), when John plays his equilibrium strategy, E.

John - **AC**	Tom - **E**	Tom - **C**
Manu - **E**	(0.368,0.313,0.319)	(0.130,0.403,0.466)
Manu- **C**	(0.368,0.312,0.320)	(0.258,0.396,0.346)

Table 2: Expectations, (Tom, John, Manu), when John plays his anti-collusive strategy, AC.

If we fix John's strategy to be E, Table 1 shows that this is similar to Game 2, the Prisoner's Dilemma-like game that we analysed earlier. Tom and Manu increase their payoffs by each playing C, but then Manu has an incentive to defect from the alliance by playing E, which exploits Tom. Tom should then start playing E as well, leaving both of them worse off than if they'd colluded. We can conclude that the three player equilibrium remains an equilibrium in this metagame, even if collusion between Tom and Manu is allowed. In addition, Table 2 shows that John has a strong incentive to switch to AC if Tom plays C. Tom then has an incentive to switch back to E.

All of this assumes that the players are rational and know each other's strategies. If only poker were that simple.

Conclusions

I hope I've shown you that multiplayer games and nonzero sum, two player games have solution structures that are far more complex than those of zero sum, two player games. Collusion, both explicit and implicit, is always a possibility with more than two participants, and even when we can calculate the equilibrium solution or solutions, the game can proceed as a complex dance between competing strategies. Even so, the dynamics of a real game are likely to be far harder to unravel than the rational responses to known opposing strategies that we've analysed here. It seems to me that many poker players treat the game as if it were a two player zero sum game, when it clearly isn't. The exploitative techniques propounded by players like TT and Manu in our book, which are often contrasted to the 'GTO' strategies suggested by some, are actually entirely consistent with the lessons we can learn from multiplayer game theory and need to be included in any player's understanding of the game, alongside the concepts of balance and range-based thinking that emerge from an appreciation of the structure of equilibrium solutions.