Jeopardy 2019

Intro/Background

Over a decade ago I noticed that contestants on Jeopardy seemed to be wagering in a suboptimal manner in certain situations, so I collected the data from j-archive, created a framework for analysis, and applied game theory to determine that my initial hunch appears to be correct: in some circumstances contestants are not playing in a strictly rational manner; instead, they appear to use simple heuristic rules of thumb to guide their actions. I wrote up my findings into an extended abstract, presented it at the Stony Brook Game Theory conference, and used this credit to bolster my applications for Econ PhD programs. What I didn't do, however, was ever actually write this into a proper paper for publication in a relevant journal. This is mostly because descriptive, explanatory papers are not typically considered worthy of publication, so without a deeper hook there wasn't much I could do with this work at the time.

One benefit of taking almost 10 years off from this project is that the dataset has continued to grow; Jeopardy is as popular and competitive as ever before. And, thanks to the Python programming language, I was able to convert what was previously a manual data collection effort into an automated scraper. So the time seems right to dust this analysis off and present two of the core results in this post: first, that going into Final Jeopardy in close games the player in second place wagers too much too often; and second,in less competitive games, the leading player overuses the heuristic of wagering just enough to cover the second place player doubling up.

Note to self- Rewrite above focused around heuristics and mixed strategy Nash Equilibria, and how randomly selecting between heuristics is difficult

The data:

The data for this analysis consists of the over 4,115 regular-season games of Jeopardy played between September 1997 and July 2019, scraped from the website j-archive.com.

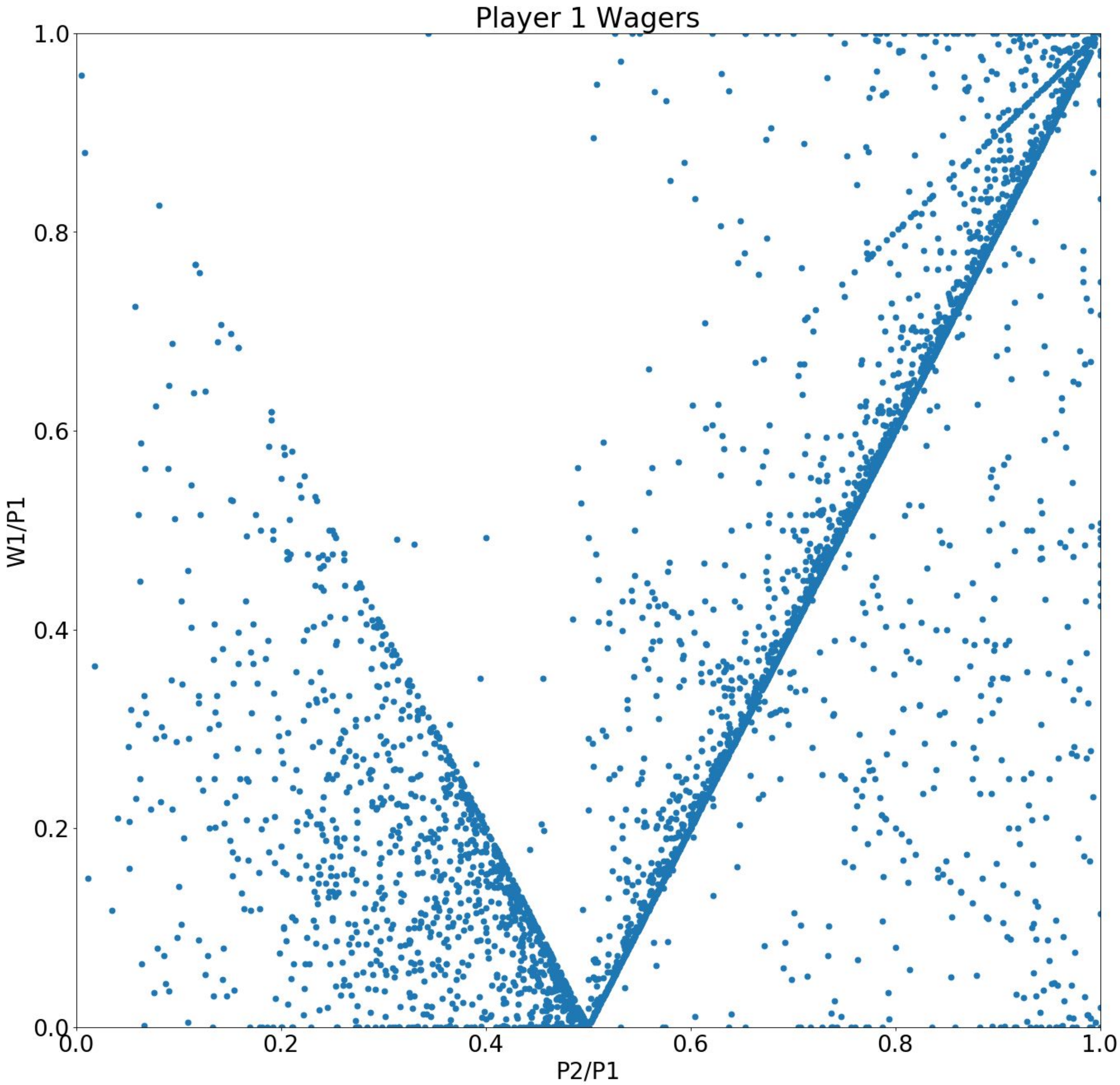


Figure 1: Scatterplot showing how Player 1 wagered across all games in the dataset

Notation:

The players are named based on their relative positions going into Final Jeopardy (FJ). Player 1 is the player currently in the lead, Player 2 is the player in second place, and Player 3 is the player in last place. Let PX (X=1,2 or 3) represent the score of Player X going into FJ, and WX represents player X's wager.

Heuristics:

A heuristic is a fancy word meaning “rule of thumb”. In behavioral economics, it is recognized that it’s often not feasible in reality for humans to perform the calculations required to rationally maximize their decision making, and often fall back on heuristics to guide their actions. In a time-constrained, high stress environment like Jeopardy it shouldn’t be surprising that the contestants are relying on heuristics to make their wagers. It is the premise of this paper that these heuristics sometimes lead players astray, acting in a way that seems irrational.

An example of a heuristic in Jeopardy is player 1 wagering just enough to cover player 2 wagering everything and getting the question correct. In terms of the above notation, this heuristic can be expressed as $W1>2*P2-P1$. This heuristic is so strong it is clearly visible in figure X: it is the upward sloping line along which most of the wagers are clustered. For player 2, a major heuristic is to wager just enough to get above Player 1’s current score ($W2>P1-P2$). This corresponds to the clear downward sloping curve visible below in figure X.

Each of these heuristics can be derived from the wager space described in the next section, with one exception, the other upward sloping line clearly visible in figure 1, described by $W1/P1=P2/P1$, or more simply $W1=P2$. This would seem to be an example of the “availability heuristic”, player 1 is choosing one of the numbers available to it, player 2’s current score, as its wager. This is odd because wagering that amount comes with no strategic significance, and yet it’s a common enough choice that’s clearly visible. There’s a similar heuristic involving round numbers, players will often select wagers so that their winning total will come out to be a round number like 20,000 (or they’ll wager a round number like 10,000). Due to how the plots are normalized these won’t be visually apparent, and maybe I should make a separate section for these non-strategic heuristics.

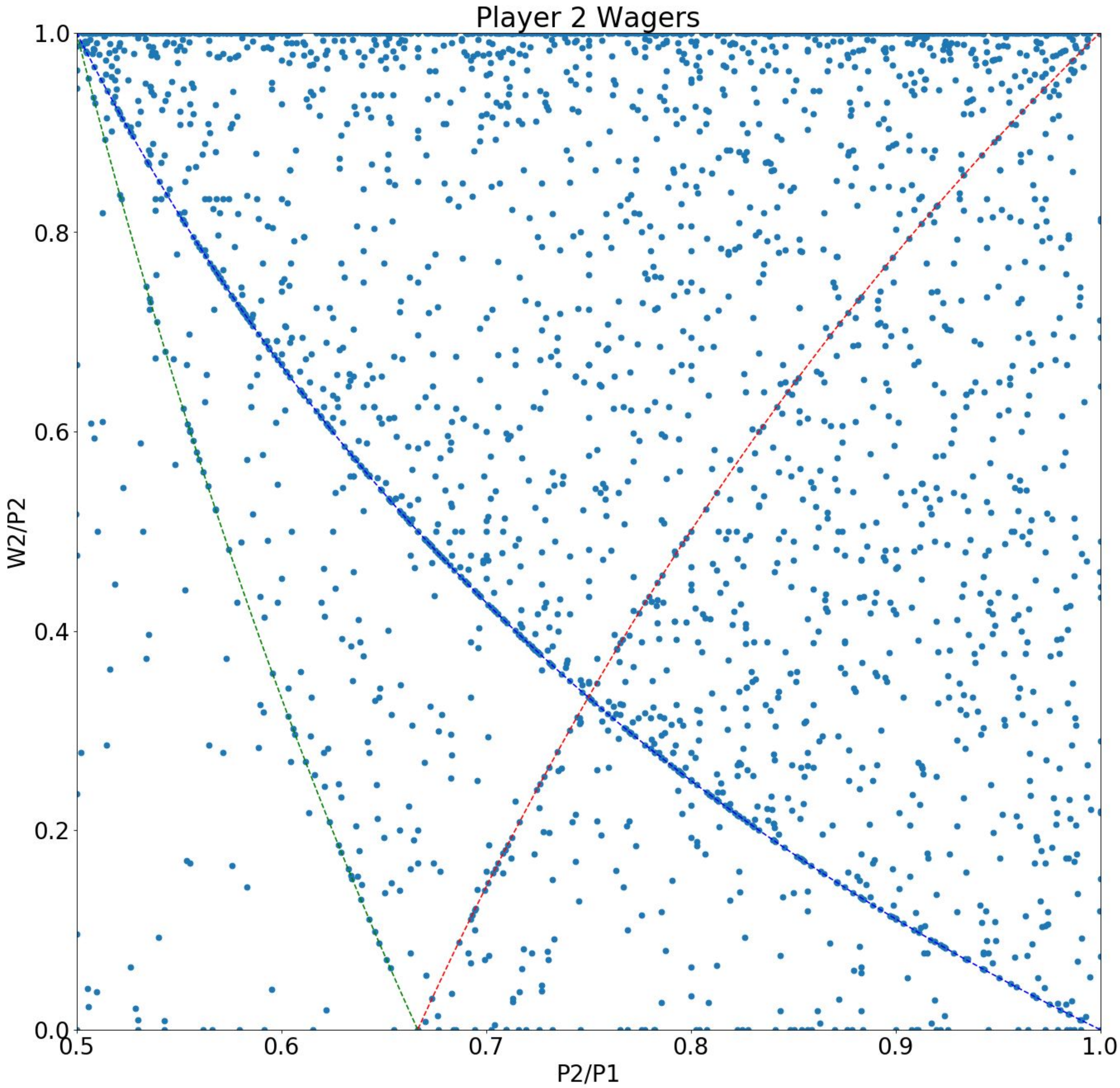


Figure 2: Plot of Player 2 Wagers

The Wager Spacer-

In the simplified, two player model of the Final Jeopardy, the wagering space can be represented as a rectangle with four distinct regions- R1, in which Player 2 wins only if they get the answer right regardless of how Player 1 answers, R2 where Player 2 wins if P1 answers incorrectly regardless of how Player 2 answers, R3 in which Player 2 only wins if they answer correctly while Player 1 answers incorrectly, and R4, where Player 1 wins regardless of how either player answers. These regions are shown in figure X. For Player 2, R1 and R2 are the regions with the highest probability of winning, so the strategic game is a matter of trying to end up in one of those two regions, not R3 or R4.

	P1 W, P2 W	W, R	R,W	R,R
R1	P1	P2	P1	P2
R2	P2	P2	P1	P1
R3	P1	P2	P1	P1
R4	P1	P1	P1	P1

Table Showing who wins in each region for each combination of the two players getting the right or wrong answer.

Walk through an example? Could also give stats on how often games end up in the different regions.

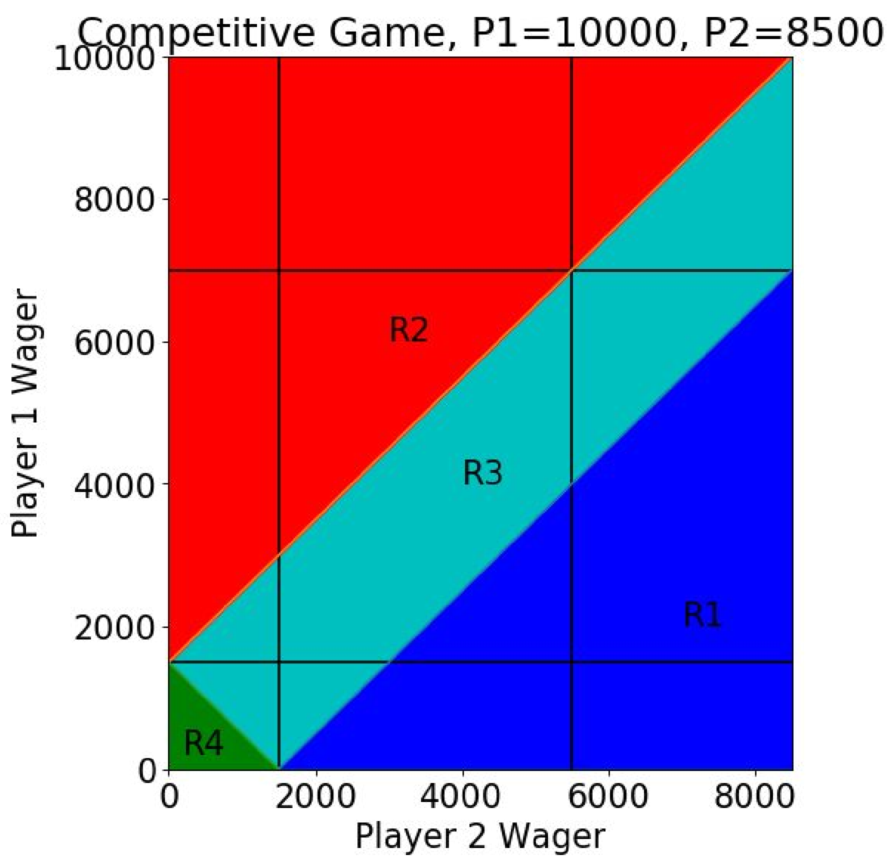


Figure 3: The Wager Space

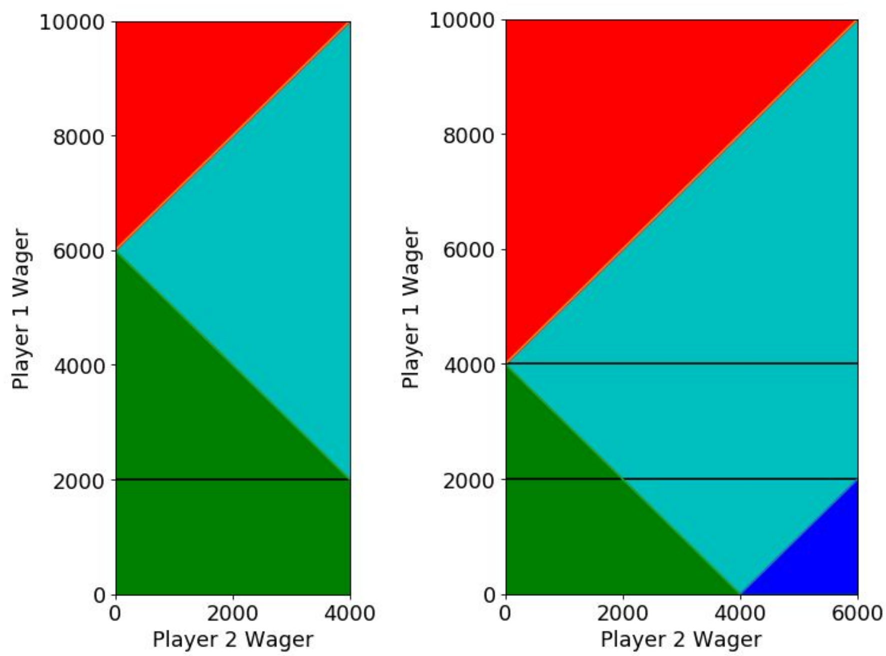
(Right, I didn't at all explain how the vertical and horizontal lines the above plot correspond to different wagering strategies, and how their counter strategies are found by where they intersect with the diagonal region border lines. At the very least should note that the two vertical lines in the above region plot correspond to the blue and red curved lines in figure 2)

Runaway and Near Runaway Games:

A game of jeopardy is called a "runaway" when player 1 has more than double player 2's score (P2/P1<1/2), meaning that with the proper wager (W1<P1-2*P2) player 1 is guaranteed to win regardless of how any of the other contestants answer the question. Free of any competitive considerations the expectation is that player 1 selects a wager within this range based on how confident they are of answering correctly, and this is what we see in the data from the 1312 runaway games in the data set. Player 1 answers correctly 50.4% of the time, and the average wager is 53% of the way between 0 and their maximum safe bid. Furthermore, in a logit model [def add a footnote for this one] with the answer outcome as the dependent variable, the wager % is the most significant explanatory variable at over 99% significance, indicating that players are wagering their perceived probability of answering correctly based on the revealed category.

Maybe discuss why it's not irrational to wager above that heuristic, and that the rarity of it is evidence for heuristics over rationality.

Wagering above the maximum safe wager is not irrational if one places a high value on winning more cash relative to the perceived value of winning the game (and thus winning the opportunity to play another game), a quantity we shall denote as V. In only 18 (1.4%) runaway games did P1 wager above the maximum safe wager, implying either that the value of V swamps the value of the cash wagers, or that players are just defaulting to the most obvious available heuristic.



Wager Spaces for Runaway (left) and Near Runaway Games.

A near runaway is a game in which player 1 can wager such that regardless of how the other players wager, player 1 will only lose if they get it wrong while one of the other players gets it right. This occurs when Player 2 has between ½ and ⅓ of player 1s score. In terms of the wager discussed above, in a runaway game player 1 can unilaterally keep the game in R4, and in a near runaway they can force R3. X% of the games are runaways, and Y% are near runaways.

In both runaway and near runaway games player 1 has a range of wagers in which the probability of winning does not change (because the region is the same), so the players are free to wager based on their expectation of getting the question correct, that is they bet more when they expect to answer correctly based on the revealed category, and less when they expect to get it wrong. And for Runaway games this is indeed what we observe, they wager across the full possible range, and their wager is strongly predictive of whether they get it right or not. But in near runaway games this is not what we see at all. As shown below in figure X, most of the wagers for near runaway games are right at the bottom of their range.

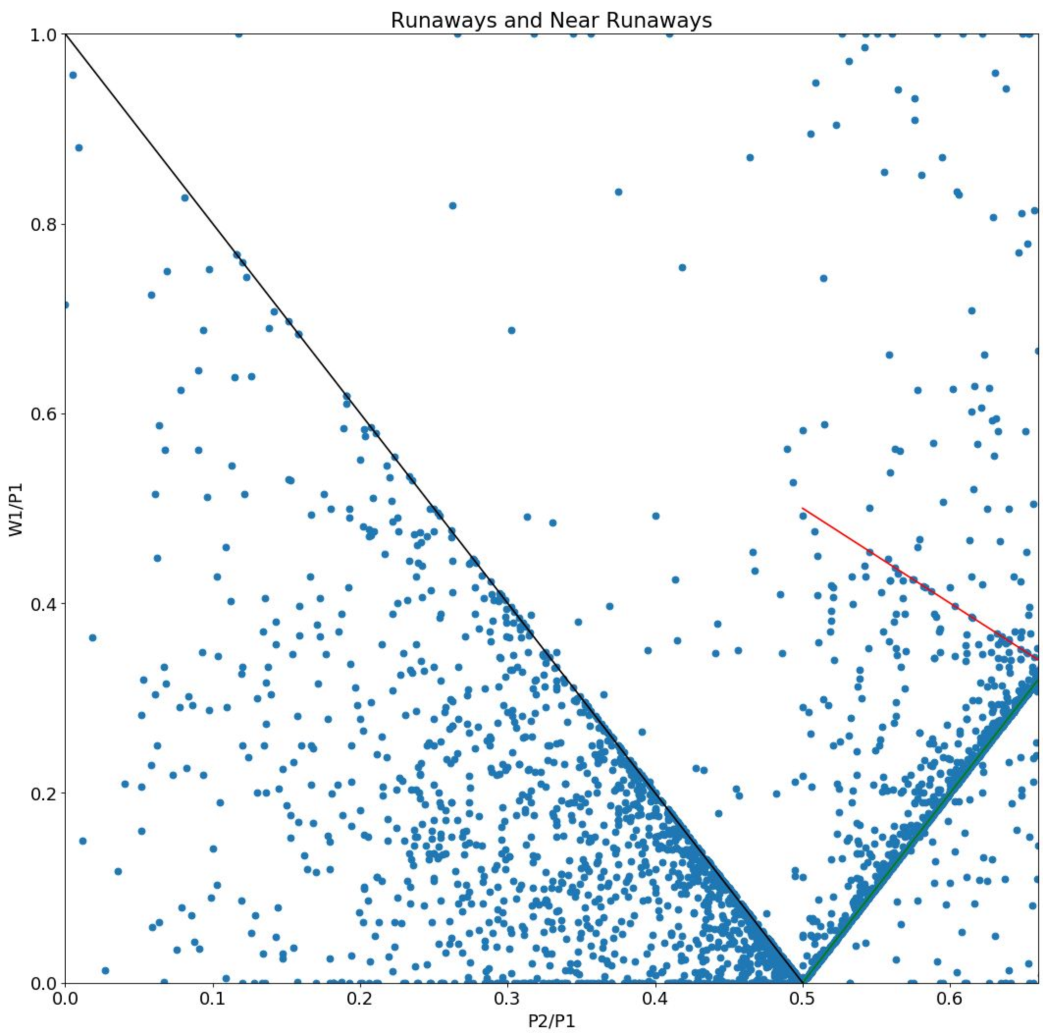


Figure X- Consider adding color to indicate when they lost.

It seems like in near runaway player 1 does not know they can wager up to the red line, and instead the “wager enough to cover player 2 doubling up” heuristics dominates their decision making. If the players were “rational” this would not happen, but of course the contestants are human beings, not calculating machines.

It is interesting to note that if player 1s in near runaway games had the same risk preferences as the player 1s in runaway games, then we would expect to see them wagering closer to the max of the range than the minimum, the opposite of what we see in the data. This is because when they answer incorrectly there’s a chance player 1s wager within the range will be irrelevant (if player 2 wagered correctly and large enough), so player 1 in these games doesn’t pay the cost for a larger wager.

[do I need to explain this more? Show the different expected values? Yea I think so, consider moving the following to a footnote or something]

$EV\_runaway=(Pr(1)-Pr(0))*W+P1+V=0.007*W+P1+V$

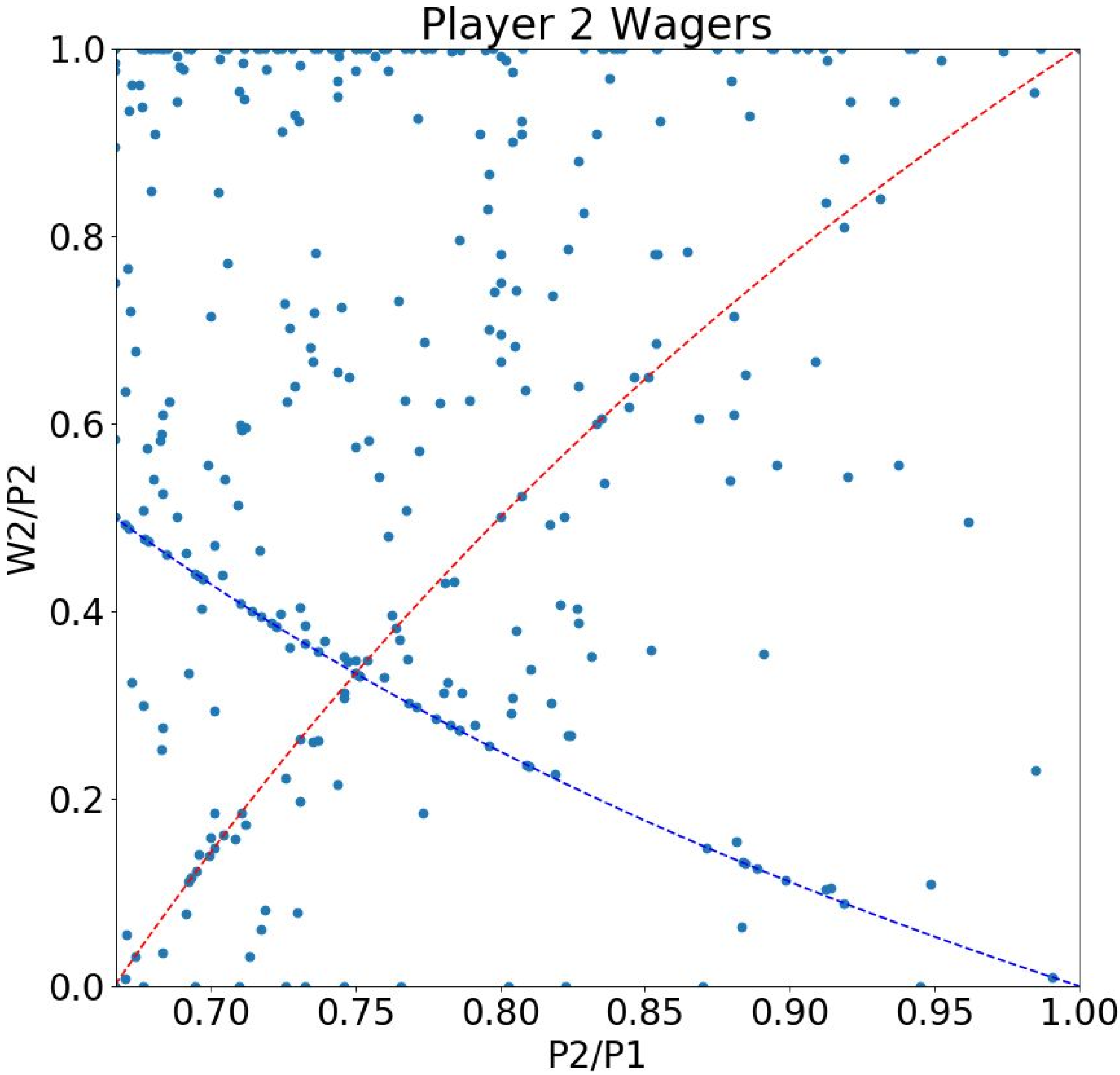
$EV\_NR=(Pr(1)-Pr(01))*W+P1+(1-Pr(01))*V=0.32*W+P1+0.8*V$

[But does just having the larger coefficient on W prove that W should be larger?]

Competitive Two Player Games:

In the intro I claimed that the observation/theory that motivated this all in the first place is that Player 2 is wagering too much too often. To investigate this claim let’s limit our analysis to non-runaway or near runaway games in which player 3 is so far behind as to not be a factor, defined here as P3 being so low that Player 1 can wager to cover Player 2, get it wrong, and still not fall behind a player 3 that bet everything and got the question correct. 358 games meet this criteria, about 9% of the total data set.

In these competitive games, player 1 ends up winning 74%, and player 2 wins the rest (except for two games in which player 3 won because player 1 bet too much and got it wrong). In terms of the regions defined above, 63% are in R3, 31.5% in R2, 5% in R1, and one game in R4.



The above figure shows the wagers for player 2 in competitive games, the red line is the division between big/small wagers as defined above, and the blue line represents how much player 2 needs to wager to get above player 1s current score.

To simplify the analysis, let's assume that each player is only trying to decide between a large or a small wager. For player 1, a wager is large if it is enough to cover player 2 wagering everything (2*P2-P1), and small wager is anything less than that. Then for Player 2, a large wager is such that if Player 1 wagered large, the game ends up in R3 (W2>3*P2-2*P1). The possible outcomes in terms of the wager regions are as follows

	Player 2 Low	Player 2 High

Player 1 Low	R3	R1
Player 1 High	R2	R3

If we assume for a moment that all that the contestants care about is winning the game, (and not their cash amount), then the payoff matrix based on the observed probability of players answering correctly is

	Player 2 Low	Player 2 High
Player 1 Low	0.85	0.5
Player 1 High	0.57	0.85

The Nash Equilibrium for the above payoff matrix is for Player 1 to wager low 44% of the time, high 56% of the time, and for Player 2 to do the opposite, low 56%, high 44%. At this equilibrium player 1 is expected to win 70% of the games, lower than the 74% of games that they actually won. So player one is overperforming, indicating that player 2 is the one making errors.

In the sample of 358 competitive games, player 1 wagers large 87% of the time, and player 2 wagers large in 70% of games, very different from the Nash equilibrium calculated above. This brings us to the key question of this paper, why is player 2 wagering large so frequently? Given that player 1 covers player 2 doubling up nearly 90% of the time, why does player 2 not react by wagering small more often?

One possibility is that the presence of player 3 is influencing how player 2 is wagering. The below plot indicates that there are games where player 2 is selecting their wager based on the current score of player 3, note how often a wager is right on the black line (analogous to the runaway games above, this line represents the wager at which player 2 will fall below player 3 if player 3 wagers everything and gets it right while player 2 answers incorrectly. Given the relatively small difference between the 2nd and 3rd place prizes ($2000 and $1000) this shouldn't really be a factor in player 2s wagering strategy, but apparently in some cases it is. Player 2 could be using this as an anchor for their wagers, playing "safe" against player 3 while inadvertently making a "large".

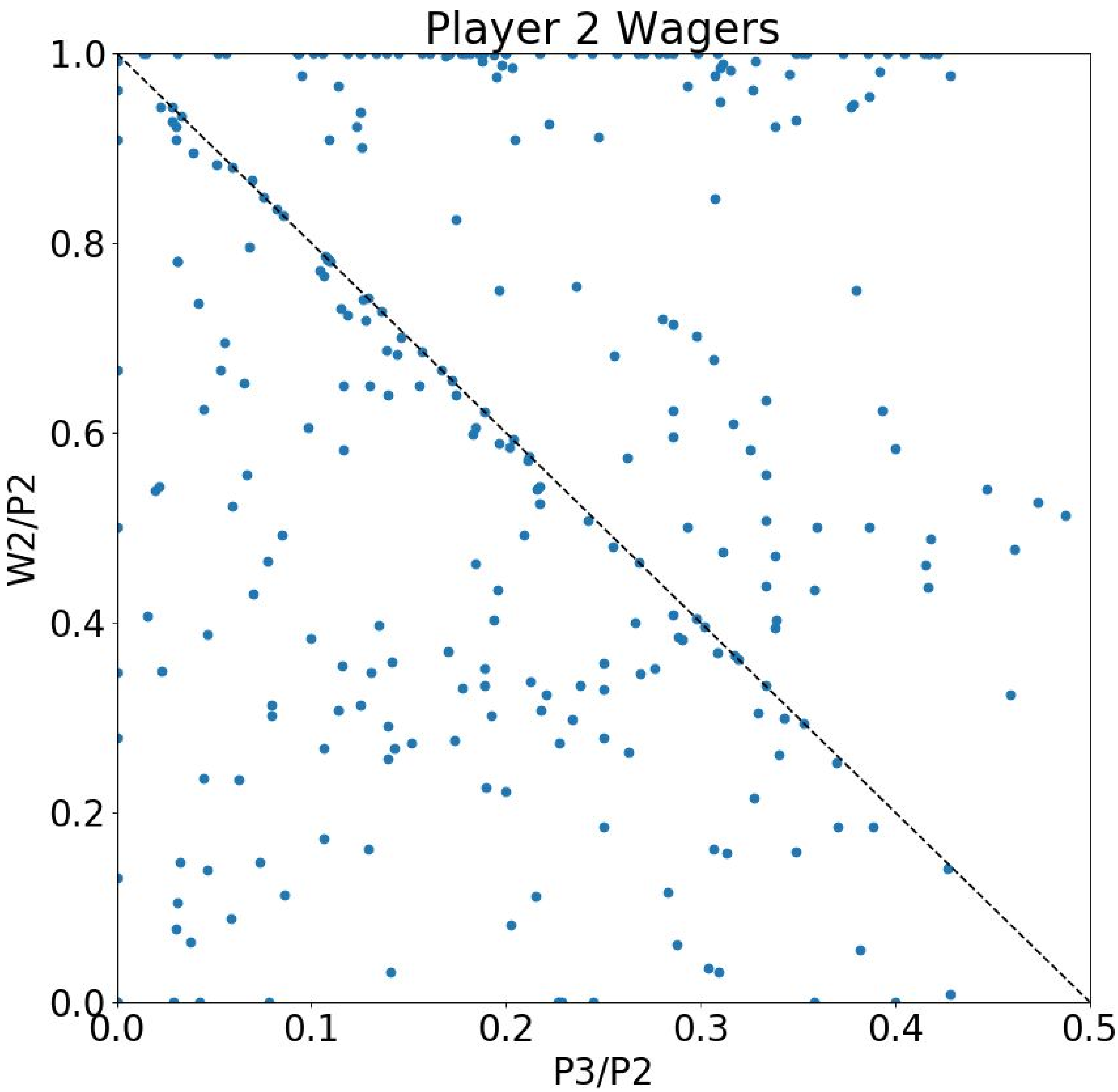


Figure X: Wagers of Player 2 in “Competitive” Games, viewed against P3

[Analyze % high/low for each player by year, see if there are any trends?]

Using the same logit regression modelling technique used in analyzing the runaway/near runaway games, we can see that player 2’s wager is statistically significant in whether they get the FJ question correct, meaning their wager amount is influenced by their perception of the difficulty. Interestingly player 1’s wager is not significant in modelling their likelihood of answering correctly. It is possible/likely that player 2 is failing to consider that when they answer correctly, there is a 73% of player 1 also answering correctly. (This paragraph doesn't seem like enough)

Perhaps there is a deeper fundamental flaw in expecting the average actions of hundreds of different players in these competitive games to resemble a mixed-strategy nash equilibrium, which would require the actions of hundreds of different players in aggregate to roughly match the prescribed split between high and low wagers. For example, the Nash Equilibrium described above has player 2 wagering low 56% of the time, high 44%, so each contestant in the player 2 position in a competitive game would need to randomly select between those two options with those given probabilities. Humans are notoriously bad at randomizing, and numerous behavioral factors could bias them towards playing the high wager option, such as overconfidence. Also in terms of player 2's likelihood of winning, there isn't that much of a difference between the current win rate, 25%, and the 30% percent calculated above if they were to play optimally, so there may simply not be enough pressure to overcome the bias for large wagers.

Conclusion:

Making a strategic choice in a stressful situation like the final round of Jeopardy is difficult, so it isn't surprising that players seem to use a set of heuristics to simplify their decision. Usually these heuristics work just fine, but sometimes they can lead players to make suboptimal decisions, as illustrated by the comparison of runaway and near runaway games. Heuristics don't tell the whole story though, it is difficult to use them to explain the persistent tendency of second place players in competitive games wagering large amounts. This is a bit disappointing, since this is the phenomenon that motivated this analysis in the first place. Regardless, I hope this was at least interesting, and maybe I'll clean it up into a more proper paper.