Results comparing mean achieved value against total occupancy and initial underdog competition value

February 27, 2025

1 Introduction

In the experiments, we search a very large space in order to look for interesting results. We are varying over a wide range of variables, which makes it quite difficult to interpret the data we're getting back from the experiments.

1.1 Experiment variables

Below are all the variables we vary and the values they can take:
System-level variables

1. *total_occupancy_high*: this variable is used to determine what the target for the number of admissions is per institution. This number represents the the total amount of admissions across all institutions as a ratio of the total number of candidates which are evaluated as 'high' on conventional metrics.
   possible values: [0.8, 0.9, 0.95, 0.98, 0.99, 1, 1.01, 1.02, 1.05, 1.1, 1.2]

2. *game_modes*: this variable determines whether we're using the average value of attendees or only taking the top-k for our comparison
   possible values: [expected, top_k]

Player-level variables

1. *win_values_underdog*: represents the probability that in the event of a candidate is invited by both institutions that the underdog institution is chosen over the favourite.
   possible values: [0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5]

2. *fav_blind*: represents if the favourite is able to distinguish between metrics (if True then it can't distinguish).
   possible values: [True, False]

3. *naivety*: represents if the favourite is naive, or if they're also able to best respond until equilibrium is found.
   possible values: [True, False]

Category-level variables

1. *log_or_normal*: used to determine if the probability distribution of value of a candidate belonging to this category will be drawn from a lognormal or a normal distribution.
   possible values: [log, normal]

2. *percent_pop_high_mean*: what percentage of the total population is rated 'high' on conventional metrics.
   possible values: [0.1, 0.2, 0.3, 0.4]

3. *high_low_ratio_mean*: the ratio between $\mu_{high} : \mu_{low}$.
   possible values: [1.2, 1.4, 1.6, 1.8, 2.0]

4. *percent_pop_high_variance*: what percentage of the total population is rated 'high' on unconvetional metrics.
   possible values: [0.1, 0.2, 0.3, 0.4]

5. *high_low_ratio_std*: the ratio between $\sigma_{high} : \sigma_{low}$.
   possible values: [1.2, 1.4, 1.6, 1.8, 2.0]

6. *mean_std_ratios*: the ratio between $\mu_{default} : \sigma_{default}$.
   possible values: [0.5, 0.75, 1, 1.25, 1.5]

These are all of the variables we vary over. In addition, we transform μ, σ when in the *lognormal* setting to prevent the numbers from overflowing but preserving some desirable characteristics. We fix the total number of students at (or slightly below due to rounding in some circumstances) 120, $k = 0.2 *$ desired admits and $\mu_{default}$ as 10.

2 Observations

We will go through some observations from the early run and discuss what needs to be changed.

2.1 Sometimes, we're attaining negative reward

Our reported figure was $result = \frac{reward_{underdog}}{\sum reward_{player}} -$(desired candidates - real candidates)^2. The penalty term was introduced to motivate aiming for the exactly correct number of students, but has resulted in undesirable behaviour in extreme cases where variance can be very large In the case when $\forall\, player \in Players$, $reward_{player} \geq 0$, we would expect $result \in [0, 1]$. However, in some cases we are getting values outside of this range, so we investigate and find the following:

```
Rows with negative underdog_mean: 1921
Percentage of total: 0.30%

Rows with underdog_mean > 1: 2248
Percentage of total: 0.35%

Crosstab of game_mode vs lognormal for combined
   conditions:
lognormal    log  normal  Total
game_mode
expected    1016     510   1526
top_k       2398     245   2643
Total       3414     755   4169
```

The first suspicion was that either agent receiving a negative reward should be due to the penalty term, however, this would not make the *lognormal* categories and *top_k* game mode more likely - **it should be noted that this is exactly the game type that should yield the greatest variance, and that overall it's still a rare feature in all experiments**. Standard deviations are calculated alongside mean values based on the final result, and consequently we'd expect these values to also be between [0,1] generally speaking. However, upon investigating, not a single case has such standard deviations. Below we show a subset of the unique values of the standard deviations, then try to filter for $[std < 5]$.

```
[102.13410644  40.44831902  60.93773101 ...
   43.43851227  27.13159925
  66.47319649]

Filtered rows with negative underdog_mean and std <=
   5:
Empty DataFrame
```

At this moment I'm not quite sure what's causing this behaviour and it will need further investigation. I have 2 ideas for what to check at the moment:

1. Eliminate the penalty term in the evaluation altogether since it's not being used to run gradient updates or anything like this.

2. Try to find what variable these terms are concentrated in by slicing the data according to all the different variables.

2.2 There's a strong tendency towards the neighbourhood of 0.5

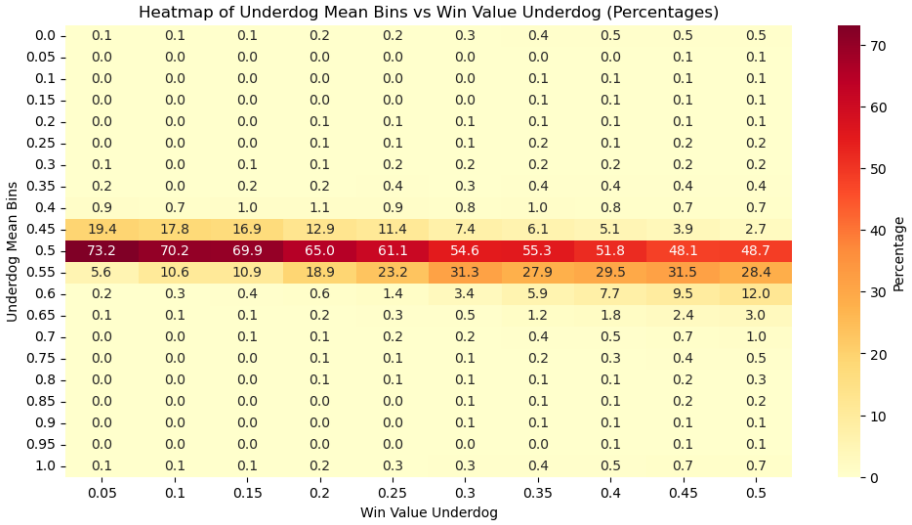


Figure 1: This image shows for each initial *win value underdog* what percentage of games end in a particular mean value bin. This is across all settings.

It should not necessarily be surprising that the underdog is actually coming out ahead in many of these cases considering the underdog never plays naively, nor blind. *It does need to be noted that these runs didn't include the cases where the total occupancy in high was greater than 1, which is where we expect the underdog to start having a natural disadvantage.*

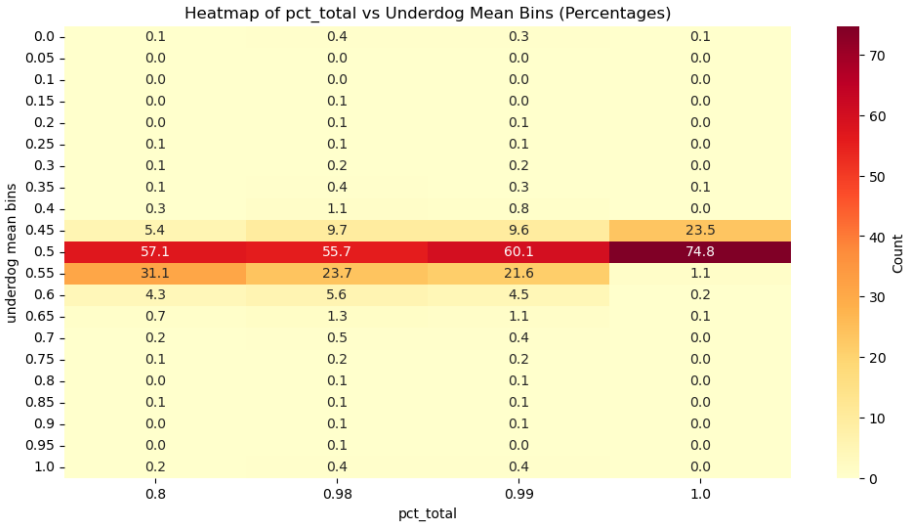


Figure 2: In this image we can see the underdog being able to exploit the favourite specifically in the cases where *total occupancy* is smaller.

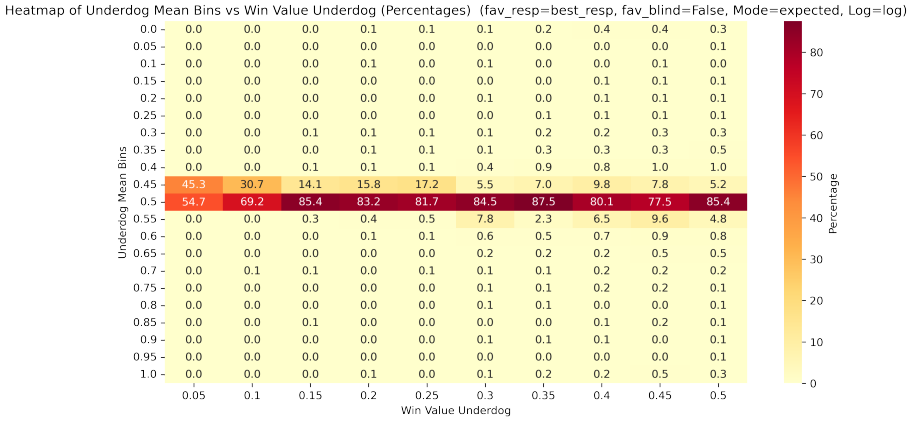
2.3 Getting a feel for standard deviation in different settings

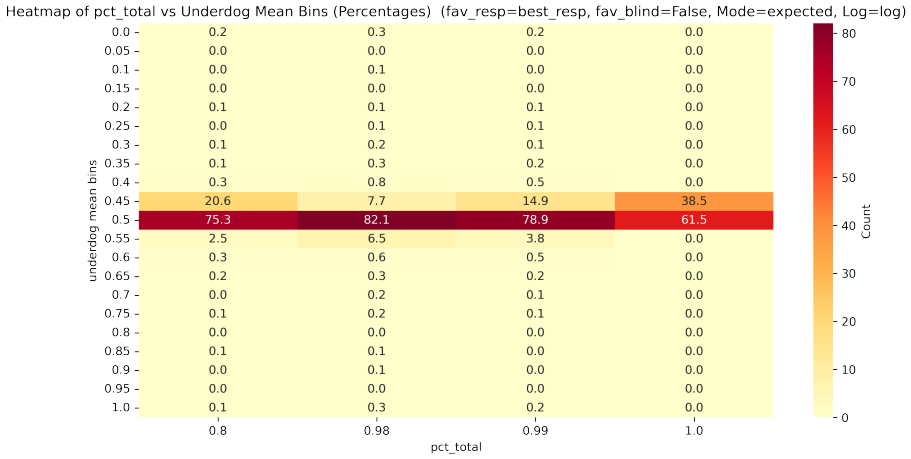
An idea that we've touched on in passing is that the standard deviation really varies based on the different categories, particularly as we're looking at long-tail events in top-k and lognormal conditions. To this end we construct a matrix to see see the average standard deviation (after filtering out the extreme values we saw earlier in 2.1). The result is:

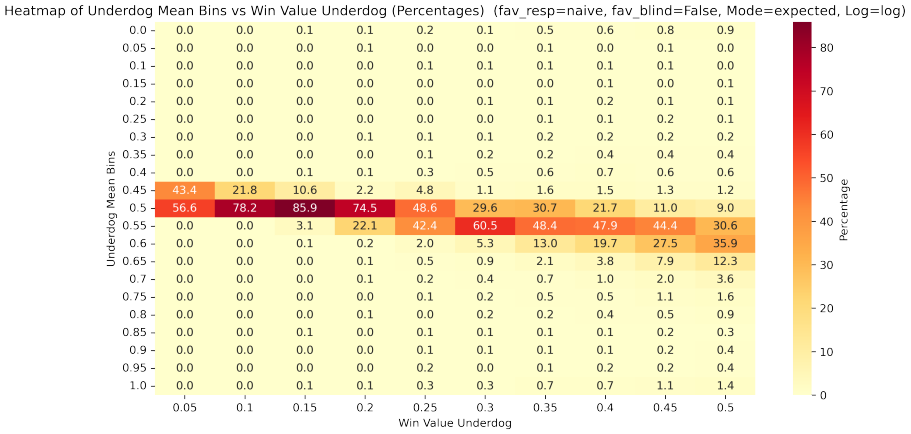
```
Average underdog_std by game_mode and lognormal:
lognormal        log    normal
game_mode
expected    0.464763  0.264238
top_k       0.816327  0.191654
```

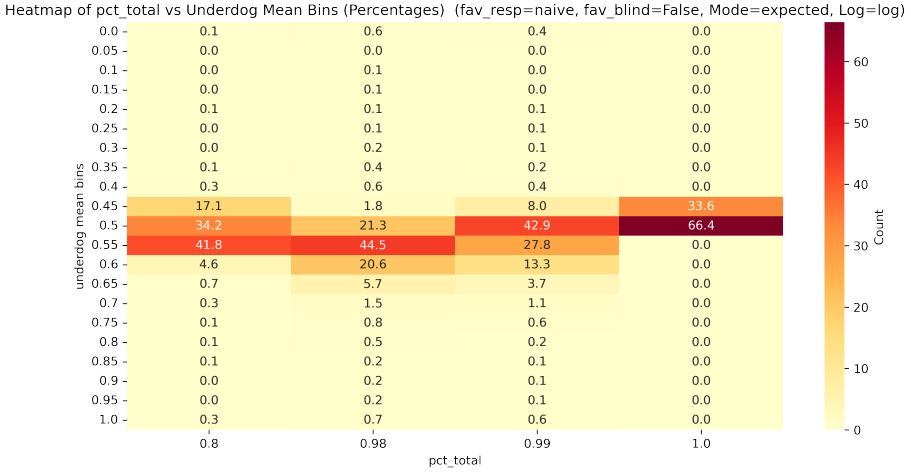
As expected *top-k* and *lognormal* have the highest deviations, in fact they are so large that any sort of response is somewhat meaningless, as in any one particular run almost any outcome can conceivably occur.

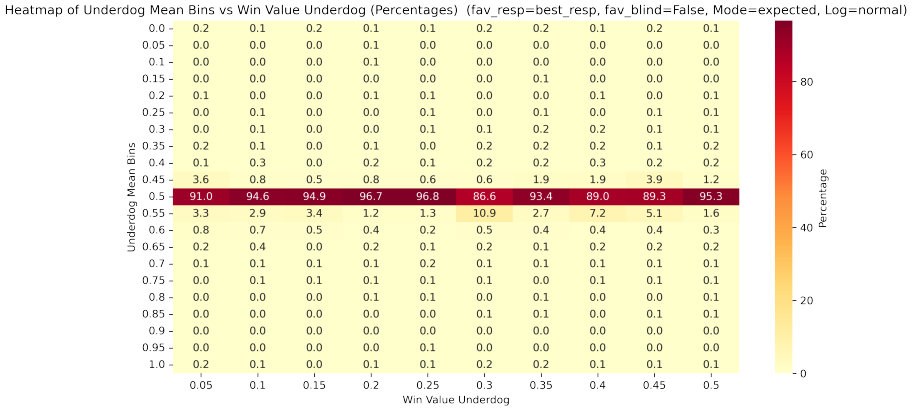
3 Optional: heatmaps that give more granular understanding of the particular scenarios, mostly for reference

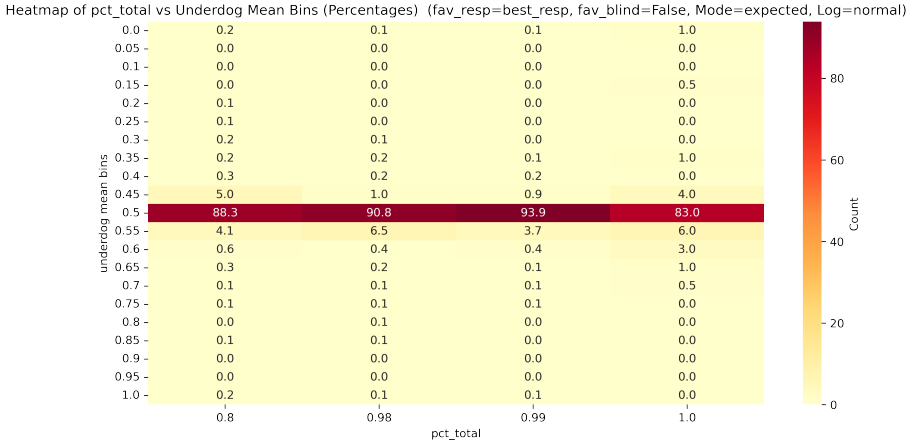


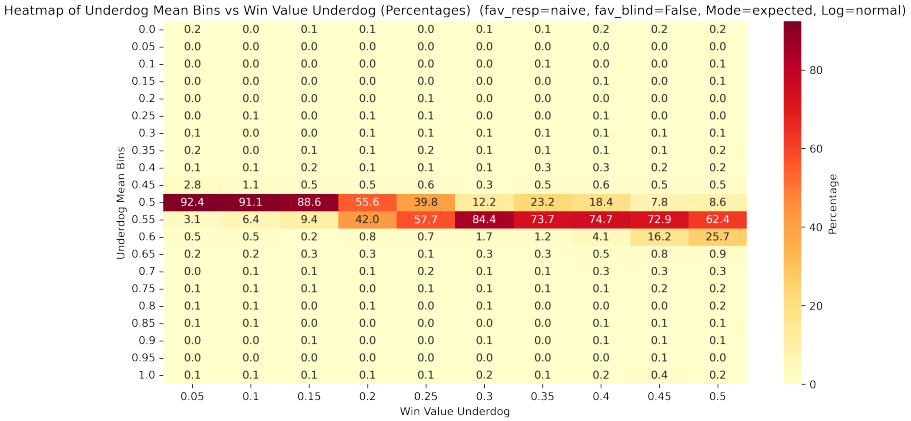


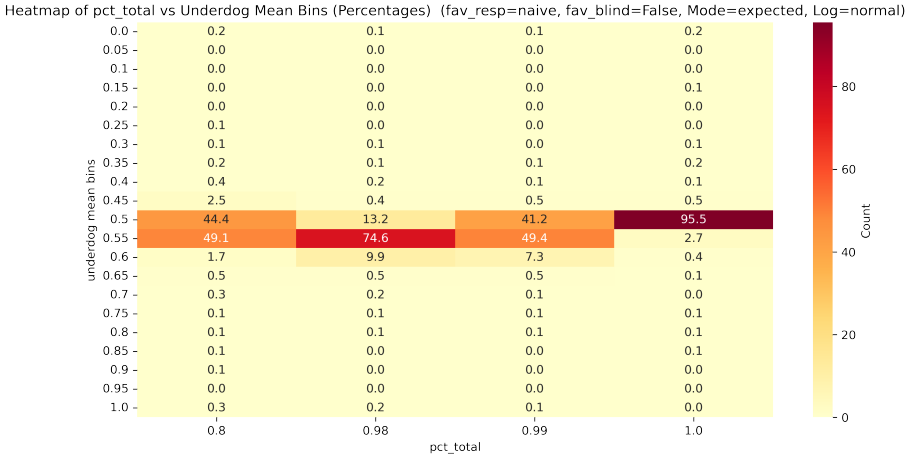

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=naive, fav_blind=False, Mode=expected, Log=log)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.0 0.0 0.1 0.1 0.2 0.1 0.5 0.6 0.8 0.9
0.05 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.0 0.1 0.0
0.1 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.0
0.15 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1
0.2 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.2 0.1 0.1
0.25 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.2 0.1
0.3 0.0 0.0 0.0 0.1 0.1 0.1 0.2 0.2 0.2 0.2
0.35 0.0 0.0 0.0 0.0 0.1 0.2 0.2 0.4 0.4 0.4
0.4 0.0 0.0 0.1 0.1 0.3 0.5 0.6 0.7 0.6 0.6
0.45 43.4 21.8 10.6 2.2 4.8 1.1 1.6 1.5 1.3 1.2
0.5 56.6 78.2 85.9 74.5 48.6 29.6 30.7 21.7 11.0 9.0
0.55 0.0 0.0 3.1 22.1 42.4 60.5 48.4 47.9 44.4 30.6
0.6 0.0 0.0 0.1 0.2 2.0 5.3 13.0 19.7 27.5 35.9
0.65 0.0 0.0 0.0 0.1 0.5 0.9 2.1 3.8 7.9 12.3
0.7 0.0 0.0 0.0 0.1 0.2 0.4 0.7 1.0 2.0 3.6
0.75 0.0 0.0 0.0 0.0 0.1 0.2 0.5 0.5 1.1 1.6
0.8 0.0 0.0 0.0 0.1 0.0 0.2 0.2 0.4 0.5 0.9
0.85 0.0 0.0 0.1 0.0 0.1 0.1 0.1 0.1 0.2 0.3
0.9 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.2 0.4
0.95 0.0 0.0 0.0 0.0 0.2 0.0 0.1 0.2 0.2 0.4
1.0 0.0 0.0 0.1 0.1 0.3 0.3 0.7 0.7 1.1 1.4
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 10 20 30 40 50 60 70 80

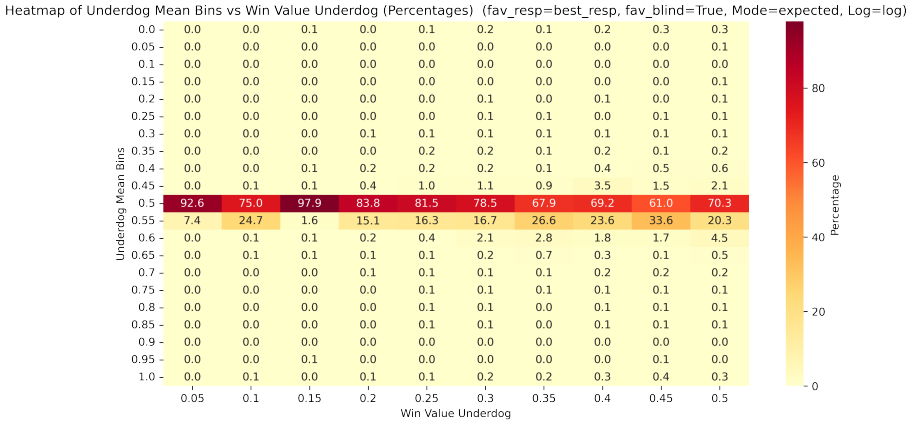

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=naive, fav_blind=False, Mode=expected, Log=log)
underdog mean bins
pct_total
Count
0.0 0.1 0.6 0.4 0.0
0.05 0.0 0.0 0.0 0.0
0.1 0.0 0.1 0.0 0.0
0.15 0.0 0.1 0.0 0.0
0.2 0.1 0.1 0.1 0.0
0.25 0.0 0.1 0.1 0.0
0.3 0.0 0.2 0.1 0.0
0.35 0.1 0.4 0.2 0.0
0.4 0.3 0.6 0.4 0.0
0.45 17.1 1.8 8.0 33.6
0.5 34.2 21.3 42.9 66.4
0.55 41.8 44.5 27.8 0.0
0.6 4.6 20.6 13.3 0.0
0.65 0.7 5.7 3.7 0.0
0.7 0.3 1.5 1.1 0.0
0.75 0.1 0.8 0.6 0.0
0.8 0.1 0.5 0.2 0.0
0.85 0.1 0.2 0.1 0.0
0.9 0.0 0.2 0.1 0.0
0.95 0.0 0.2 0.1 0.0
1.0 0.3 0.7 0.6 0.0
0.8 0.98 0.99 1.0
0 10 20 30 40 50 60

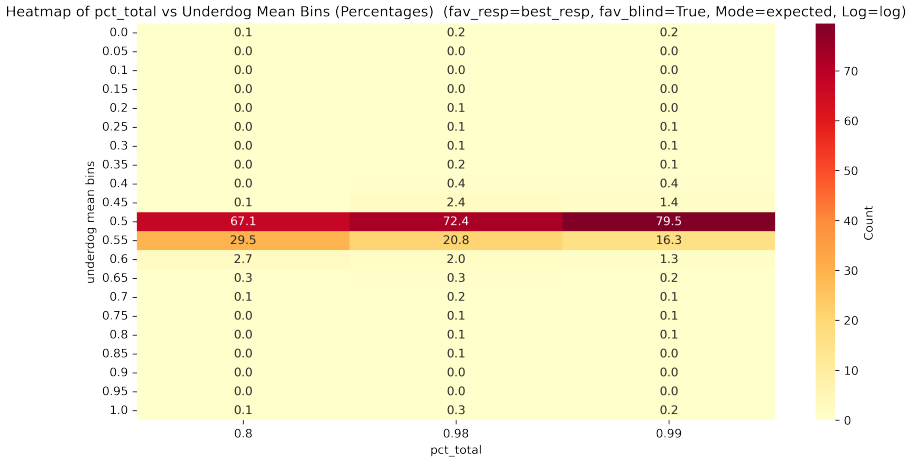

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=best_resp, fav_blind=False, Mode=expected, Log=normal)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.2 0.1 0.2 0.1 0.1 0.2 0.2 0.1 0.2 0.1
0.05 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0 0.0 0.0
0.15 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0
0.2 0.1 0.0 0.0 0.1 0.1 0.0 0.0 0.1 0.0 0.1
0.25 0.0 0.1 0.0 0.0 0.1 0.0 0.1 0.0 0.1 0.1
0.3 0.0 0.1 0.0 0.0 0.1 0.1 0.2 0.2 0.1 0.1
0.35 0.2 0.1 0.0 0.1 0.0 0.2 0.2 0.2 0.1 0.2
0.4 0.1 0.3 0.0 0.2 0.1 0.2 0.2 0.3 0.2 0.2
0.45 3.6 0.8 0.5 0.8 0.6 0.6 1.9 1.9 3.9 1.2
0.5 91.0 94.6 94.9 96.7 96.8 86.6 93.4 89.0 89.3 95.3
0.55 3.3 2.9 3.4 1.2 1.3 10.9 2.7 7.2 5.1 1.6
0.6 0.8 0.7 0.5 0.4 0.2 0.5 0.4 0.4 0.4 0.3
0.65 0.2 0.4 0.0 0.2 0.1 0.2 0.1 0.2 0.2 0.2
0.7 0.1 0.1 0.1 0.1 0.2 0.1 0.1 0.1 0.1 0.1
0.75 0.0 0.1 0.1 0.1 0.1 0.1 0.0 0.1 0.1 0.1
0.8 0.0 0.0 0.0 0.1 0.1 0.0 0.1 0.0 0.0 0.1
0.85 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.0 0.1 0.1
0.9 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.95 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0 0.0
1.0 0.2 0.1 0.0 0.1 0.1 0.2 0.2 0.1 0.1 0.1
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 20 40 60 80

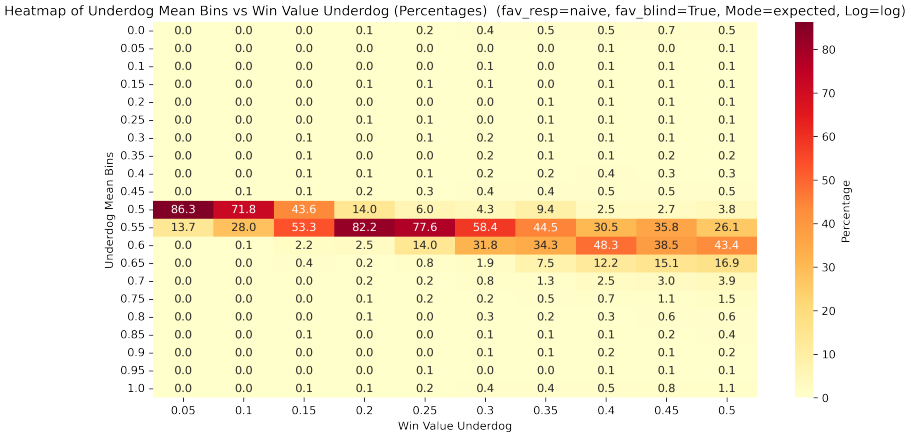

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=best_resp, fav_blind=False, Mode=expected, Log=normal)
underdog mean bins
pct_total
Count
0.0 0.2 0.1 0.1 1.0
0.05 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.0
0.15 0.0 0.0 0.0 0.5
0.2 0.1 0.0 0.0 0.0
0.25 0.1 0.0 0.0 0.0
0.3 0.2 0.1 0.0 0.0
0.35 0.2 0.2 0.1 1.0
0.4 0.3 0.2 0.2 0.0
0.45 5.0 1.0 0.9 4.0
0.5 88.3 90.8 93.9 83.0
0.55 4.1 6.5 3.7 6.0
0.6 0.6 0.4 0.4 3.0
0.65 0.3 0.2 0.1 1.0
0.7 0.1 0.1 0.1 0.5
0.75 0.1 0.1 0.1 0.0
0.8 0.0 0.1 0.0 0.0
0.85 0.1 0.1 0.0 0.0
0.9 0.0 0.0 0.0 0.0
0.95 0.0 0.0 0.0 0.0
1.0 0.2 0.1 0.1 0.0
0.8 0.98 0.99 1.0
0 20 40 60 80

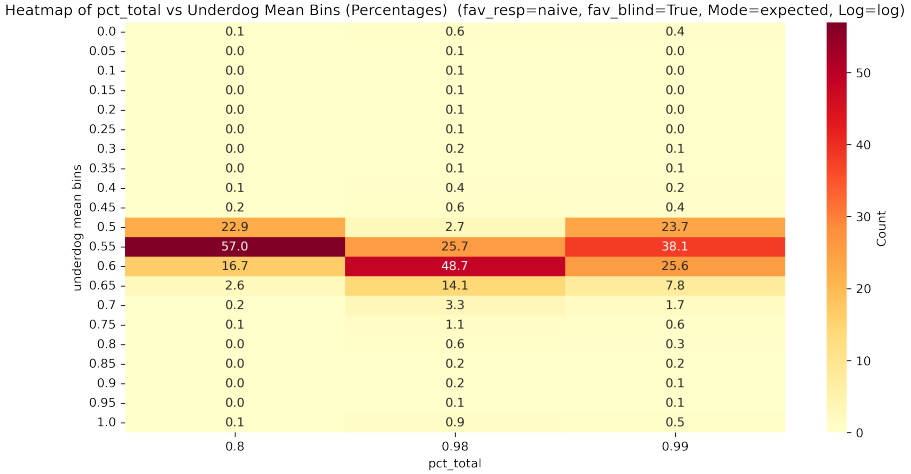

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=naive, fav_blind=False, Mode=expected, Log=normal)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.2 0.0 0.1 0.1 0.0 0.1 0.1 0.2 0.2 0.2
0.05 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.1
0.15 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1
0.2 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0 0.0
0.25 0.0 0.1 0.0 0.1 0.1 0.0 0.0 0.1 0.0 0.0
0.3 0.1 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.1 0.1
0.35 0.2 0.0 0.1 0.1 0.2 0.1 0.1 0.1 0.1 0.2
0.4 0.1 0.1 0.2 0.1 0.1 0.1 0.3 0.3 0.2 0.2
0.45 2.8 1.1 0.5 0.5 0.6 0.3 0.5 0.6 0.5 0.5
0.5 92.4 91.1 88.6 55.6 39.8 12.2 23.2 18.4 7.8 8.6
0.55 3.1 6.4 9.4 42.0 57.7 84.4 73.7 74.7 72.9 62.4
0.6 0.5 0.5 0.2 0.8 0.7 1.7 1.2 4.1 16.2 25.7
0.65 0.2 0.2 0.3 0.3 0.1 0.3 0.3 0.5 0.8 0.9
0.7 0.0 0.1 0.1 0.1 0.2 0.1 0.1 0.3 0.3 0.3
0.75 0.1 0.1 0.1 0.0 0.1 0.1 0.1 0.1 0.2 0.2
0.8 0.1 0.1 0.0 0.1 0.0 0.1 0.0 0.0 0.0 0.2
0.85 0.1 0.1 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1
0.9 0.0 0.0 0.1 0.0 0.0 0.1 0.0 0.1 0.1 0.1
0.95 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0
1.0 0.1 0.1 0.1 0.1 0.1 0.2 0.1 0.2 0.4 0.2
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 20 40 60 80

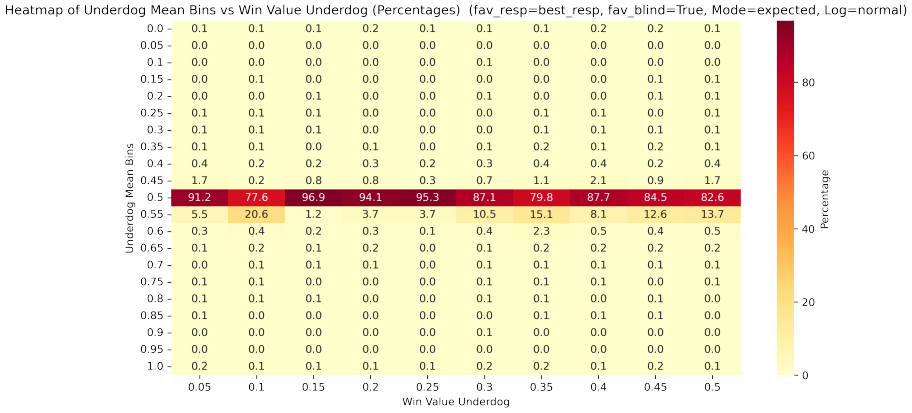

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=naive, fav_blind=False, Mode=expected, Log=normal)
underdog mean bins
pct_total
Count
0.0 0.2 0.1 0.1 0.2
0.05 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.0
0.15 0.0 0.0 0.0 0.1
0.2 0.0 0.0 0.0 0.0
0.25 0.1 0.0 0.0 0.0
0.3 0.1 0.1 0.0 0.1
0.35 0.2 0.1 0.1 0.2
0.4 0.4 0.2 0.1 0.1
0.45 2.5 0.4 0.5 0.5
0.5 44.4 13.2 41.2 95.5
0.55 49.1 74.6 49.4 2.7
0.6 1.7 9.9 7.3 0.4
0.65 0.5 0.5 0.5 0.1
0.7 0.3 0.2 0.1 0.0
0.75 0.1 0.1 0.1 0.1
0.8 0.1 0.1 0.1 0.1
0.85 0.1 0.0 0.0 0.1
0.9 0.1 0.0 0.0 0.0
0.95 0.0 0.0 0.0 0.0
1.0 0.3 0.2 0.1 0.0
0.8 0.98 0.99 1.0
0 20 40 60 80

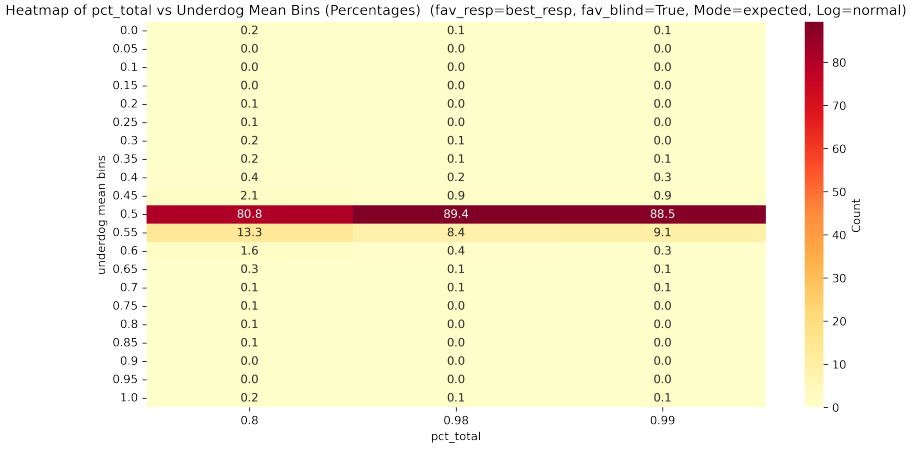

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=best_resp, fav_blind=True, Mode=expected, Log=log)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.0 0.0 0.1 0.0 0.1 0.2 0.1 0.2 0.3 0.3
0.05 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1
0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.15 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1
0.2 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.0 0.1
0.25 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.0 0.1 0.1
0.3 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.1 0.1
0.35 0.0 0.0 0.0 0.0 0.2 0.2 0.1 0.2 0.1 0.2
0.4 0.0 0.0 0.1 0.2 0.2 0.2 0.1 0.4 0.5 0.6
0.45 0.0 0.1 0.1 0.4 1.0 1.1 0.9 3.5 1.5 2.1
0.5 92.6 75.0 97.9 83.8 81.5 78.5 67.9 69.2 61.0 70.3
0.55 7.4 24.7 1.6 15.1 16.3 16.7 26.6 23.6 33.6 20.3
0.6 0.0 0.1 0.1 0.2 0.4 2.1 2.8 1.8 1.7 4.5
0.65 0.0 0.1 0.1 0.1 0.1 0.2 0.7 0.3 0.1 0.5
0.7 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.2 0.2 0.2
0.75 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.1
0.8 0.0 0.0 0.0 0.0 0.1 0.1 0.0 0.1 0.1 0.1
0.85 0.0 0.0 0.0 0.0 0.1 0.1 0.0 0.1 0.1 0.1
0.9 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.95 0.0 0.0 0.1 0.0 0.0 0.0 0.0 0.0 0.1 0.0
1.0 0.0 0.1 0.0 0.1 0.1 0.2 0.2 0.3 0.4 0.3
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 20 40 60 80

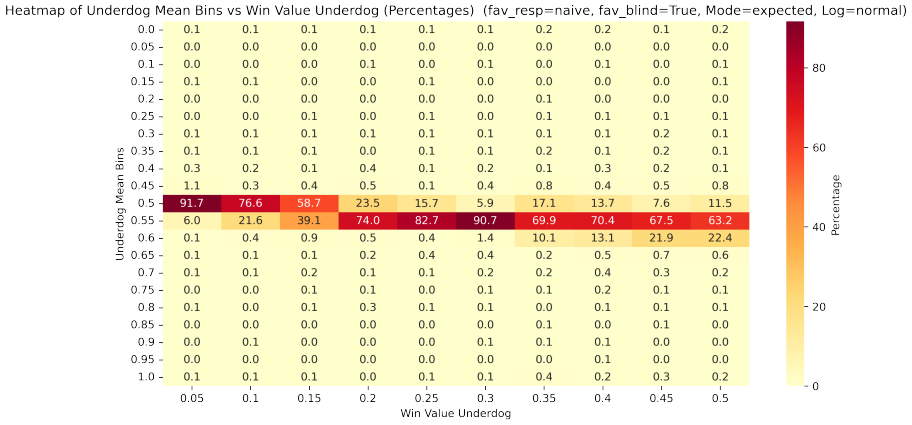

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=best_resp, fav_blind=True, Mode=expected, Log=log)
underdog mean bins
pct_total
Count
0.0 0.1 0.2 0.2
0.05 0.0 0.0 0.0
0.1 0.0 0.0 0.0
0.15 0.0 0.0 0.0
0.2 0.0 0.1 0.0
0.25 0.0 0.1 0.1
0.3 0.0 0.1 0.1
0.35 0.0 0.2 0.1
0.4 0.0 0.4 0.4
0.45 0.1 2.4 1.4
0.5 67.1 72.4 79.5
0.55 29.5 20.8 16.3
0.6 2.7 2.0 1.3
0.65 0.3 0.3 0.2
0.7 0.1 0.2 0.1
0.75 0.0 0.1 0.1
0.8 0.0 0.1 0.1
0.85 0.0 0.1 0.0
0.9 0.0 0.0 0.0
0.95 0.0 0.0 0.0
1.0 0.1 0.3 0.2
0.8 0.98 0.99
0 10 20 30 40 50 60 70

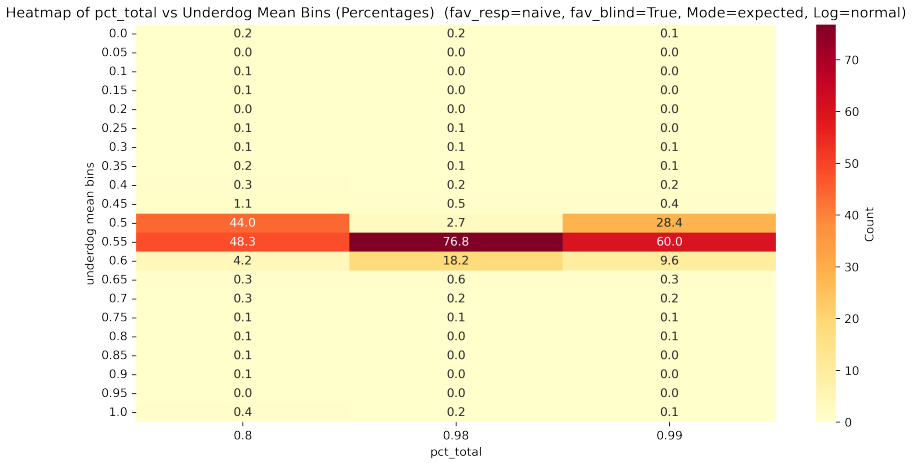

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=naive, fav_blind=True, Mode=expected, Log=log)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.0 0.0 0.0 0.1 0.2 0.4 0.5 0.5 0.7 0.5
0.05 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1
0.1 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.1
0.15 0.0 0.0 0.0 0.1 0.1 0.1 0.0 0.1 0.1 0.1
0.2 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1
0.25 0.0 0.0 0.0 0.1 0.1 0.0 0.1 0.1 0.1 0.1
0.3 0.0 0.0 0.1 0.0 0.1 0.2 0.1 0.1 0.1 0.1
0.35 0.0 0.0 0.1 0.0 0.0 0.2 0.1 0.1 0.2 0.2
0.4 0.0 0.0 0.1 0.1 0.1 0.2 0.2 0.4 0.3 0.3
0.45 0.0 0.1 0.1 0.2 0.3 0.4 0.4 0.5 0.5 0.5
0.5 86.3 71.8 43.6 14.0 6.0 4.3 9.4 2.5 2.7 3.8
0.55 13.7 28.0 53.3 82.2 77.6 58.4 44.5 30.5 35.8 26.1
0.6 0.0 0.1 2.2 2.5 14.0 31.8 34.3 48.3 38.5 43.4
0.65 0.0 0.0 0.4 0.2 0.8 1.9 7.5 12.2 15.1 16.9
0.7 0.0 0.0 0.0 0.2 0.2 0.8 1.3 2.5 3.0 3.9
0.75 0.0 0.0 0.0 0.1 0.2 0.2 0.5 0.7 1.1 1.5
0.8 0.0 0.0 0.0 0.1 0.0 0.3 0.2 0.3 0.6 0.6
0.85 0.0 0.0 0.1 0.0 0.0 0.1 0.1 0.1 0.2 0.4
0.9 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.2 0.1 0.2
0.95 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.1 0.1
1.0 0.0 0.0 0.1 0.1 0.2 0.4 0.4 0.5 0.8 1.1
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 10 20 30 40 50 60 70 80

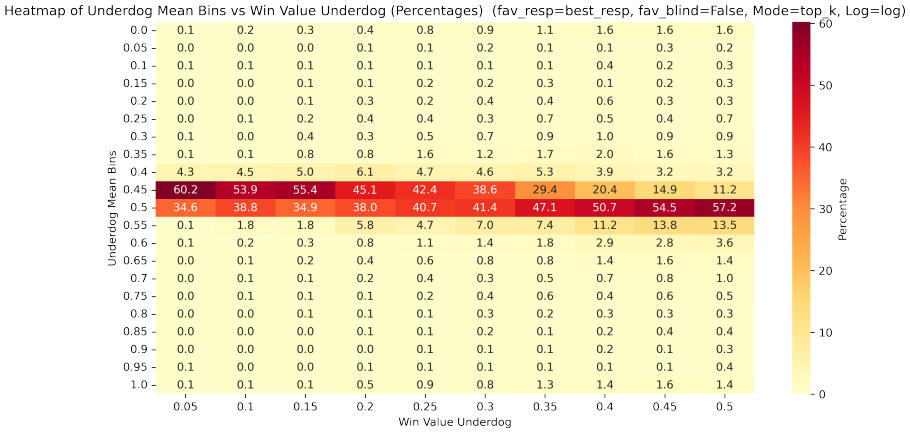

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=naive, fav_blind=True, Mode=expected, Log=log)
underdog mean bins
pct_total
Count
0.0 0.1 0.6 0.4
0.05 0.0 0.1 0.0
0.1 0.0 0.1 0.0
0.15 0.0 0.1 0.0
0.2 0.0 0.1 0.0
0.25 0.0 0.1 0.0
0.3 0.0 0.2 0.1
0.35 0.0 0.1 0.1
0.4 0.1 0.4 0.2
0.45 0.2 0.6 0.4
0.5 22.9 2.7 23.7
0.55 57.0 25.7 38.1
0.6 16.7 48.7 25.6
0.65 2.6 14.1 7.8
0.7 0.2 3.3 1.7
0.75 0.1 1.1 0.6
0.8 0.0 0.6 0.3
0.85 0.0 0.2 0.2
0.9 0.0 0.2 0.1
0.95 0.0 0.1 0.1
1.0 0.1 0.9 0.5
0.8 0.98 0.99
0 10 20 30 40 50

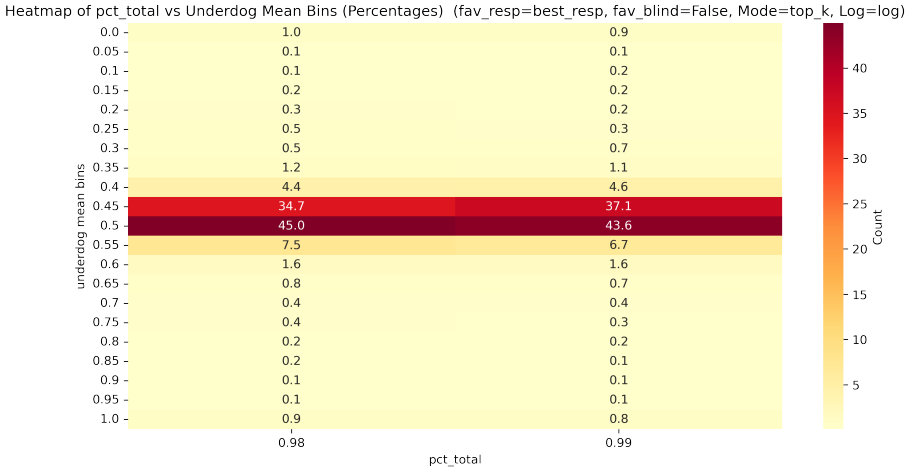

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=best_resp, fav_blind=True, Mode=expected, Log=normal)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.1 0.1 0.1 0.2 0.1 0.1 0.1 0.2 0.2 0.1
0.05 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0
0.15 0.0 0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1
0.2 0.0 0.0 0.1 0.0 0.0 0.1 0.0 0.0 0.1 0.1
0.25 0.1 0.1 0.1 0.0 0.0 0.0 0.1 0.1 0.0 0.1
0.3 0.1 0.1 0.1 0.0 0.0 0.0 0.1 0.1 0.1 0.1
0.35 0.1 0.1 0.0 0.1 0.0 0.1 0.2 0.1 0.2 0.1
0.4 0.4 0.2 0.2 0.3 0.2 0.3 0.4 0.4 0.2 0.4
0.45 1.7 0.2 0.8 0.8 0.3 0.7 1.1 2.1 0.9 1.7
0.5 91.2 77.6 96.9 94.1 95.3 87.1 79.8 87.7 84.5 82.6
0.55 5.5 20.6 1.2 3.7 3.7 10.5 15.1 8.1 12.6 13.7
0.6 0.3 0.4 0.2 0.3 0.1 0.4 2.3 0.5 0.4 0.5
0.65 0.1 0.2 0.1 0.2 0.0 0.1 0.2 0.2 0.2 0.2
0.7 0.0 0.1 0.1 0.1 0.0 0.1 0.1 0.1 0.1 0.1
0.75 0.1 0.1 0.0 0.0 0.0 0.1 0.1 0.1 0.0 0.1
0.8 0.1 0.1 0.1 0.0 0.0 0.1 0.1 0.0 0.1 0.0
0.85 0.1 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.0
0.9 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0
0.95 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
1.0 0.2 0.1 0.1 0.1 0.1 0.2 0.2 0.1 0.2 0.1
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 20 40 60 80

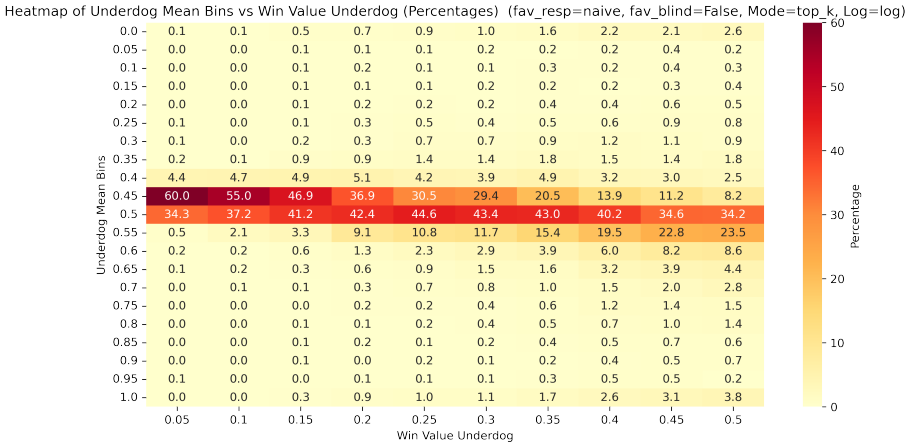

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=best_resp, fav_blind=True, Mode=expected, Log=normal)
underdog mean bins
pct_total
Count
0.0 0.2 0.1 0.1
0.05 0.0 0.0 0.0
0.1 0.0 0.0 0.0
0.15 0.0 0.0 0.0
0.2 0.1 0.0 0.0
0.25 0.1 0.0 0.0
0.3 0.2 0.1 0.0
0.35 0.2 0.1 0.1
0.4 0.4 0.2 0.3
0.45 2.1 0.9 0.9
0.5 80.8 89.4 88.5
0.55 13.3 8.4 9.1
0.6 1.6 0.4 0.3
0.65 0.3 0.1 0.1
0.7 0.1 0.1 0.1
0.75 0.1 0.0 0.0
0.8 0.1 0.0 0.0
0.85 0.1 0.0 0.0
0.9 0.0 0.0 0.0
0.95 0.0 0.0 0.0
1.0 0.2 0.1 0.1
0.8 0.98 0.99
0 10 20 30 40 50 60 70 80

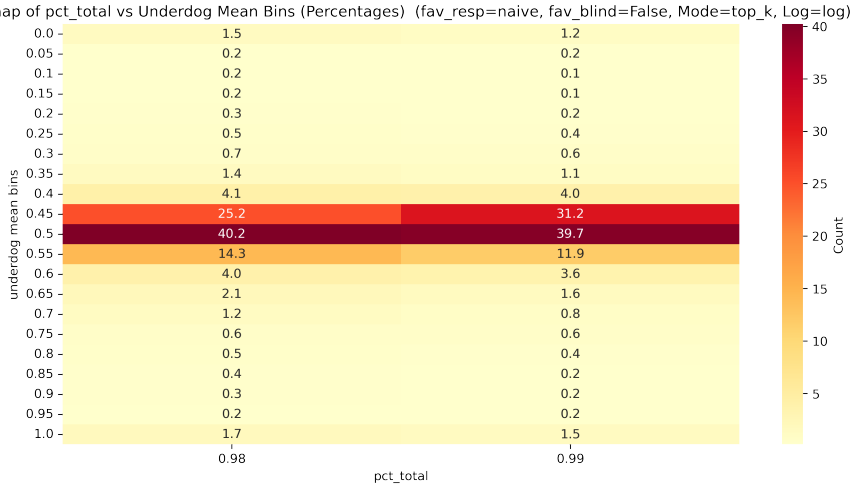

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=naive, fav_blind=True, Mode=expected, Log=normal)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.1 0.1 0.1 0.1 0.1 0.1 0.2 0.2 0.1 0.2
0.05 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.1 0.0 0.1 0.0 0.1 0.0 0.1
0.15 0.1 0.1 0.0 0.0 0.1 0.0 0.0 0.0 0.0 0.1
0.2 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0
0.25 0.0 0.0 0.1 0.0 0.1 0.0 0.1 0.1 0.1 0.1
0.3 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.2 0.1
0.35 0.1 0.1 0.1 0.0 0.1 0.1 0.2 0.1 0.2 0.1
0.4 0.3 0.2 0.1 0.4 0.1 0.2 0.1 0.3 0.2 0.1
0.45 1.1 0.3 0.4 0.5 0.1 0.4 0.8 0.4 0.5 0.8
0.5 91.7 76.6 58.7 23.5 15.7 5.9 17.1 13.7 7.6 11.5
0.55 6.0 21.6 39.1 74.0 82.7 90.7 69.9 70.4 67.5 63.2
0.6 0.1 0.4 0.9 0.5 0.4 1.4 10.1 13.1 21.9 22.4
0.65 0.1 0.1 0.1 0.2 0.4 0.4 0.2 0.5 0.7 0.6
0.7 0.1 0.1 0.2 0.1 0.1 0.2 0.2 0.4 0.3 0.2
0.75 0.0 0.0 0.1 0.1 0.0 0.1 0.1 0.2 0.1 0.1
0.8 0.1 0.0 0.1 0.3 0.1 0.1 0.0 0.1 0.1 0.1
0.85 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.0
0.9 0.0 0.1 0.0 0.0 0.0 0.1 0.1 0.1 0.0 0.0
0.95 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0
1.0 0.1 0.1 0.1 0.0 0.1 0.1 0.4 0.2 0.3 0.2
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 20 40 60 80

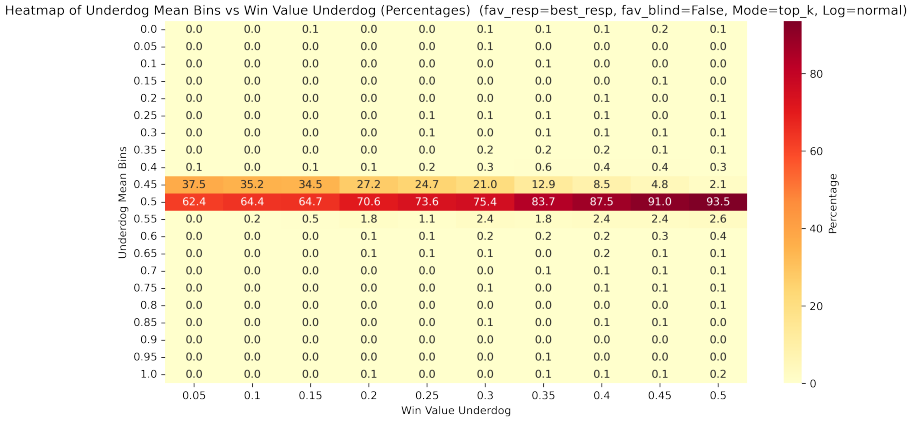

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=naive, fav_blind=True, Mode=expected, Log=normal)
underdog mean bins
pct_total
Count
0.0 0.2 0.2 0.1
0.05 0.0 0.0 0.0
0.1 0.1 0.0 0.0
0.15 0.1 0.0 0.0
0.2 0.0 0.0 0.0
0.25 0.1 0.1 0.0
0.3 0.1 0.1 0.1
0.35 0.2 0.1 0.1
0.4 0.3 0.2 0.2
0.45 1.1 0.5 0.4
0.5 44.0 2.7 28.4
0.55 48.3 76.8 60.0
0.6 4.2 18.2 9.6
0.65 0.3 0.6 0.3
0.7 0.3 0.2 0.2
0.75 0.1 0.1 0.1
0.8 0.1 0.0 0.1
0.85 0.1 0.0 0.0
0.9 0.1 0.0 0.0
0.95 0.0 0.0 0.0
1.0 0.4 0.2 0.1
0.8 0.98 0.99
0 10 20 30 40 50 60 70

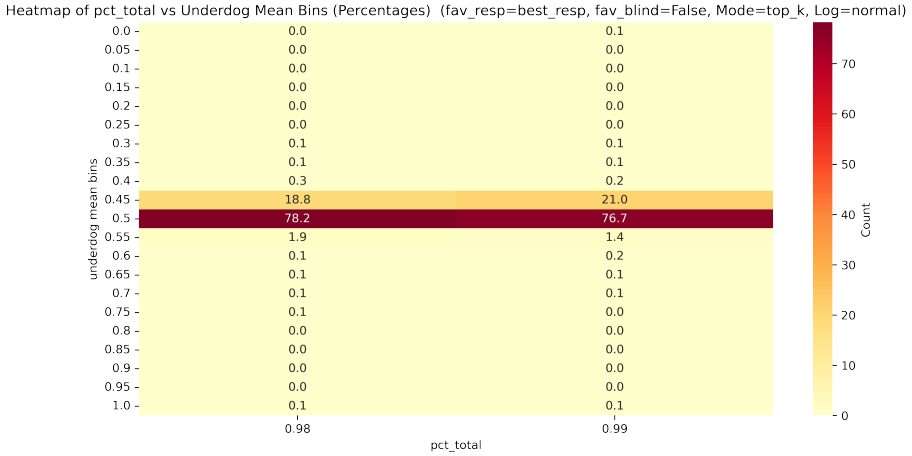

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=best_resp, fav_blind=False, Mode=top_k, Log=log)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.1 0.2 0.3 0.4 0.8 0.9 1.1 1.6 1.6 1.6
0.05 0.0 0.0 0.0 0.1 0.1 0.2 0.1 0.1 0.3 0.2
0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.1 0.4 0.2 0.3
0.15 0.0 0.0 0.1 0.1 0.2 0.2 0.3 0.1 0.2 0.3
0.2 0.0 0.0 0.1 0.3 0.2 0.4 0.4 0.6 0.3 0.3
0.25 0.0 0.1 0.2 0.4 0.4 0.3 0.7 0.5 0.4 0.7
0.3 0.1 0.0 0.4 0.3 0.5 0.7 0.9 1.0 0.9 0.9
0.35 0.1 0.1 0.8 0.8 1.6 1.2 1.7 2.0 1.6 1.3
0.4 4.3 4.5 5.0 6.1 4.7 4.6 5.3 3.9 3.2 3.2
0.45 60.2 53.9 55.4 45.1 42.4 38.6 29.4 20.4 14.9 11.2
0.5 34.6 38.8 34.9 38.0 40.7 41.4 47.1 50.7 54.5 57.2
0.55 0.1 1.8 1.8 5.8 4.7 7.0 7.4 11.2 13.8 13.5
0.6 0.1 0.2 0.3 0.8 1.1 1.4 1.8 2.9 2.8 3.6
0.65 0.0 0.1 0.2 0.4 0.6 0.8 0.8 1.4 1.6 1.4
0.7 0.0 0.1 0.1 0.2 0.4 0.3 0.5 0.7 0.8 1.0
0.75 0.0 0.1 0.1 0.1 0.2 0.4 0.6 0.4 0.6 0.5
0.8 0.0 0.0 0.1 0.1 0.1 0.3 0.2 0.3 0.3 0.3
0.85 0.0 0.0 0.0 0.1 0.1 0.2 0.1 0.2 0.4 0.4
0.9 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.2 0.1 0.3
0.95 0.1 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.1 0.4
1.0 0.1 0.1 0.1 0.5 0.9 0.8 1.3 1.4 1.6 1.4
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
60 50 40 30 20 10 0

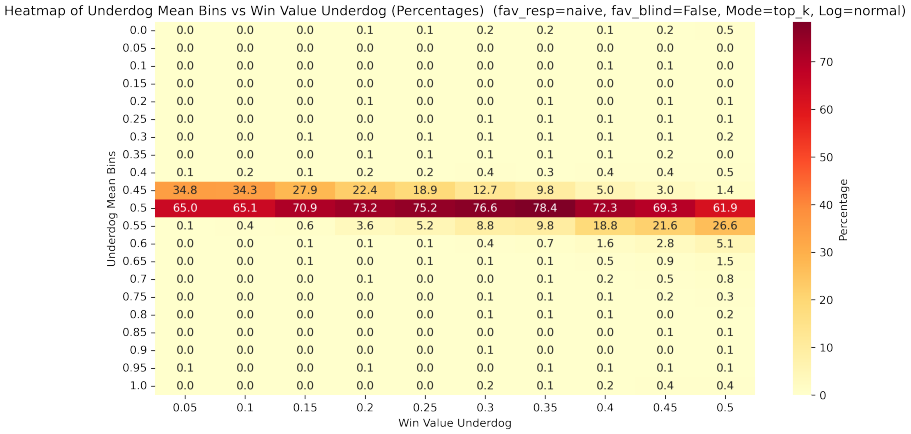

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=best_resp, fav_blind=False, Mode=top_k, Log=log)
underdog mean bins
pct_total
Count
0.0 1.0 0.9
0.05 0.1 0.1
0.1 0.1 0.2
0.15 0.2 0.2
0.2 0.3 0.2
0.25 0.5 0.3
0.3 0.5 0.7
0.35 1.2 1.1
0.4 4.4 4.6
0.45 34.7 37.1
0.5 45.0 43.6
0.55 7.5 6.7
0.6 1.6 1.6
0.65 0.8 0.7
0.7 0.4 0.4
0.75 0.4 0.3
0.8 0.2 0.2
0.85 0.2 0.1
0.9 0.1 0.1
0.95 0.1 0.1
1.0 0.9 0.8
0.98 0.99
40 35 30 25 20 15 10 5

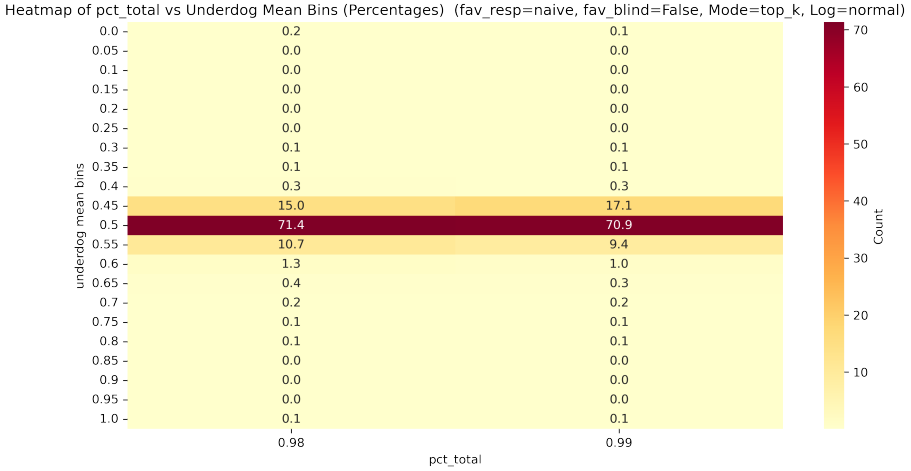

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=naive, fav_blind=False, Mode=top_k, Log=log)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.1 0.1 0.5 0.7 0.9 1.0 1.6 2.2 2.1 2.6
0.05 0.0 0.0 0.1 0.1 0.1 0.2 0.2 0.2 0.4 0.2
0.1 0.0 0.0 0.1 0.2 0.1 0.1 0.3 0.2 0.4 0.3
0.15 0.0 0.0 0.1 0.1 0.1 0.2 0.2 0.2 0.3 0.4
0.2 0.0 0.0 0.1 0.2 0.2 0.2 0.4 0.4 0.6 0.5
0.25 0.1 0.0 0.1 0.3 0.5 0.4 0.5 0.6 0.9 0.8
0.3 0.1 0.0 0.2 0.3 0.7 0.7 0.9 1.2 1.1 0.9
0.35 0.2 0.1 0.9 0.9 1.4 1.4 1.8 1.5 1.4 1.8
0.4 4.4 4.7 4.9 5.1 4.2 3.9 4.9 3.2 3.0 2.5
0.45 60.0 55.0 46.9 36.9 30.5 29.4 20.5 13.9 11.2 8.2
0.5 34.3 37.2 41.2 42.4 44.6 43.4 43.0 40.2 34.6 34.2
0.55 0.5 2.1 3.3 9.1 10.8 11.7 15.4 19.5 22.8 23.5
0.6 0.2 0.2 0.6 1.3 2.3 2.9 3.9 6.0 8.2 8.6
0.65 0.1 0.2 0.3 0.6 0.9 1.5 1.6 3.2 3.9 4.4
0.7 0.0 0.1 0.1 0.3 0.7 0.8 1.0 1.5 2.0 2.8
0.75 0.0 0.0 0.0 0.2 0.2 0.4 0.6 1.2 1.4 1.5
0.8 0.0 0.0 0.1 0.1 0.2 0.4 0.5 0.7 1.0 1.4
0.85 0.0 0.0 0.1 0.2 0.1 0.2 0.4 0.5 0.7 0.6
0.9 0.0 0.0 0.1 0.0 0.2 0.1 0.2 0.4 0.5 0.7
0.95 0.1 0.0 0.0 0.1 0.1 0.1 0.3 0.5 0.5 0.2
1.0 0.0 0.0 0.3 0.9 1.0 1.1 1.7 2.6 3.1 3.8
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
60 50 40 30 20 10 0

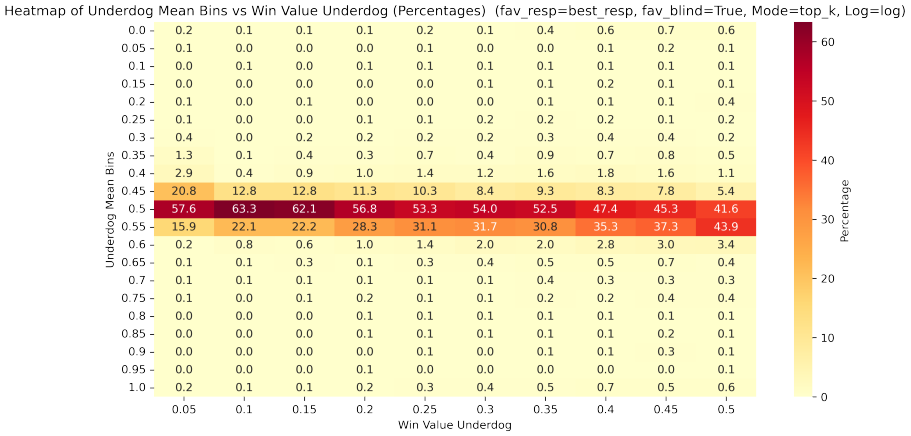

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=naive, fav_blind=False, Mode=top_k, Log=log)
underdog mean bins
pct_total
Count
0.0 1.5 1.2
0.05 0.2 0.2
0.1 0.2 0.1
0.15 0.2 0.1
0.2 0.3 0.2
0.25 0.5 0.4
0.3 0.7 0.6
0.35 1.4 1.1
0.4 4.1 4.0
0.45 25.2 31.2
0.5 40.2 39.7
0.55 14.3 11.9
0.6 4.0 3.6
0.65 2.1 1.6
0.7 1.2 0.8
0.75 0.6 0.6
0.8 0.5 0.4
0.85 0.4 0.2
0.9 0.3 0.2
0.95 0.2 0.2
1.0 1.7 1.5
0.98 0.99
40 35 30 25 20 15 10 5

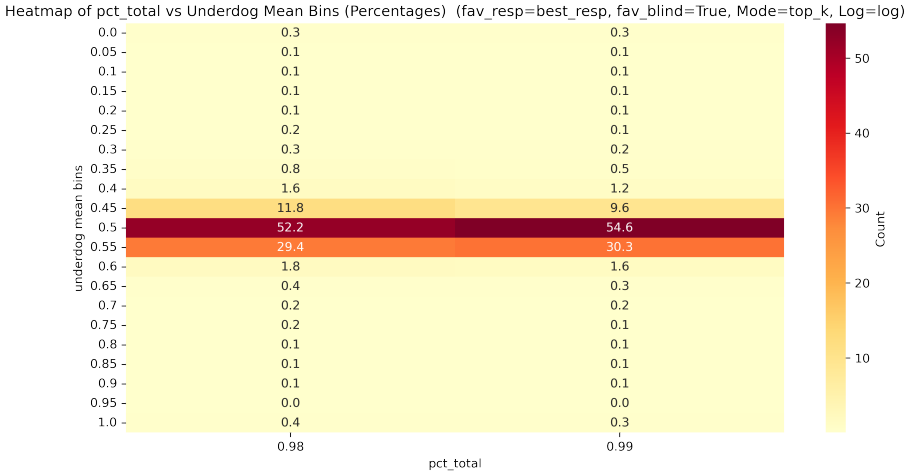

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=best_resp, fav_blind=False, Mode=top_k, Log=normal)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.1 0.1 0.2 0.1
0.05 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0
0.15 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0
0.2 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1
0.25 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.0 0.1
0.3 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.1 0.1
0.35 0.0 0.0 0.0 0.0 0.0 0.2 0.2 0.2 0.1 0.1
0.4 0.1 0.0 0.1 0.1 0.2 0.3 0.6 0.4 0.4 0.3
0.45 37.5 35.2 34.5 27.2 24.7 21.0 12.9 8.5 4.8 2.1
0.5 62.4 64.4 64.7 70.6 73.6 75.4 83.7 87.5 91.0 93.5
0.55 0.0 0.2 0.5 1.8 1.1 2.4 1.8 2.4 2.4 2.6
0.6 0.0 0.0 0.0 0.1 0.1 0.2 0.2 0.2 0.3 0.4
0.65 0.0 0.0 0.0 0.1 0.1 0.1 0.0 0.2 0.1 0.1
0.7 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1
0.75 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.1
0.8 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1
0.85 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.0
0.9 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.95 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0
1.0 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.1 0.1 0.2
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
80
60
40
20
0

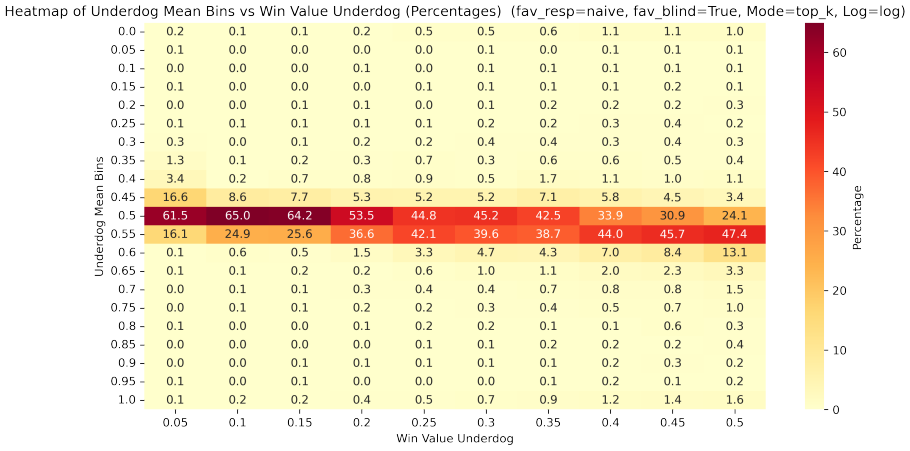

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=best_resp, fav_blind=False, Mode=top_k, Log=normal)
underdog mean bins
pct_total
Count
0.0 0.0 0.1
0.05 0.0 0.0
0.1 0.0 0.0
0.15 0.0 0.0
0.2 0.0 0.0
0.25 0.0 0.0
0.3 0.1 0.1
0.35 0.1 0.1
0.4 0.3 0.2
0.45 18.8 21.0
0.5 78.2 76.7
0.55 1.9 1.4
0.6 0.1 0.2
0.65 0.1 0.1
0.7 0.1 0.1
0.75 0.1 0.0
0.8 0.0 0.0
0.85 0.0 0.0
0.9 0.0 0.0
0.95 0.0 0.0
1.0 0.1 0.1
0.98 0.99
70
60
50
40
30
20
10
0

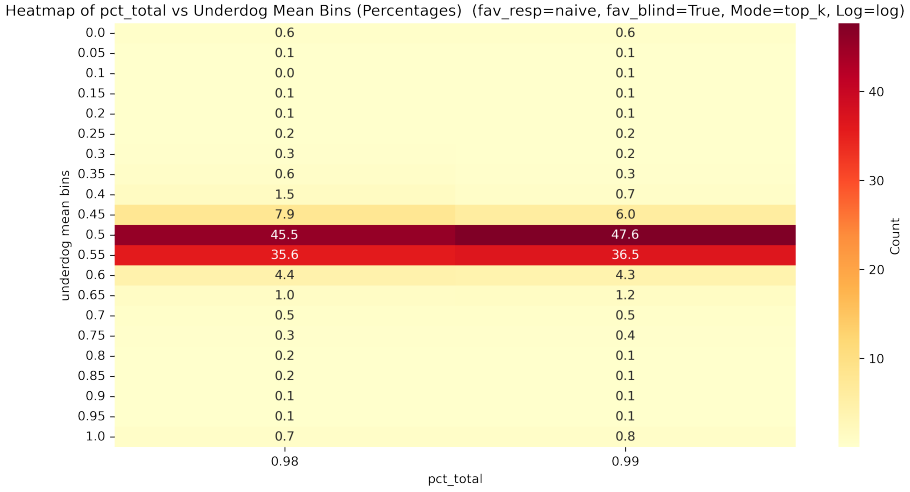

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=naive, fav_blind=False, Mode=top_k, Log=normal)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.0 0.0 0.0 0.1 0.1 0.2 0.2 0.1 0.2 0.5
0.05 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.0
0.15 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.2 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.0 0.1 0.1
0.25 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1
0.3 0.0 0.0 0.1 0.0 0.1 0.1 0.1 0.1 0.1 0.2
0.35 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.2 0.0
0.4 0.1 0.2 0.1 0.2 0.2 0.4 0.3 0.4 0.4 0.5
0.45 34.8 34.3 27.9 22.4 18.9 12.7 9.8 5.0 3.0 1.4
0.5 65.0 65.1 70.9 73.2 75.2 76.6 78.4 72.3 69.3 61.9
0.55 0.1 0.4 0.6 3.6 5.2 8.8 9.8 18.8 21.6 26.6
0.6 0.0 0.0 0.1 0.1 0.1 0.4 0.7 1.6 2.8 5.1
0.65 0.0 0.0 0.1 0.0 0.1 0.1 0.1 0.5 0.9 1.5
0.7 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.2 0.5 0.8
0.75 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.2 0.3
0.8 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.0 0.2
0.85 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1
0.9 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.1
0.95 0.1 0.0 0.0 0.1 0.0 0.0 0.1 0.1 0.1 0.1
1.0 0.0 0.0 0.0 0.0 0.0 0.2 0.1 0.2 0.4 0.4
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 10 20 30 40 50 60 70

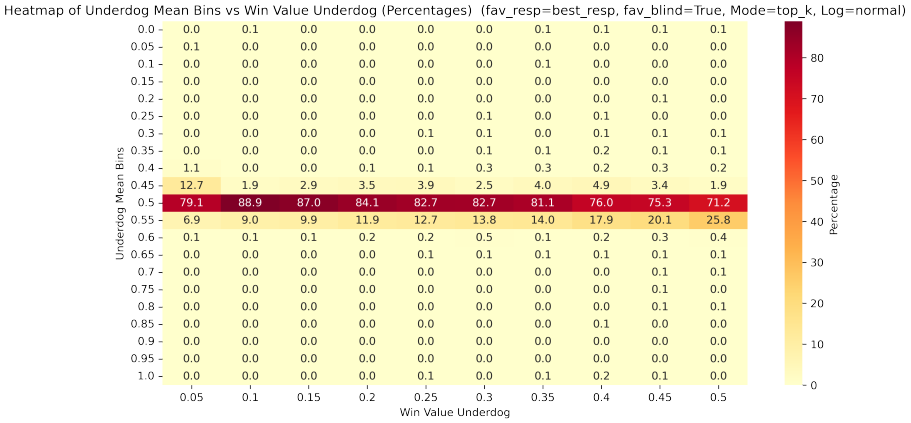

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=naive, fav_blind=False, Mode=top_k, Log=normal)
underdog mean bins
pct_total
Count
0.0 0.2 0.1
0.05 0.0 0.0
0.1 0.0 0.0
0.15 0.0 0.0
0.2 0.0 0.0
0.25 0.0 0.0
0.3 0.1 0.1
0.35 0.1 0.1
0.4 0.3 0.3
0.45 15.0 17.1
0.5 71.4 70.9
0.55 10.7 9.4
0.6 1.3 1.0
0.65 0.4 0.3
0.7 0.2 0.2
0.75 0.1 0.1
0.8 0.1 0.1
0.85 0.0 0.0
0.9 0.0 0.0
0.95 0.0 0.0
1.0 0.1 0.1
0.98 0.99
10 20 30 40 50 60 70

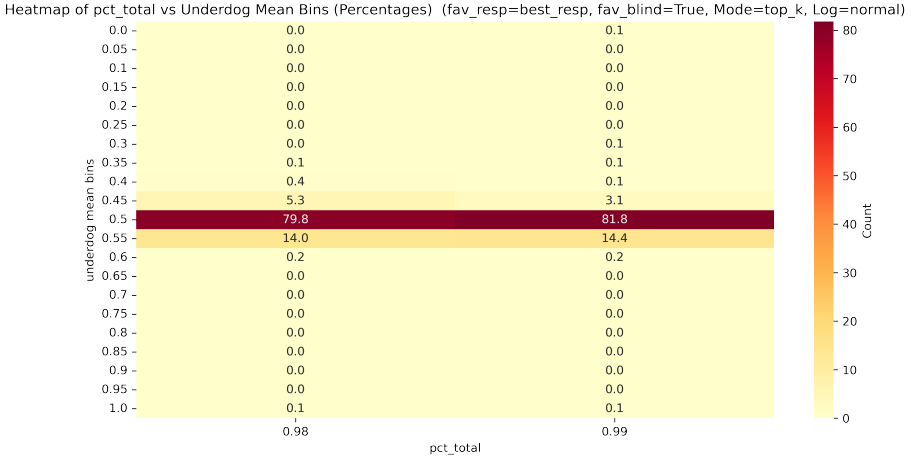

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=best_resp, fav_blind=True, Mode=top_k, Log=log)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.2 0.1 0.1 0.1 0.2 0.1 0.4 0.6 0.7 0.6
0.05 0.1 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.2 0.1
0.1 0.0 0.1 0.0 0.1 0.1 0.1 0.1 0.1 0.1 0.1
0.15 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.2 0.1 0.1
0.2 0.1 0.0 0.1 0.0 0.0 0.0 0.1 0.1 0.1 0.4
0.25 0.1 0.0 0.0 0.1 0.1 0.2 0.2 0.2 0.1 0.2
0.3 0.4 0.0 0.2 0.2 0.2 0.2 0.3 0.4 0.4 0.2
0.35 1.3 0.1 0.4 0.3 0.7 0.4 0.9 0.7 0.8 0.5
0.4 2.9 0.4 0.9 1.0 1.4 1.2 1.6 1.8 1.6 1.1
0.45 20.8 12.8 12.8 11.3 10.3 8.4 9.3 8.3 7.8 5.4
0.5 57.6 63.3 62.1 56.8 53.3 54.0 52.5 47.4 45.3 41.6
0.55 15.9 22.1 22.2 28.3 31.1 31.7 30.8 35.3 37.3 43.9
0.6 0.2 0.8 0.6 1.0 1.4 2.0 2.0 2.8 3.0 3.4
0.65 0.1 0.1 0.3 0.1 0.3 0.4 0.5 0.5 0.7 0.4
0.7 0.1 0.1 0.1 0.1 0.1 0.1 0.4 0.3 0.3 0.3
0.75 0.1 0.0 0.1 0.2 0.1 0.1 0.2 0.2 0.4 0.4
0.8 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.1 0.1
0.85 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.2 0.1
0.9 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.3 0.1
0.95 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0 0.0 0.1
1.0 0.2 0.1 0.1 0.2 0.3 0.4 0.5 0.7 0.5 0.6
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 10 20 30 40 50 60

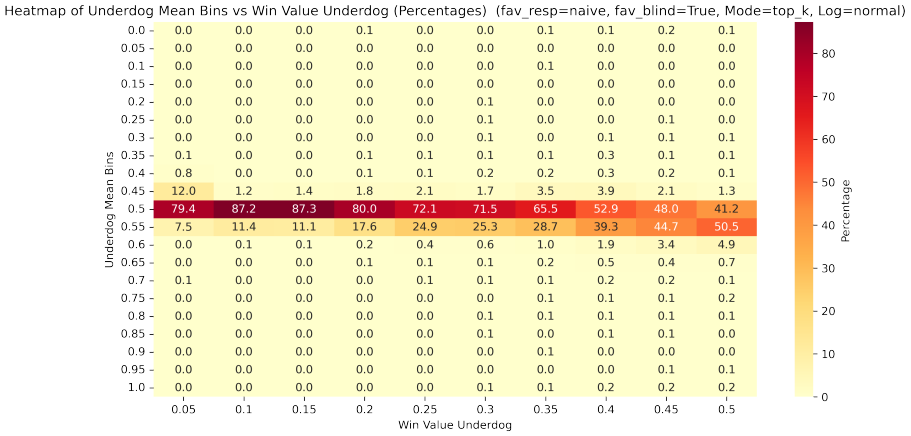

Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=best_resp, fav_blind=True, Mode=top_k, Log=log)
underdog mean bins
pct_total
Count
0.0 0.3 0.3
0.05 0.1 0.1
0.1 0.1 0.1
0.15 0.1 0.1
0.2 0.1 0.1
0.25 0.2 0.1
0.3 0.3 0.2
0.35 0.8 0.5
0.4 1.6 1.2
0.45 11.8 9.6
0.5 52.2 54.6
0.55 29.4 30.3
0.6 1.8 1.6
0.65 0.4 0.3
0.7 0.2 0.2
0.75 0.2 0.1
0.8 0.1 0.1
0.85 0.1 0.1
0.9 0.1 0.1
0.95 0.0 0.0
1.0 0.4 0.3
0.98 0.99
10 20 30 40 50

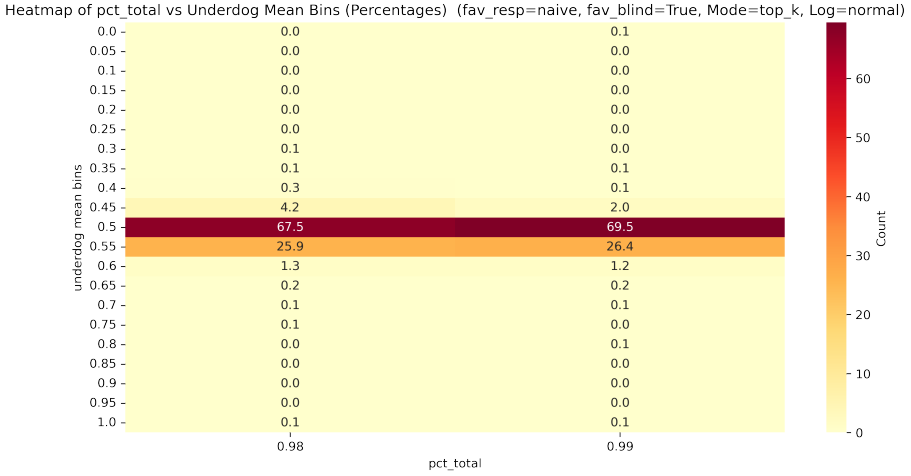

Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=naive, fav_blind=True, Mode=top_k, Log=log)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.2 0.1 0.1 0.2 0.5 0.5 0.6 1.1 1.1 1.0
0.05 0.1 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.1
0.1 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.1 0.1 0.1
0.15 0.1 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.2 0.1
0.2 0.0 0.0 0.1 0.1 0.0 0.1 0.2 0.2 0.2 0.3
0.25 0.1 0.1 0.1 0.1 0.1 0.2 0.2 0.3 0.4 0.2
0.3 0.3 0.0 0.1 0.2 0.2 0.4 0.4 0.3 0.4 0.3
0.35 1.3 0.1 0.2 0.3 0.7 0.3 0.6 0.6 0.5 0.4
0.4 3.4 0.2 0.7 0.8 0.9 0.5 1.7 1.1 1.0 1.1
0.45 16.6 8.6 7.7 5.3 5.2 5.2 7.1 5.8 4.5 3.4
0.5 61.5 65.0 64.2 53.5 44.8 45.2 42.5 33.9 30.9 24.1
0.55 16.1 24.9 25.6 36.6 42.1 39.6 38.7 44.0 45.7 47.4
0.6 0.1 0.6 0.5 1.5 3.3 4.7 4.3 7.0 8.4 13.1
0.65 0.1 0.1 0.2 0.2 0.6 1.0 1.1 2.0 2.3 3.3
0.7 0.0 0.1 0.1 0.3 0.4 0.4 0.7 0.8 0.8 1.5
0.75 0.0 0.1 0.1 0.2 0.2 0.3 0.4 0.5 0.7 1.0
0.8 0.1 0.0 0.0 0.1 0.2 0.2 0.1 0.1 0.6 0.3
0.85 0.0 0.0 0.0 0.0 0.1 0.1 0.2 0.2 0.2 0.4
0.9 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.2 0.3 0.2
0.95 0.1 0.0 0.1 0.0 0.0 0.0 0.1 0.2 0.1 0.2
1.0 0.1 0.2 0.2 0.4 0.5 0.7 0.9 1.2 1.4 1.6
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 10 20 30 40 50 60


Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=naive, fav_blind=True, Mode=top_k, Log=log)
underdog mean bins
pct_total
Count
0.0 0.6 0.6
0.05 0.1 0.1
0.1 0.0 0.1
0.15 0.1 0.1
0.2 0.1 0.1
0.25 0.2 0.2
0.3 0.3 0.2
0.35 0.6 0.3
0.4 1.5 0.7
0.45 7.9 6.0
0.5 45.5 47.6
0.55 35.6 36.5
0.6 4.4 4.3
0.65 1.0 1.2
0.7 0.5 0.5
0.75 0.3 0.4
0.8 0.2 0.1
0.85 0.2 0.1
0.9 0.1 0.1
0.95 0.1 0.1
1.0 0.7 0.8
0.98 0.99
10 20 30 40


Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=best_resp, fav_blind=True, Mode=top_k, Log=normal)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.0 0.1 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1
0.05 0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0
0.15 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.2 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0
0.25 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.0 0.0
0.3 0.0 0.0 0.0 0.0 0.1 0.1 0.0 0.1 0.1 0.1
0.35 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.2 0.1 0.1
0.4 1.1 0.0 0.0 0.1 0.1 0.3 0.3 0.2 0.3 0.2
0.45 12.7 1.9 2.9 3.5 3.9 2.5 4.0 4.9 3.4 1.9
0.5 79.1 88.9 87.0 84.1 82.7 82.7 81.1 76.0 75.3 71.2
0.55 6.9 9.0 9.9 11.9 12.7 13.8 14.0 17.9 20.1 25.8
0.6 0.1 0.1 0.1 0.2 0.2 0.5 0.1 0.2 0.3 0.4
0.65 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1 0.1
0.7 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1
0.75 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0
0.8 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1
0.85 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0
0.9 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.95 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
1.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.2 0.1 0.0
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 10 20 30 40 50 60 70 80


Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=best_resp, fav_blind=True, Mode=top_k, Log=normal)
underdog mean bins
pct_total
Count
0.0 0.0 0.1
0.05 0.0 0.0
0.1 0.0 0.0
0.15 0.0 0.0
0.2 0.0 0.0
0.25 0.0 0.0
0.3 0.0 0.1
0.35 0.1 0.1
0.4 0.4 0.1
0.45 5.3 3.1
0.5 79.8 81.8
0.55 14.0 14.4
0.6 0.2 0.2
0.65 0.0 0.0
0.7 0.0 0.0
0.75 0.0 0.0
0.8 0.0 0.0
0.85 0.0 0.0
0.9 0.0 0.0
0.95 0.0 0.0
1.0 0.1 0.1
0.98 0.99
0 10 20 30 40 50 60 70 80


Heatmap of Underdog Mean Bins vs Win Value Underdog (Percentages) (fav_resp=naive, fav_blind=True, Mode=top_k, Log=normal)
Underdog Mean Bins
Win Value Underdog
Percentage
0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.1 0.2 0.1
0.05 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.1 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0
0.15 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.2 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0 0.0
0.25 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.1 0.0
0.3 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.1
0.35 0.1 0.0 0.0 0.1 0.1 0.1 0.1 0.3 0.1 0.1
0.4 0.8 0.0 0.0 0.1 0.1 0.2 0.2 0.3 0.2 0.1
0.45 12.0 1.2 1.4 1.8 2.1 1.7 3.5 3.9 2.1 1.3
0.5 79.4 87.2 87.3 80.0 72.1 71.5 65.5 52.9 48.0 41.2
0.55 7.5 11.4 11.1 17.6 24.9 25.3 28.7 39.3 44.7 50.5
0.6 0.0 0.1 0.1 0.2 0.4 0.6 1.0 1.9 3.4 4.9
0.65 0.0 0.0 0.0 0.1 0.1 0.1 0.2 0.5 0.4 0.7
0.7 0.1 0.0 0.0 0.0 0.1 0.1 0.1 0.2 0.2 0.1
0.75 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.2
0.8 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.1 0.1 0.1
0.85 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.1 0.1 0.0
0.9 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.0 0.0 0.0
0.95 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1
1.0 0.0 0.0 0.0 0.0 0.0 0.1 0.1 0.2 0.2 0.2
0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5
0 10 20 30 40 50 60 70 80


Heatmap of pct_total vs Underdog Mean Bins (Percentages) (fav_resp=naive, fav_blind=True, Mode=top_k, Log=normal)
underdog mean bins
pct_total
Count
0.0 0.0 0.1
0.05 0.0 0.0
0.1 0.0 0.0
0.15 0.0 0.0
0.2 0.0 0.0
0.25 0.0 0.0
0.3 0.1 0.0
0.35 0.1 0.1
0.4 0.3 0.1
0.45 4.2 2.0
0.5 67.5 69.5
0.55 25.9 26.4
0.6 1.3 1.2
0.65 0.2 0.2
0.7 0.1 0.1
0.75 0.1 0.0
0.8 0.0 0.1
0.85 0.0 0.0
0.9 0.0 0.0
0.95 0.0 0.0
1.0 0.1 0.1
0.98 0.99
0 10 20 30 40 50 60