Estimating Cheating Rates in Titled Tuesday

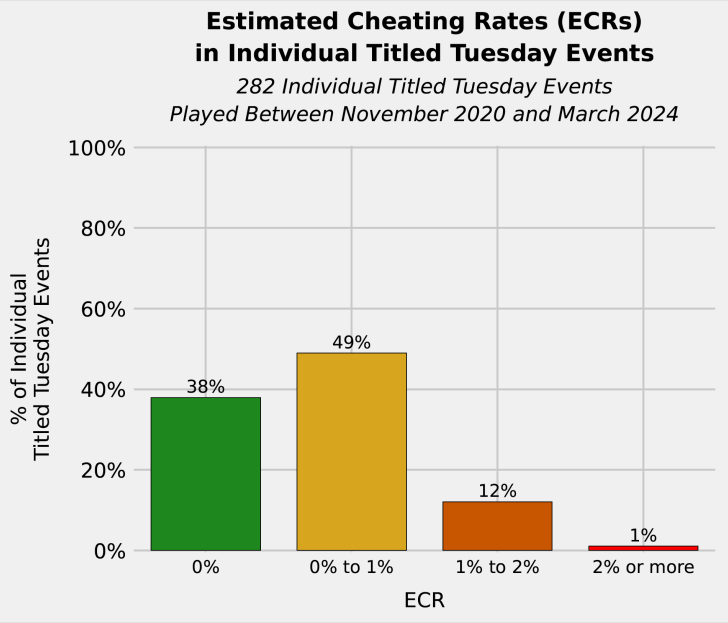
Published: July 2024

This report was researched and written by Chess.com’s Fair Play Team, a group of over 30 dedicated professionals comprising researchers, statisticians, data scientists, analysts, event proctors, and engineers. The team, which includes three grandmasters and ten titled players, works diligently to protect the integrity of the game. The report was overseen by Chess.com’s Director of Fair Play, FM Dan Rozovsky.

Introduction

In our first report on cheating in Titled Tuesday, we compared upset rates in over the board (OTB) blitz to upset rates in Titled Tuesday and found that cheating appears to be far more limited in Titled Tuesday than many believed. We also recognized that cheating does happen.

In this second report, we examine the issue of cheating in greater detail by digging into player performances in Titled Tuesday. We approach this by developing a statistical model that predicts the outcomes of Titled Tuesday games, and using it to simulate hundreds of thousands of historical Titled Tuesday games. It is unsurprising to find that every Titled Tuesday has players who overperform relative to expectations. Most overperformances do not indicate cheating and are explained by natural variance. To evaluate the prevalence of cheating, the key question is how many overperformances are *expected* compared to how many *actually* occur. We term the difference between these two the number of surplus overperformances – and it is at the heart of this report. **Our key takeaway, visualized in the graph below, is that the estimated cheating rate (ECR) – which is the number of surplus overperformances divided by the total participants in a Titled Tuesday – is almost always below 2%, but generally greater than zero**.



In other words, we can estimate that a large majority of Titled Tuesday events have one or more players cheating in them, but these players regularly make up a small fraction of Titled Tuesday participants. Importantly, in the one-year period from March 20, 2023 to March 19, 2024, the ECRs closely align with the actions taken by our Fair Play team, who closed 1.1% of Titled Tuesday participants for cheating that happened *in Titled Tuesday* in that same one-year period.

While not every overperformance is cheating, by calculating surplus overperformance we can provide the chess community with our best estimate of the extent of cheating that is occurring in Titled Tuesday.

Methodology

To compute the surplus overperformance metric we first need to establish a range of likely outcomes for each player in each Titled Tuesday. From there we can determine: (a) how many players we ***expect*** to perform outside of their ranges due to natural variance, and (b) how many players *actually* perform outside of their range. The difference between (b) and (a), divided by the total number of Titled Tuesday participants, will serve as an estimated rate of cheating, i.e., *the estimated percentage of Titled Tuesday participants who cheat.*

Let's break down how we do this step-by-step:

1. Train a statistical model that provides us with probabilities that a player will win, draw, or lose against any opponent in Titled Tuesday.
2. Use the probabilities from the statistical model we train to simulate (replay) every Titled Tuesday game thousands of times.
3. Use the simulated game outcomes to establish 95% prediction intervals for each player in each Titled Tuesday. **For example, if a player has 95% prediction interval of 4.5 – 7.5, then that means that in approximately 95% of simulations the player ended that Titled Tuesday with a total simulated score between 4.5 and 7.5. Another way of saying this would be that our model predicts that approximately 95% of the time that player will score between 4.5 and 7.5 points <u>against the opponents they actually faced</u> in that particular Titled Tuesday.**
4. Count the number of players who *in reality* score better than their respective 95% prediction intervals. For example, if a player's prediction interval is 4.5 – 7.5 and in reality they score 8 points, then they are an overperformer.
5. Next, reuse the same simulated Titled Tuesday outcomes to determine how many overperformers we should expect to see just due to natural variance. With hundreds of participants in every Titled Tuesday, this number turns out to always be greater than 0.
6. Finally, compute the difference between the number of actual overperformers and the number of expected overperformers and call this the number of surplus overperformers. This number of surplus overperformers represents our best estimate of how many players cheat in each Titled Tuesday. **We can also express this as a percentage of all Titled Tuesday participants, which we call the estimated cheating rate (ECR). For example, if 500 players join a Titled Tuesday, 15 overperform, while we only expected 10 to overperform, then we would calculate that there are 15 - 10 = 5 surplus overperformers, giving an ECR of $\frac{5}{500}$, or 1%.**

Training a Win/Draw/Loss Model

The first step in our process is to train a statistical model that takes the relevant information about a player and their opponent as inputs and, based on what has happened historically in Titled Tuesday games, provides probabilities that the player will win, draw, or lose the game. We call this kind of model a Win/Draw/Loss (W/D/L) model.

The most important inputs into a W/D/L model are player ratings. In our earlier report we primarily relied on FIDE ratings (both blitz and classical) to compare OTB blitz and online play. In this report we're only considering performances in Titled Tuesday, so we now incorporate Chess.com blitz ratings as well, and our analysis leverages a W/D/L model – which we call our Advanced W/D/L model – that uses all three ratings simultaneously.

Note that when we repeat our analysis using simpler W/D/L models that use Chess.com blitz ratings only, it reaches similar conclusions. For more details see **W/D/L Models** in the Appendix. For readers wondering why we trained a model at all, instead of borrowing from the expected score formulas from Elo, please see **Why Not Use Elo?** in the Appendix.

Estimating Rates of Cheating

How can we use a W/D/L model to estimate the percentage of players cheating in Titled Tuesday? The short answer: by simulating the games that were played within the event thousands of times.

To understand exactly what that means, let's walk through an example step-by-step. For this example, we'll use the Advanced W/D/L model to **estimate how many players cheated in the Late Titled Tuesday that took place on July 18, 2023**. As an illustrative example, we start by focusing on a single player's performance (in this case Hikaru Nakamura, the most successful player in Titled Tuesday history), and then we do the same for all participants in the event.

Simulating Games for an Individual Player

The Advanced W/D/L model we trained provides the following per-game chances Hikaru would win, draw, or lose:

Hikaru Nakamura's W/D/L Probabilities

Late Titled Tuesday on July 18, 2023

Round	Opponent	Win Probability	Draw Probability	Loss Probability	Actual Result
1	IM Irishchessman16 (2507)	89%	6%	5%	Win
2	IM wannabe2700 (2603)	92%	4%	4%	Win
3	FM Fantastic_Power (2736)	88%	7%	5%	Win
4	GM penguingm1 (2884)	86%	8%	6%	Win
5	GM vi_pranav (2978)	55%	26%	19%	Win
6	IM 0gZPanda (2983)	80%	10%	10%	Win
7	GM Jospem (3066)	64%	23%	13%	Win
8	GM Oleksandr_Bortnyk (3036)	46%	34%	20%	Draw
9	IM MITerryble (2978)	73%	16%	11%	Win
10	GM jefferyx (2991)	48%	38%	14%	Draw
11	GM VladimirKramnik (2990)	65%	23%	12%	Win

Notes:

1. Opponent ratings reflect **pre**game values.
2. W/D/L probabilities generated by Chess.com's **Advanced** W/D/L model.

We leverage these probabilities to **simulate** each of Hikaru's games. In our case, "simulating" involves writing a program to effectively roll a 3-sided die thousands of times – weighted according to the probabilities that Hikaru would win, draw, or lose. For example, in Hikaru's round 1 game, the program will randomly select one of the following:

- 1 – which is the number of points Hikaru earns for a win – with a 89% probability;
- 0.5 – which is the number of points Hikaru earns for a draw – with a 6% probability; or
- 0 – which is the number of points Hikaru earns for a loss – with a 5% probability.

Let's say the program selects a 1, i.e., Hikaru wins. We'll keep track of the result and say that Hikaru is scoring 1 out of 1 in this simulated Titled Tuesday.

We move on to Hikaru's round 2 game and this time we ask the program to randomly select one of the following:

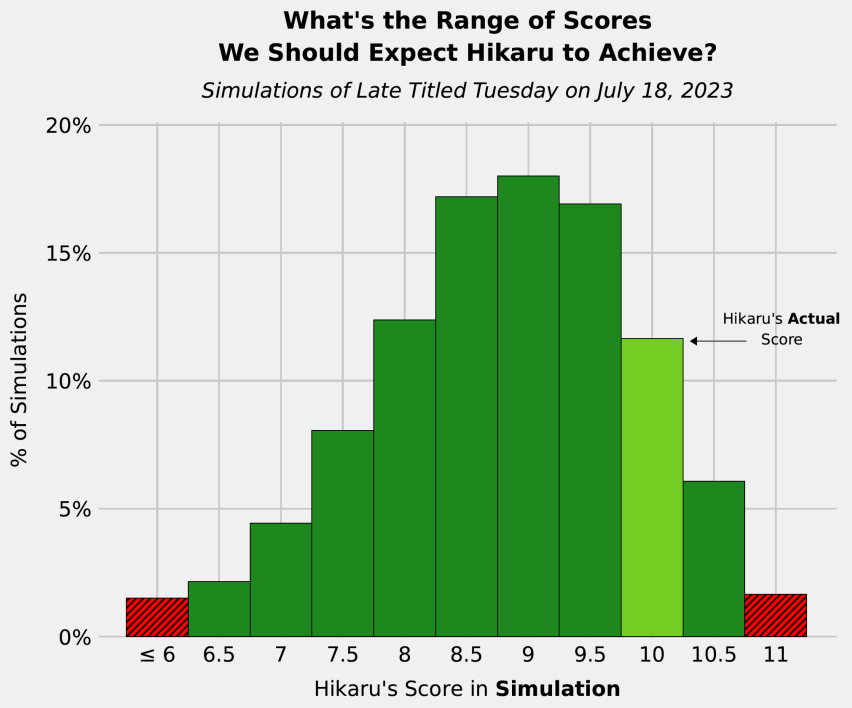
- 1 with a 92% probability;
- 0.5 with a 4% probability; or
- 0 with a 4% probability.

This time, the program selects a 0, i.e., Hikaru loses. This is an unlikely outcome (4%), but when we run a simulation the results do not have to match what actually happened in the event. So now Hikaru is scoring 1 out of 2 in this simulated Titled Tuesday.

We continue this process for Hikaru's 9 remaining games, and the program arrives at a total score for the event (e.g. 8.5/11). **That single simulated score isn't useful on its own, but by repeating the same simulation process thousands of times, we can make powerful observations by looking at the *distribution* of Hikaru's simulated scores.**

Indeed, the graph below shows that when we repeat the simulation process thousands of times we see that:

(a) Hikaru most commonly scores 9 points out of 11;
(b) Hikaru scores 10 points – his real-life score – or better 19% of the time; and
(c) ***Approximately 95% of the time, Hikaru scores between 6.5/11 and 10.5/11***. Please note:
  (i) We explain in the Appendix **What a "95% Prediction Interval" Really Means** why we use a range that covers *approximately* 95% of Hikaru's scores instead of *exactly* 95%.
  (ii) This range reflects the particularly high level of opposition Hikaru faced in this specific Titled Tuesday. In practice, Hikaru is very unlikely to score 6.5 points in a Titled Tuesday because if he performed poorly, he would face weaker opposition. We elaborate on this point further in Appendix **Why Hikaru's Prediction Interval Runs Low**.



This is powerful information. According to our Advanced W/D/L model, Hikaru's actual score of 10/11 was better than his expected result against his real life lineup of opponents, but was still *within the range* of scores that fall within his 95% prediction interval. If Hikaru scored outside of his 95% prediction interval in real life – below 6.5 or above 10.5 in this case – we would call it an **underperformance or an overperformance**.

Simulating Games for all Titled Tuesday Participants

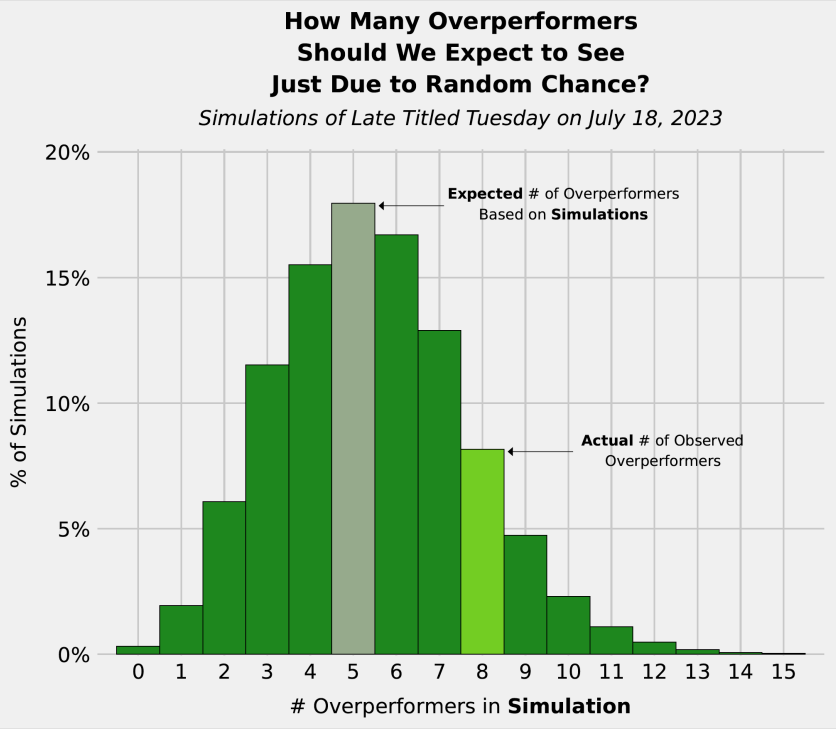
Next let's broaden our focus beyond just one player, and run the exact same simulation for *all* 438 participants who played in the Titled Tuesday (see **Inclusion Criterion** in the Appendix for how we arrived at 438). That means every individual who played in the Late Titled Tuesday on July 18, 2023 will have their own 95% prediction interval based on thousands of simulations just like the one we conducted for Hikaru.

When we do this for the Late Titled Tuesday on July 18, 2023, we see that 8 players outperform their respective 95% prediction intervals.

Expected and Surplus Overperformers

Does this mean that these 8 overperformers cheated? **No. When you have 438 participants, as in this particular Titled Tuesday, what would be statistically improbable was if not a single player overperformed.** On their own, each of these 8 overperformances is impressive and unexpected. What we need to do is zoom out, look at the event as a whole, and ask: how many overperformers is cause for concern?

To answer this question, we need to know how many overperformers to *expect due to natural variance.* We can figure this out by leveraging the same simulations we ran to compute each individual player's 95% prediction interval. The mechanics of this process are more complex (see Appendix **Expected Overperformers Methodology** for a detailed explanation) but it leaves us with an important new distribution: **the distribution of overperformers we may observe due to natural variance,** shown in the graph below, for the Late Titled Tuesday on July 18, 2023.

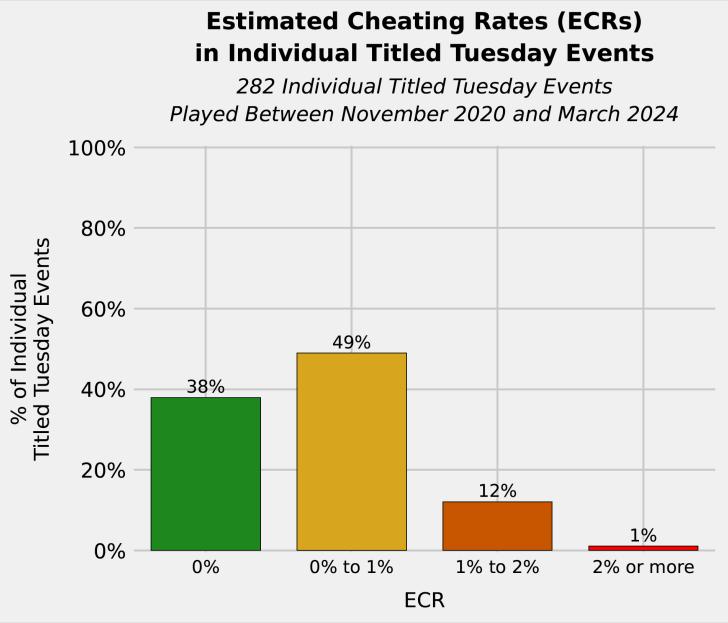


Now we can put the 8 overperformers from the Late Titled Tuesday on July 18, 2023 into context. Based on the simulation we expected 5 overperformers for that event, but the actual event had 8. **That means there were 3 surplus overperformers – or, 3 more overperformers than statistically expected. Dividing that number of surplus overperformers by the number of participants, 438, gives us an estimated cheating rate (ECR) of 0.68%.**

Note, this does not mean that (a) 3 of the 8 overperformers *definitely* cheated, or (b) 3 out of the larger group of 438 participants *definitely* cheated. The number of expected overperformances is not fixed, but varies stochastically. It's statistically possible that fewer than 3 players cheated – in fact, the distribution above shows that 17% of simulations included 8 or more overperformers just due to natural variance – and it's also statistically possible that *more* than 3 players cheated. When it comes to any individual player's result, further investigation is required into strength of play, strength of *opponent* play, and other factors to conclusively determine whether cheated occurred. What this analysis enables us to do is *estimate* an *overall* cheating rate of 0.68% in the July 18, 2023 Late Titled Tuesday.

Findings

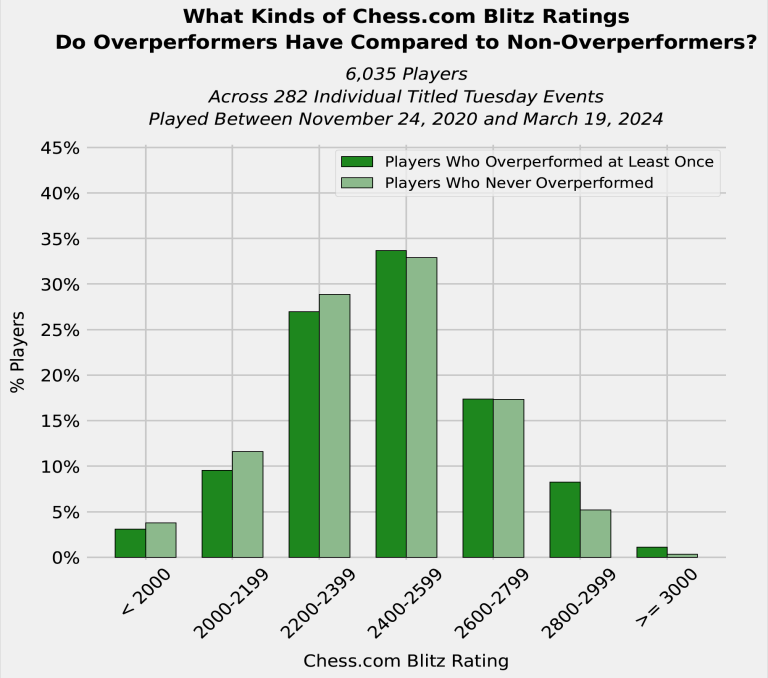
In the above example, we concluded that it's reasonable to estimate a cheating rate of 0.68% in the Late Titled Tuesday on July 18, 2023. That's only one example, though. What we really want to know is the ECR across *many* Titled Tuesdays. The graph below, also displayed in the introduction, shows how ECR varies across all individual Titled Tuesdays played between November 24, 2020 and March 19, 2024.



As noted in the introduction, **slightly less than two-thirds of Titled Tuesday events have ECRs above 0%, and 99% of Titled Tuesday events have ECRs below 2%**. See Appendix **ECR by Year** and **ECR by W/D/L Model** for additional views of ECR.

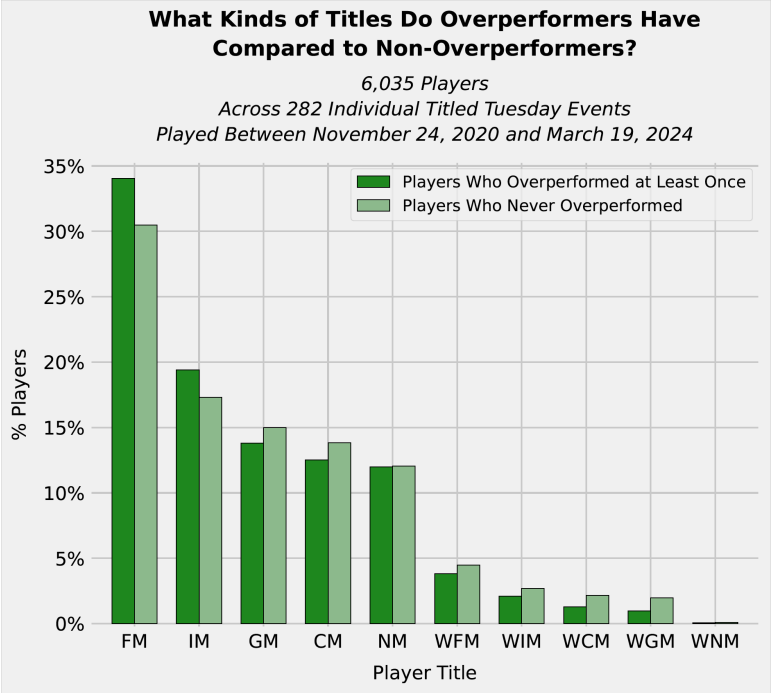
It's important to remember that the methodology used to calculate ECR does not indicate exactly which players may have cheated. Instead, it broadly identifies overperformers, and then uses the number of *surplus* overperformers to arrive at an ECR. But we can still ask: how does the cohort of overperformers compare to the rest of the players in Titled Tuesday?

The graph below shows the rating distribution of the two groups.

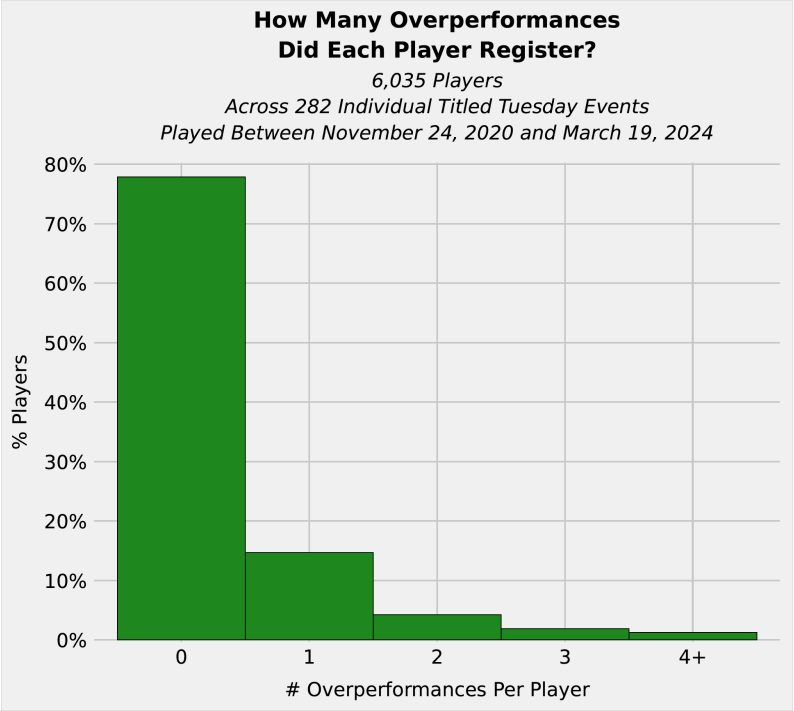


We see that overperformers most typically have a Chess.com blitz rating between 2400 and 2599. This is fairly unremarkable, as the median Chess.com blitz rating among non-overperformers in Titled Tuesday also falls within this range.

If we look at the titles of overperformers in the graph below, we find that overperformers are most commonly FMs. This is also unsurprising, as it's also the most common title among non-overperformers. However, players of all strengths are found to occasionally overperform, including GMs.



Next, let's look at how many times each player turns up as an overperformer. How many players *ever* qualify as an overperformer? Are any players overperforming multiple times?



We can see that 78% – or roughly three-quarters – of the 6,035 players we analyzed, never overperformed at all. Of the remaining 1,334 players who overperformed, 66% did so only one time.

Readers may also wonder what connection exists, if any, between the **overperformance** methodology used in this report and **actual closures** the Chess.com Fair Play team has made. Of the 4,340 players who participated in Titled Tuesday during the one-year period from March 20, 2023 to March 19, 2024, the Chess.com Fair Play team **closed 1.1% of them** for cheating that happened in Titled Tuesday. Breaking that down further:

- Of the 674 players *who overperformed at least once* during that period, **3.6% of them were eventually closed** for cheating that happened in Titled Tuesday
- Among the remaining 3,666 players who *never overperformed* in Titled Tuesday during the one-year period, the Chess.com Fair Play team **closed only 0.7% of them** for cheating that happened in Titled Tuesday.

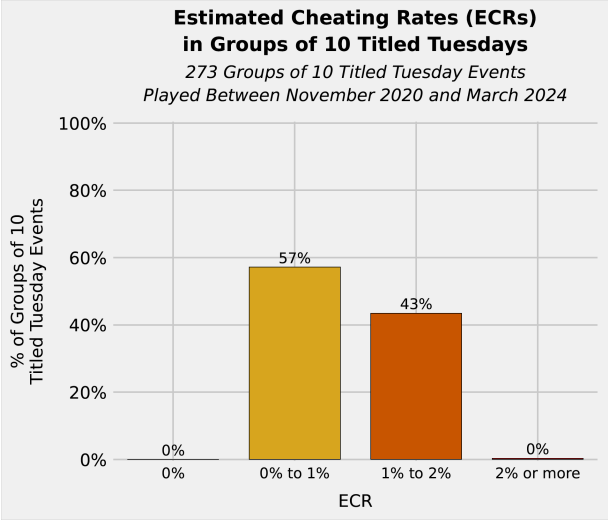
We can glean a couple of insights from this comparison:

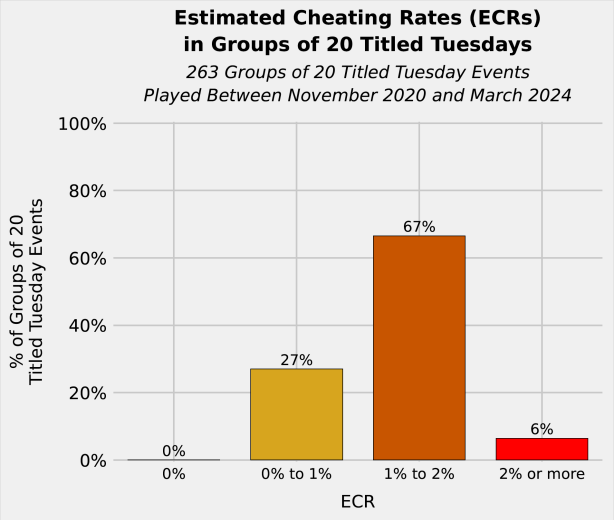
- **The Chess.com Fair Play team is 5x more likely to close an overperforming account for cheating that happened in Titled Tuesday than it is to close a non-overperforming account**
- **The large majority of overperformances are more likely the result of natural variance than cheating**.

Cheating Rates Over Longer Time Horizons

There is a potential shortcoming of the analysis thus far: players may cheat in only a few games in certain Titled Tuesdays and play cleanly in the remaining set. These players may rarely, if ever, perform to such a degree that they would outperform their respective 95% prediction interval in a single Titled Tuesday, but they may, *over the course of multiple events*, perform in the aggregate better than what our model would expect.

We can correct for this shortcoming by looking at Titled Tuesdays together in groups of 10 or 20 consecutive events (note Appendix **Simulation Adjustment**). The way we do this is by numbering Titled Tuesdays chronologically from 1 to 282. The first group consists of Titled Tuesdays numbered 1-10; the second group consists of Titled Tuesdays numbered 2-11; and so on. We do the same for groups of 20. The distributions of ECRs for these group sizes are shown in the graphs below.





When we consider multiple Titled Tuesdays via this grouping method, ECR does tend to increase, which may suggest that some players who cheat in Titled Tuesday only do so sparingly. However, ECR remains below 2% for the vast majority of Titled Tuesday groupings.

Conclusion

Our goal in this report is to provide the community with our best estimate of the rate of cheating in Titled Tuesday. We accomplished this by developing a custom model and simulating hundreds of thousands of Titled Tuesday games. We also gleaned various insights from this analysis.

It is important to emphasize that while our analysis shows that the vast majority of games and players are clean, a cheating rate of even 1% is unacceptable and a serious issue in Titled Tuesday. Within the Fair Play Team, our constant focus is on preventing cheating and closing those who have cheated. Over the last few years we have closed hundreds of titled players and will keep improving our methods to deter and catch even the most occasional cheater.

Over the next few months we will be updating our fair play policies to include stronger punishments for cheating, and plan to release new software that will provide improved oversight of all players in Titled Tuesday and other prize events. Finally, to the greater than 98% of Titled Players playing clean and fair, we appreciate you and commit to ensuring that our events remain competitive and enjoyable for all.

Appendix

Why Not Use Elo?

Elo is inadequate for our purposes for a few reasons:

1. The Elo approach does not provide specific win, draw, or loss probabilities, which makes it unsuitable for simulating Titled Tuesday games in a way that mimics reality;
2. As chess statistician Jeff Sonas has noted, the Elo expected score formula performs poorly when there are large rating differences, which happens frequently in Titled Tuesday; and
3. Elo is too simplistic of a model since it only uses rating difference as an input (and it treats all equal rating differences as equally impactful).

W/D/L Model Details

Data: We first trained our models on Titled Tuesday games played between April 2020 – when the Titled Tuesday time control shifted from 3+2 to 3+1 – and November 2020. As we simulated through historical Titled Tuesday events, we retrained our models using progressively more data. Every time we trained a model, we excluded games consisting of 8 or fewer plies and games involving players banned for cheating.

Overview of W/D/L Models

Features Group	Statistical Modeling Technique	Rating Type(s) Used
Essential	*Multinomial logistic regression*	*Chess.com blitz*
Enhanced	*Multinomial logistic regression*	*Chess.com blitz*
Advanced	*CatBoost*	***Chess.com blitz*** *and* ***FIDE blitz*** *and* ***FIDE classical***

Essential Model Features:

1. Chess.com blitz rating difference
2. Chess.com blitz rating
3. Piece color

Enhanced Model Features:

1. Chess.com blitz rating difference
2. Chess.com rating
3. Piece color
4. Interaction between Chess.com blitz rating difference and Chess.com rating

Advanced Model Features:

1. Chess.com blitz rating difference
2. Chess.com blitz rating
3. FIDE blitz rating difference
4. FIDE blitz rating
5. FIDE classical rating difference
6. FIDE classical rating
7. Piece color
8. Timeout rate in 3|0 (for user and opponent)
9. Effects of Elo bias (for user and opponent)
10. Warmed up before Titled Tuesday vs. Cold start (for user and opponent)

The Advanced model does not explicitly include an interaction term between rating and rating difference since it's a CatBoost (tree-based) model.

We may share insights we found about the relationships between each of these features and W/D/L probabilities in a future report.

What a "95% Prediction Interval" Really Means

The term "95% prediction interval" is a slight misnomer. To be more precise, what we call a "95% prediction interval" covers *at least* the middle 95% of simulated scores; in practice, though, it usually covers *more* than 95% of the middle ground. To illustrate this point, consider the underlying distribution of simulated scores for Hikaru in the Late Titled Tuesday on July 18, 2023 described in the table below:

What Does a "*95%* Prediction Interval" Really Mean?

Hikaru's Simulation Results From the Late Titled Tuesday on July 18, 2023

Total Score	# of Simulations	Probability	Cumulative Probability	Part of "95%" Prediction Interval?
4.5	5	0.05%	0.05%	No
5	15	0.15%	0.20%	No
5.5	39	0.39%	0.59%	No
6	92	0.92%	1.51%	No
6.5	215	2.15%	3.66%	***Yes***
7	443	4.43%	8.09%	***Yes***
7.5	805	8.05%	16.14%	***Yes***
8	1,237	12.37%	28.51%	***Yes***
8.5	1,719	17.19%	45.70%	***Yes***
9	1,800	18.00%	63.70%	***Yes***
9.5	1,691	16.91%	80.61%	***Yes***
10	1,165	11.65%	92.26%	***Yes***
10.5	608	6.08%	98.34%	***Yes***
11	166	1.66%	100.00%	No
Total	10,000	100.0%	-	-

Note: Simulation results based on W/D/L probabilities generated by Chess.com's **Advanced** W/D/L model.

Since Titled Tuesday scores are *discrete*, in order to capture the middle 95% of simulated scores for Hikaru we actually have to capture *more* than 95% of simulated scores – 96.83% in this case. Therefore, it's wrong to assume that Hikaru, or any other player, has a 2.5% probability of performing below their 95% prediction interval and a 2.5% probability of performing above it. In this example Hikaru actually has only a 1.51% probability of performing below 6.5 and a 1.66% probability of performing above 10.5.

This becomes a critical point to factor in when we calculate how many overperformers to expect due to natural variance.

Why Hikaru's Prediction Interval Runs Low

It can be hard to fathom Hikaru scoring only 6.5/11 in Titled Tuesday, and yet our simulations tell us this would **not** be a statistical outlier in the context of the Late Titled Tuesday on July 18, 2023. Here it's important to keep in mind that the simulation procedure we employ always uses the *actual* matchups that took place in Titled Tuesday.

For Hikaru to score 10/11 as he actually did in the Late Titled Tuesday on July 18, 2023, he had to face stiff competition; his opponents were rated 2887 (Chess.com blitz) on average, and he faced multiple strong grandmasters, including: Andrew Tang, Jose Martinez, Oleksander Bortnyk, Jeffery Xiong, and Vladimir Kramnik. *Every time* we simulate Hikaru's experience in the Late Titled Tuesday on July 18, 2023, we take these matchups as given, meaning Hikaru *always* faces this same high level of opposition in each iteration. In real life, if Hikaru had lost to Jose Martinez in round 7, for example, he would not have been paired against Oleksander Bortnyk in

round 8. But our simulation methodology does not dynamically pair players based on the simulation results of each round.

To summarize, the proper interpretation of Hikaru's 95% prediction interval is not: "The W/D/L model predicts Hikaru to score between 6.5/11 and 10.5/11 approximately 95% of the time *against the overall field of participants* in this Titled Tuesday," but instead: "The W/D/L model predicts Hikaru to score between 6.5/11 and 10.5/11 approximately 95% of the time **in the 11 games that he actually played** in this Titled Tuesday."

To further illustrate this point, consider that in the Early Titled Tuesday on July 4, 2023, Hikaru's 95% prediction interval spans from 8 – 11 according to our Advanced W/D/L model. Why do simulations produce such a drastically different 95% prediction interval for Hikaru in this Titled Tuesday? The answer is that Hikaru faced much weaker opposition; his opponents were only rated 2687 (Chess.com blitz) on average.

Inclusion Criterion

We only count players who played at least 1 game consisting of at least 9 plies.

Expected Overperformers Methodology

We can conceive of our simulation methodology as the replaying of all games played in a given Titled Tuesday, repeated thousands of times. To determine how many overperformers we expect just due to natural variance, we reuse the simulation results that served to construct 95% prediction intervals for each player. Looking at only the first iteration in which we simulated all games played in Titled Tuesday, we *compare each player's simulated score to their own 95% prediction interval*. If a player's simulated score falls outside of their 95% prediction interval, then they are an over/underperformer **in that *simulated* Titled Tuesday**. Note that we are *not* comparing each player's *actual real-life* score to their 95% prediction interval; instead we are comparing each player's *simulated* score to their 95% prediction interval.

Suppose that in the first iteration of simulating an entire Titled Tuesday, 5 players have simulated scores *above* their own 95% prediction intervals. That means that in one iteration of simulating an entire Titled Tuesday, 5 players overperformed *just due to natural variance*. After we repeat this process for each of the thousands of Titled Tuesday iterations, the resulting *distribution* of overperformers ***in simulated*** Titled Tuesdays tells us how many overperformers to expect in Titled Tuesday just due to random chance. This distribution is shown in the figure **How Many Overperformers Should We Expect to See?**.

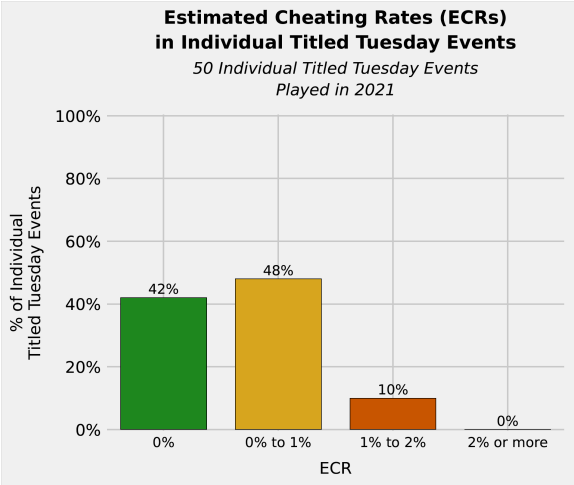
Some readers may wonder: why can't the number of expected overperformers be computed analytically based on a binomial distribution, where $p = 0.025$ and $n = 438$? As mentioned in Appendix **What a "95% Prediction Interval" Really Means**, Titled Tuesday scores are *discrete*, which means that practically speaking "95% prediction intervals" often cover *more* than 95% of likely scores. Therefore, it would be wrong to assume that each player has a 2.5% probability of overperforming. The probability can differ for each player, and it's almost always below 2.5%.

A logical next question is: why not approach the problem as a *Poisson* binomial distribution, using the appropriate probability of overperforming p_i for each player? Treating the number of overperformers as a Poisson binomial distribution still fails because the probability of one player overperforming depends on whether other players overperform. In order for some players to overperform, other players have to lose (and not overperform). When we simulate, two players "share" the result of each game, so we take this dependence into account.

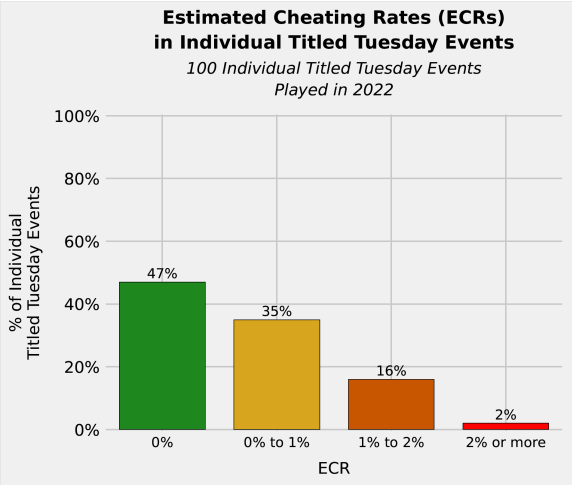
ECR by Year

ECR was above 0% in almost three-quarters of Titled Tuesdays played in 2023 – a higher share than in 2022 or in 2021 – but still almost always below 2%.

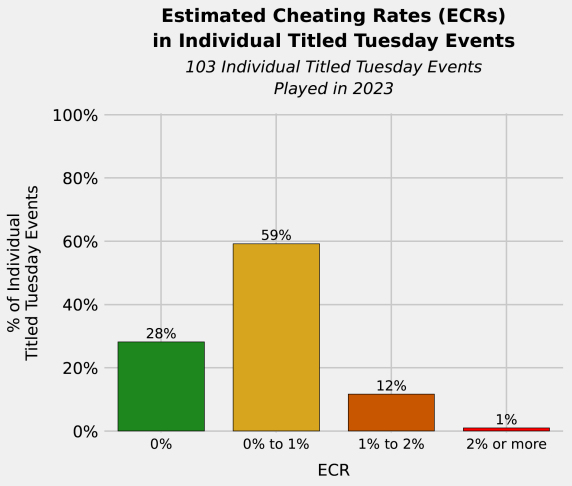
2021:



2022:



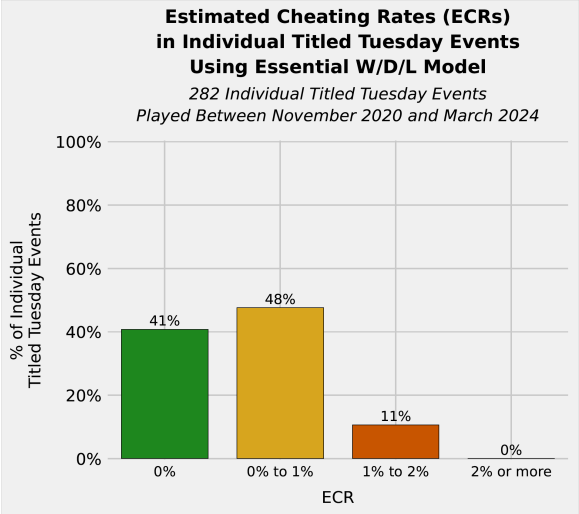
2023:



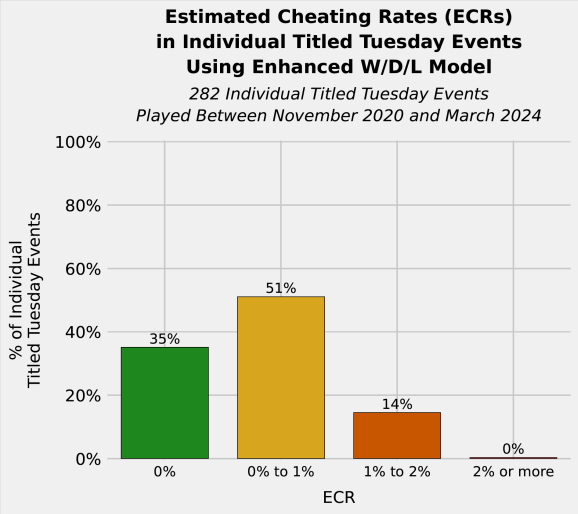
ECR by W/D/L Model

We found similar ECRs in each Titled Tuesday when using our Basic and Enhanced W/D/L models, each of which is a multinomial logistic regression that uses Chess.com blitz ratings at the exclusion of FIDE ratings. See the earlier Appendix **W/D/L Model Details** for more information.

Essential Model:



Enhanced Model:



Simulation Adjustment

Due to the computational complexity required to combine together and aggregate simulation results for *multiple* Titled Tuesdays, for this part of our analysis we simulated each game 1,000 times. When computing ECR for each *individual* Titled Tuesday event we simulated each game 10,000 times.