Is the mulligan rigged? Large study indicates that it is.

tl;dr
- Release: Bot that automatically gathers samples of mulligans.
- Data set: 4310 R1 mulligans gathered by the bot.
- Discovery 1: Mulligan bias likely to have been reduced slightly last patch.
- Discovery 2: The mulligan seems rigged to produce more extreme outcomes.
- Note: The rig is likely not a new mechanic, but the weights were increased last patch.
- Result: Up to half of the expected 'single repeats' are converted into '0/2/3-repeats'.

Context
Gwent's mulligan has been studied a lot since closed beta, but a common theme of previous studies has been data sets lacking in size. In conjunction with this study, a bot that gathers data automatically was developed. This gives us a data set that is void of human error, and far larger than anything previously available.

All of the software used in the study is made available in [this github package](#).

The raw data gathered for this study is made available in the 'raw_data' folder of the github package. Please read the readme's paragraph 'F' for an explanation of the syntax used.

The Mulligan
The two, main, logical, implementations of the mulligan phase are:
- M1: Insert mulliganed cards into the deck in-between each mulligan.
- M2: Set mulliganed cards aside. Insert them at the end of the mulligan phase.

Furthermore, the two, main, logical, variations of M1 are:
- M1a: [Choose card to mulligan] → [Insert mulliganed card] → [Draw replacement]
- M1b: [Choose card to mulligan] → [Draw replacement] → [Insert mulliganed card]

A big controversy regarding M1 is the fact that it creates *mulligan bias* (i.e., increased odds of mulliganed cards landing in the top three spots of the deck); M1b exhibiting somewhat less than M1a. M2 is void of such bias.

History
Gwent is believed to have started out with M1a as it's implementation. In the fall of 2017 a change was made, which reduced mulligan bias. CDPR comments at the time indicated a move to M1b, while studies at the time indicated that mulligan bias had been completely removed, leading to a belief in the community that Gwent, in fact, had moved to M2. While I can't say for sure that it didn't, at this point there's some indication that Gwent never moved beyond M1b.

The old studies might have fallen victim to small data sets, or existence of hidden mechanics (i.e. 'rigging'), leading to erroneous conclusions.

This winter, a revert to M1a seemed to occur, supported by a moderately large study. Finally, in the previous (April) patch, the mulligan was changed yet again.

Methodology
Data is gathered in practice mode with a deck of 26 different Bronze specials, using Calveit as leader. After mulliganing three cards, Calveit is played to reveal the top three spots of the deck.

Findings
Disclaimer: While the data set is unlikely to contain errors, keep in mind that my analysis could.

A summary of some of the most important metrics is given in this spreadsheet.

The following table compares the expected proportion of repeats in top three, under the main mulligan implementations, to the observed proportions.

	M1a	M1b	M2	Observed
No repeats	34.93%	41.92%	51.08%	54.39%
1 repeat	47.64%	45.67%	41.78%	25.85%
2 repeats	16.12%	11.75%	6.97%	16.68%
3 repeats	1.30%	0.66%	0.18%	3.09%

In the winter-study, a comparison like the one above was used to discard M1b and M2 as possibilities. Allthough M1a was a somewhat bad fit, it was within reasonable expectation, which led to the conclusion that Gwent had reverted to M1a.

This time around, it is immediately apparent that the observations cannot possibly be produced by any of the logical implementations alone. It seems that a large portion of what was supposed to be instances of '1 repeat' is re-distributed into instances of '0/2/3 repeats'. We're forced to conclude that, while Gwent likely uses one of the above implementations as it's foundation, it is rigged to produce far more extreme outcomes than expected.

So, which one is the foundation? In the summary-spreadsheet, comparison of the exact location the cards end up in is provided. This reveals a very good fit with M1b, indicating that one of the changes of the April patch was re-reverting back to M1b. The other change being the rig.

The Rig - When?

A natural question arises: Is this hidden mechanic new? It could be, but probably isn't. In the winter-study, the observations deviated from expecation in the same manner, just not nearly as strongly. This indicates that the rig was in place back then too, but with much milder weights, and that the weights were juiced up in the April patch.

There's no way to know for how long the mulligan has been rigged, but it could've been a hidden mechanic since the very beginning. That would certainly serve to explain CDPR's reluctance to share info about the mulligan implementation.

The Rig - Consequences?
An important aspect to note is the fact that the rig works "both ways" - i.e., producing, not only more instances of double and triple repeats, but also more instances of no repeats. As a result, the average number of mulliganed cards redrawn is balanced out as expected.

However, that is not to say that there aren't consequences of such a mechanic:
- 1-repeats turning into 0-repeats easily goes by unnoticed, while increased amounts of 2- and 3-repeats are very noticable, and feels bad for players.
- 1-repeats is, generally, a decent result for players. The single repeat can be mulliganed. Turning them into 2/3-repeats is almost always a disadvantage, and the times they are turned into 0-repeats doesn't make up for it (players might lack a "bad" card to mulligan).
- The rig doesn't seem to affect the three mulliganed cards uniformly. It uses card #2 and #3 to produce the outcomes; either moving them out of-, or into, the top three spots of the deck to produce the extra 0/2/3-repeats. This causes unexpected biases - for instance, instances of 1-repeats has card #1 in top three far more often than expected.

The Rig - How?
Finally, exactly *how* is the rig implemented? I simply don't know. There's an infite ways to rig the mulligan, which means - not only is it hard to come up with an algorithm producing similar results - even if you did you'd have no way of knowing it's the right one. In the end, it doesn't really matter how it is rigged, as long as we know that it is, and what the outcomes are.