

Who Would Win in A Fight?

EXAMINING BALANCE IN INFINITE CRISIS

Shawn Connor

*Sr. Statistical Analyst
WB Games / Turbine*



The Game

- Infinite Crisis is a multiplayer online battle arena (MOBA) game set in the DC Multiverse
- Players form teams, either with their friends or with random players
- Each teammate choose a character (“Champion”) from their roster of heroes and villains
- Then... BATTLE!

The Champions

- Each Champion has unique powers and abilities
- Each Champion has a “role”, a broad category of competence
- Each player has access to a roster of Champions
 - Tutorial Champions
 - Free rotation
 - Champions bought

“Balance”

- Every Champion needs to be a viable, fun choice; no must-haves, no must-nots
- In the competitive atmosphere of MOBAs, the balance of individual Champions is taken very seriously by players
- But! Effectiveness is a result of many factors: player skill, customization opportunities, choices made by allies and opponents, and even player perception
- Designers need to be able to understand the performance of each Champion in a detailed and nuanced fashion, so they can adjust and iterate

The Data

- A long stream of event data is fired off by the game's servers during play
- We instrument everything we can, as efficiently as we can
- To drive analytics, we materialize the events in a convenient RDBMS collector
- For exploration, visualization, and collaboration, we use ad-hoc queries to pull this data to Tableau

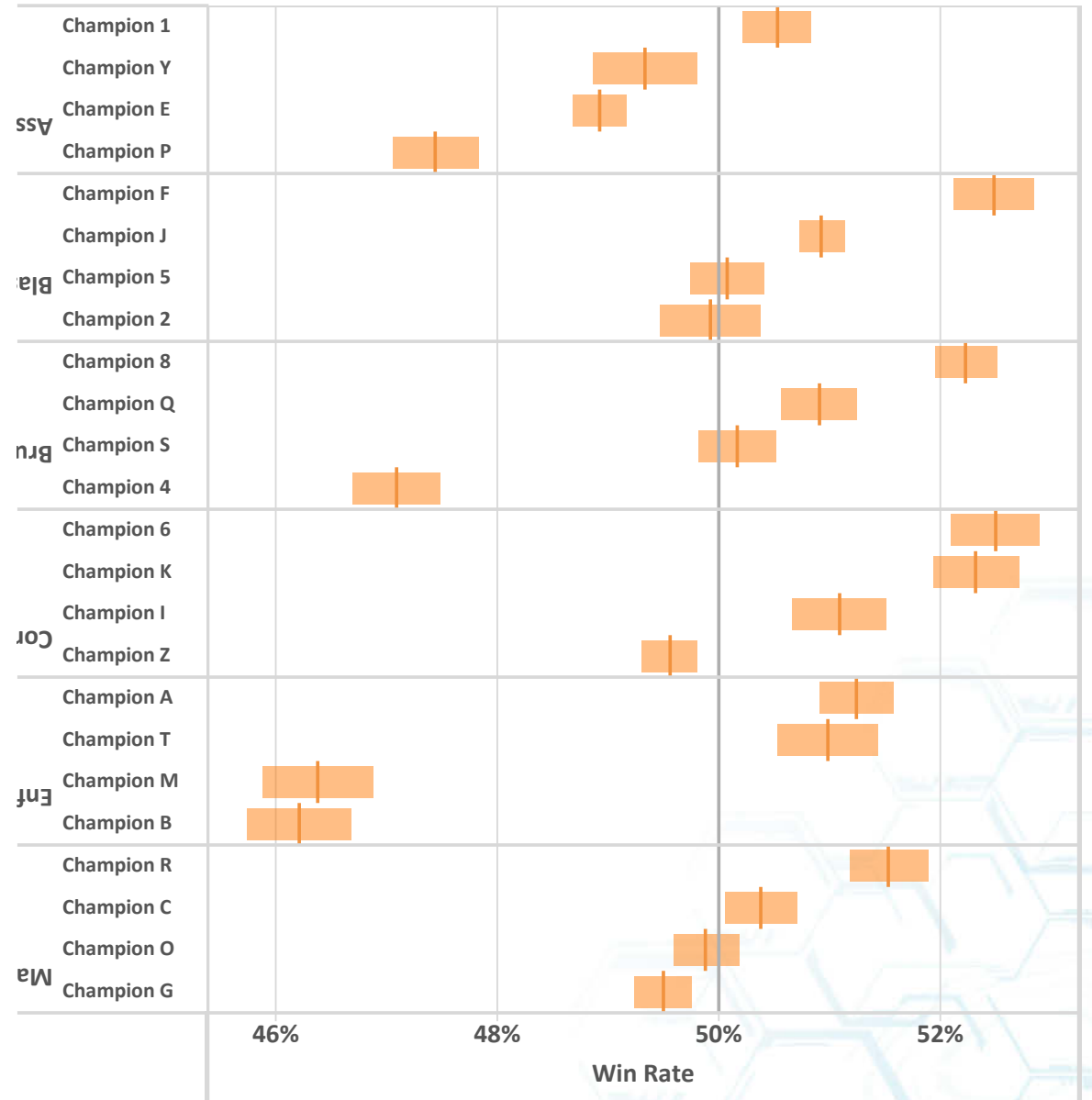
The Basics

First, some noise reduction

- Mirror Matches – Champion appears on both teams
- Incomplete Data – One or more missing scorecards
- Participation Gaps – One or more mid-match disconnects
- Time Context – Champions are always changing

The sample gets smaller, but the results more meaningful. Broad confidence intervals are okay!

(Names and roles have been randomized)

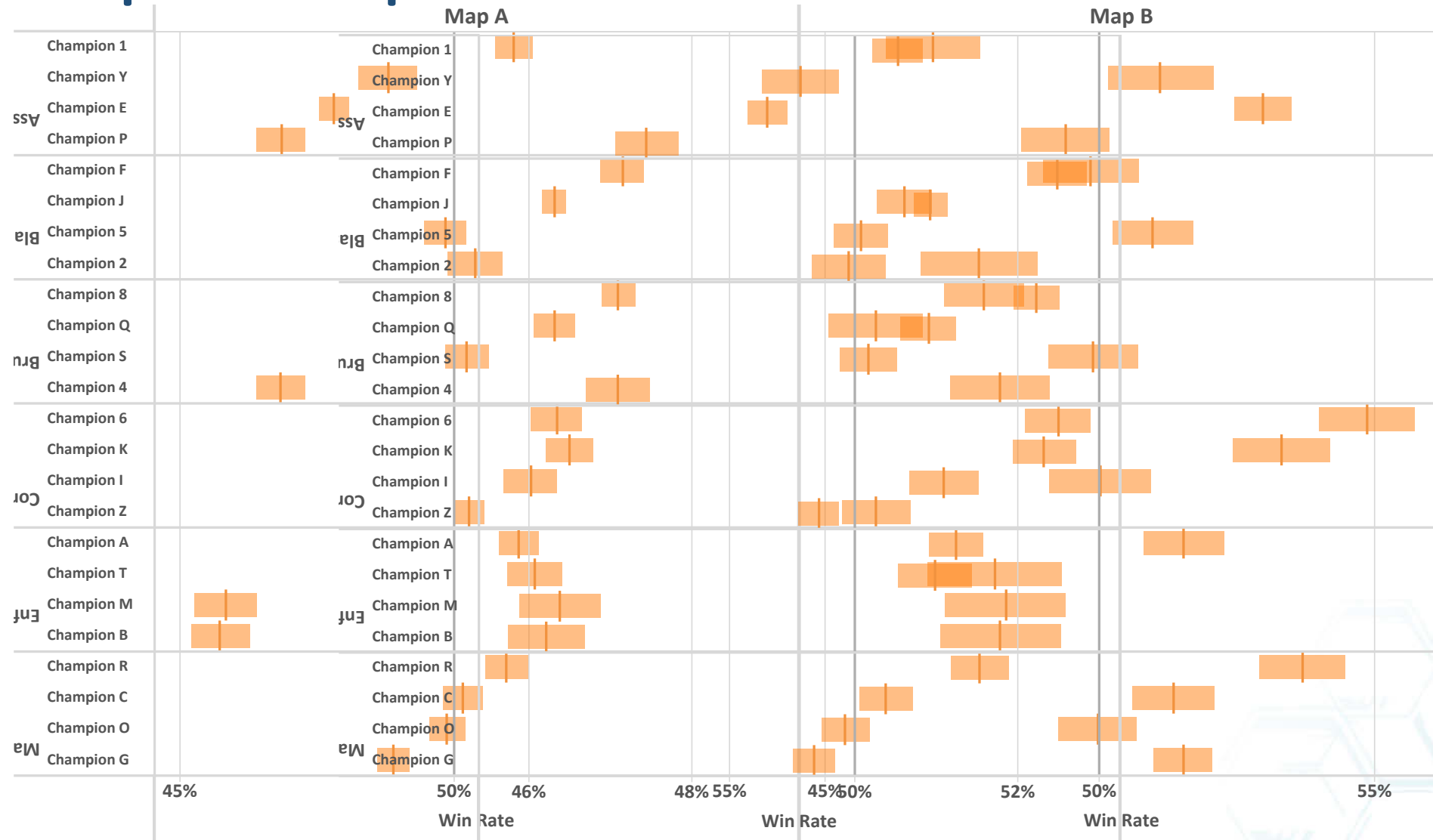


But...

There are lots of confounding factors!

- Champions perform differently in different game modes
- Some Champions are strong early, some are strong late-game
- Some Champions have steeper learning curves
- **Skilled *players* tend to win, no matter what Champion they select**

Example: Maps



Getting Tricky: Player Skill

Try as we might, teams are not always of equal skill!

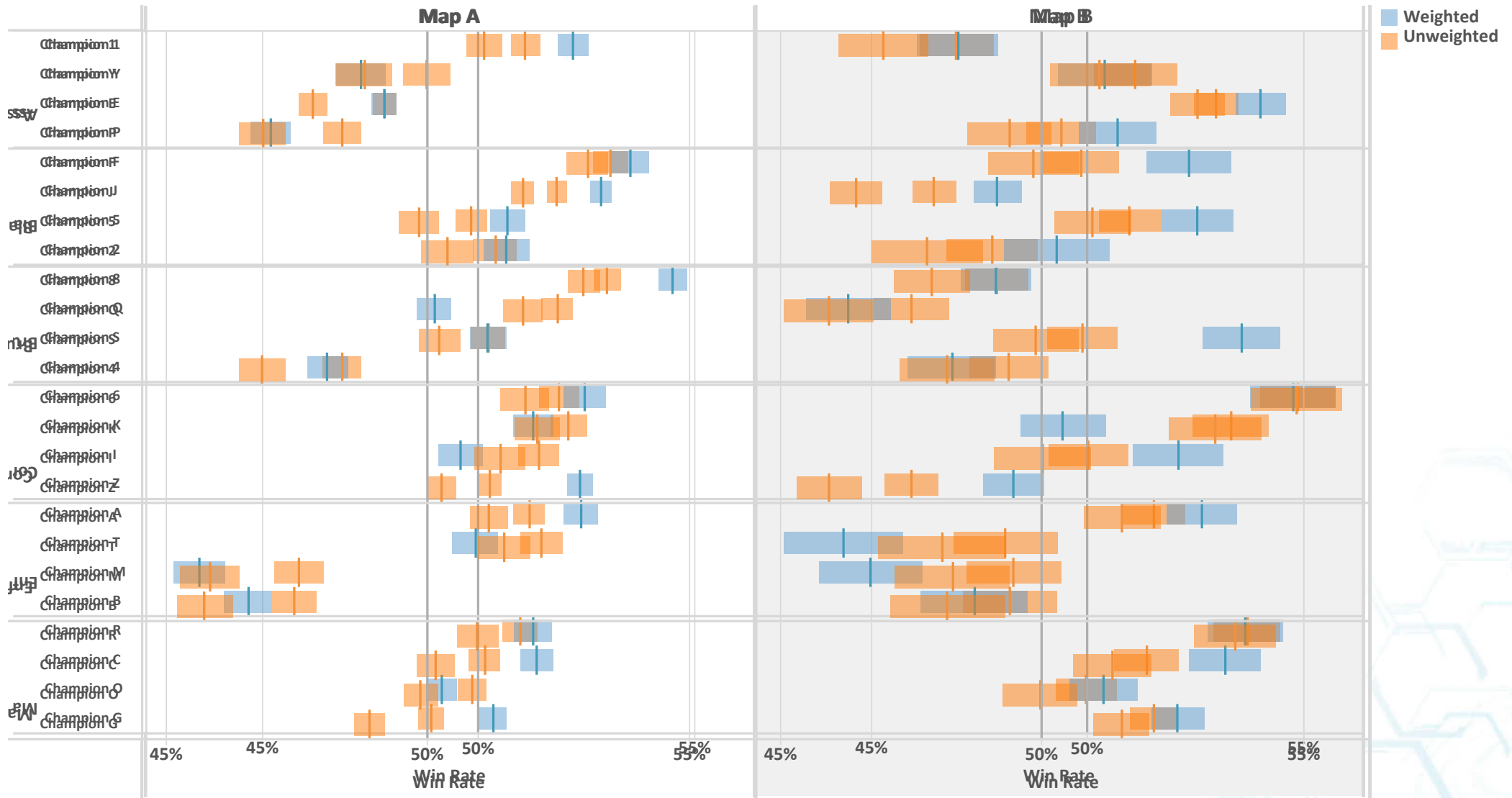
- Some players are new and haven't shown us what they can do yet
- Some players play in premade teams, overriding programmatic balancing
- Players who are extremely good (or bad) are rare, and hard to match
- Some Champions (especially those that are not currently free) are predominantly played by skilled users

Hmm... who else has this problem?

Place Your Bets

- Based on our running assessment of each player's skill, calculate the expected odds that each side will win
 - “Borrow” the same system that assesses skill and calculates rewards after the match
- Weight each win or loss by the expectation
 - When the expected happens, we have a low-weight result (i.e. low-reward bet)
 - When our expectations are challenged, we potentially learn more
 - Given large samples, improbable results should be balanced by the probable ones
 - Hooray for Bayesian stats!
- After weighting, each individual outcome is now represented by a value that can be summed, averaged, etc. *just like a binary win/loss!*

The Result



Questions We Can Begin to Answer

- Are “Overpowered” Champions really overpowered, or are they just disproportionately played by skilled players?
- Is a Champion behaving as the designer intended?
- Are anecdotal reports of imbalance quantitatively supported?
- What is the “true” utility of a Champion?

With further filtering, we can segment by match length, player skill, etc. to inform tactical balance changes

Further Study

Given this model, we have two easy paths for improving our assessments

1. Add more factors to the prior probability calculation

- Experience with the selected Champion
- Experience with other Champions of the same role
- Experience with the map
- *Anything we can know about the whole team – Teams win or lose together*

2. Create more dimensions for the results; more dimensions of “effectiveness”

- How long the match took
- Teammate and opponent Champion choices
- What “position” each user played on the map (telemetry)
- Intra-match advancement (item builds, skills chosen)
- *Anything we observe happening during or after the match – “Batman tends to win more when played like...”*

Questions?