

# Estimating the effect of 'diverse' team compositions on Dota 2 game outcomes using Inverse Probability Weighting

## Team Diversity

Dota 2 hero categorizations:

**Attack-type:** ranged, melee

**Attribute:** Intelligence, agility, Strength

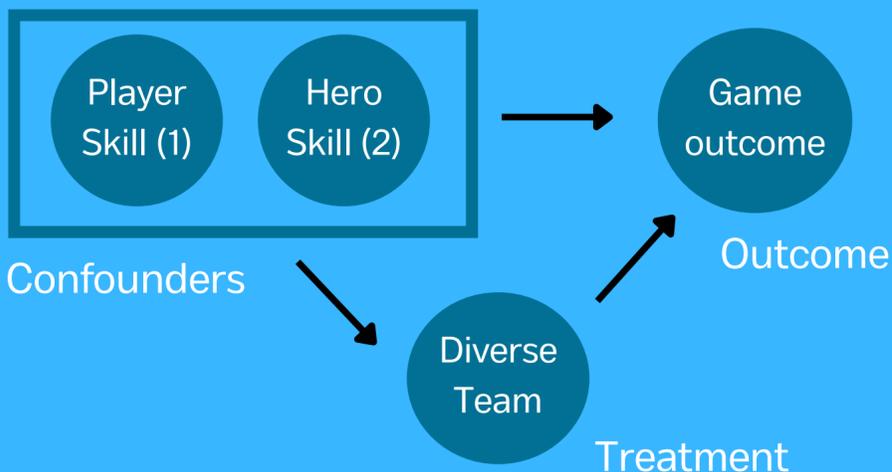
**Roles:** Carry, Nuker, Initiator, Disabler, Durable, Escape, Support, Pusher, Jungler



Highlighted diversity metrics:

3	Team contains each Attribute at least once
4	Team contains more than 12 non-support and non-carry roles
6	Metric 01 & Metric 02 & Metric 03 & Metric 04

## Confounders



- Team MMR difference (1): Average ranking of the team compared to the opponent
- Average team Hero score (2)

## Inverse Probability Weighting

- **Logistic Regression Model** -> probability of treatment given confounders (  $\Pr [T | C]$  )
- **Inverse of  $\Pr [T | C]$**  -> weight a team gets assigned
- All the data reweighted -> pseudo-population
- Pseudo-population-> **association = causation\***

## 1

## Results

## 4

- Dataset formed with **397 games** (794 team compositions)
- Average Treatment Effect

$$ATE = E[Y_1 - Y_0]$$

Results with IPW correction:

nr.	ATE	Mean	StDev	95% CI	% with treat.
3	-0.3	-0.33	3.74	(-0.57, -0.1)	66,88
4	4.63	4.70	3.54	(4.48, 4.92)	54.53
6	7,72	7,80	4,48	(7.52, 8.08)	19,14

Results without IPW correction:

nr.	ATE	Mean	StDev	95% CI	% with treat.
3	0,28	0,26	3,73	(0.03, 0.49)	66,88
4	4,83	4,91	3,54	(4.69, 5.13)	54.53
6	5,70	5,76	4,48	(5.48, 6.04)	19,14

- Metrics have varying impact on game outcome
- Small ATE difference between with IPW and without IPW correction (exception metric 6)
- Are confounders influencing the treatment?

## Conclusion

## 5

- Diverse teams (depending on the metric) -> higher % of winning
- Confounders need to actively influence the treatment

## What's Next?

## 6

- **Discrete treatment vs Continuous treatment:** the relationship between the confounders and a continuous treatment might be more apparent
- **Larger sample size:** helps determine whether confounders are not actively influencing treatment or whether Logistic Regression Model was inaccurate