Estimating individual contributions to team success in women's college volleyball

Scott Powers, Luke Stancil and Naomi Consiglio

NESSIS 2023

RICE UNIVERSITY
Sport Analytics

Outline

Act 1: Estimating Point Win Probability

Technique: Markov Chain Model

Act 2: Evaluating Individual Contributions

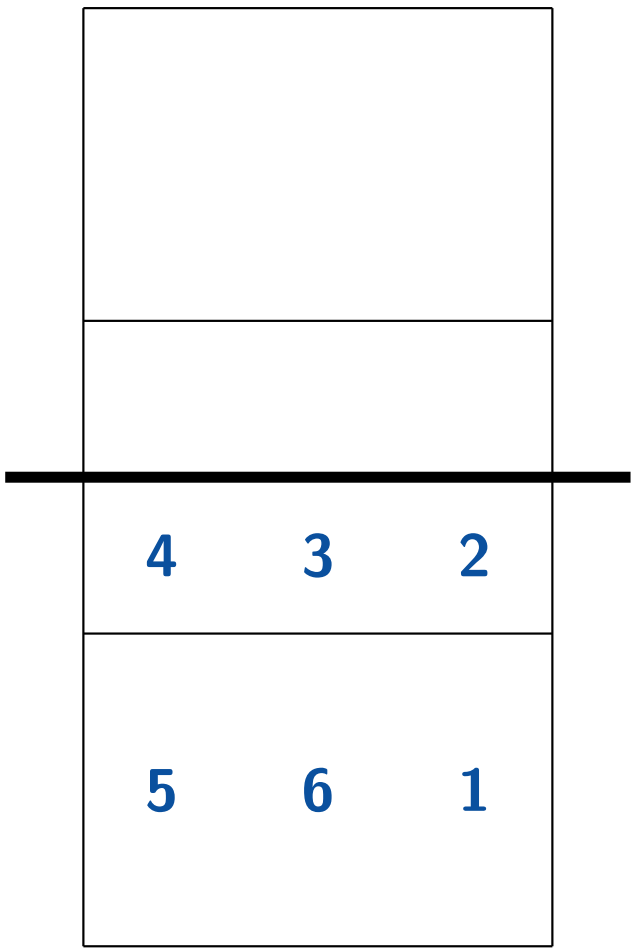
Technique: Domain Knowledge

Act 3: Adjusting for Strength of Schedule

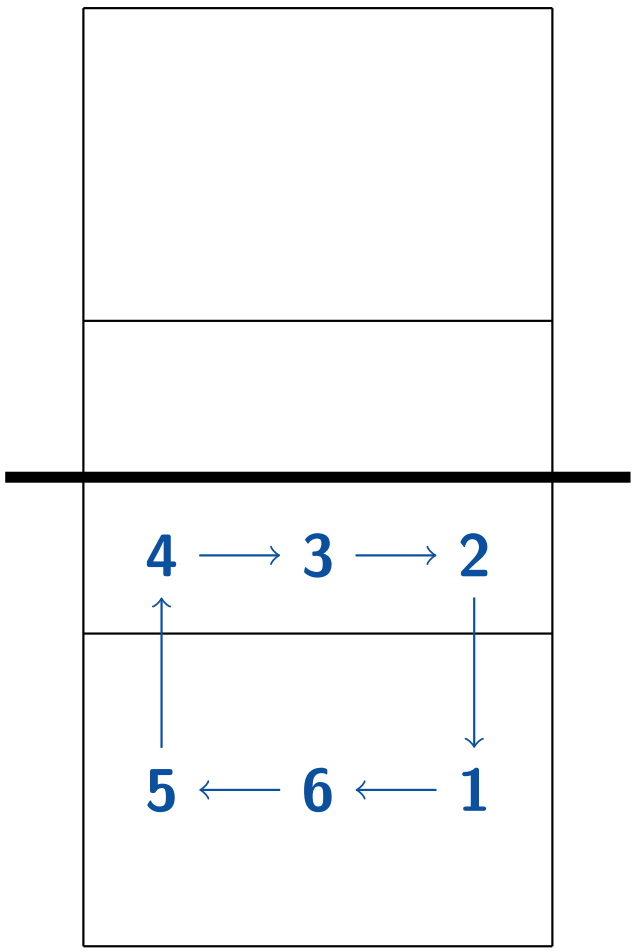
Technique: Linear Mixed-Effect Models

(Act 0: Introduction to Volleyball)

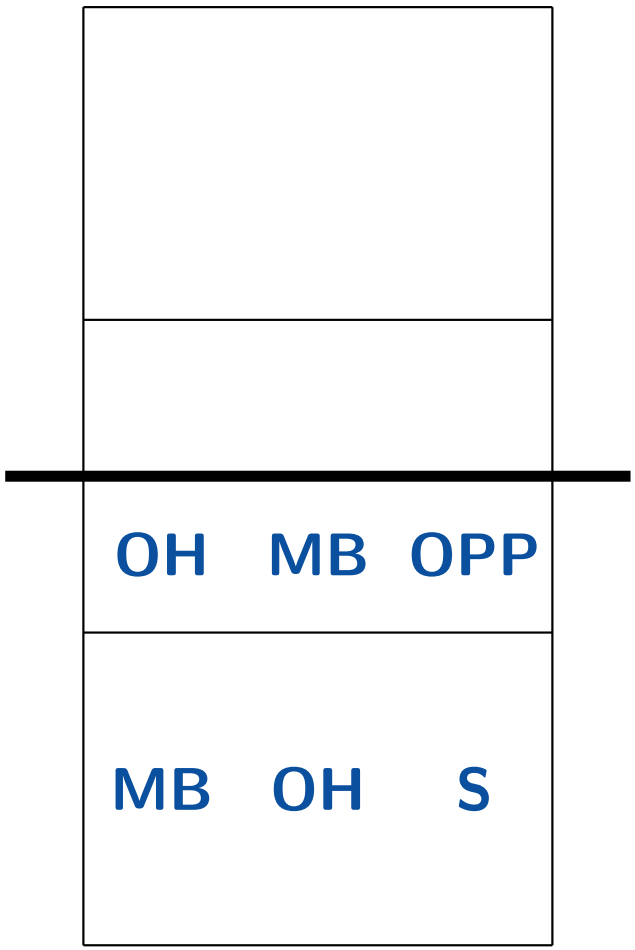
Introduction to Volleyball



Introduction to Volleyball



Introduction to Volleyball



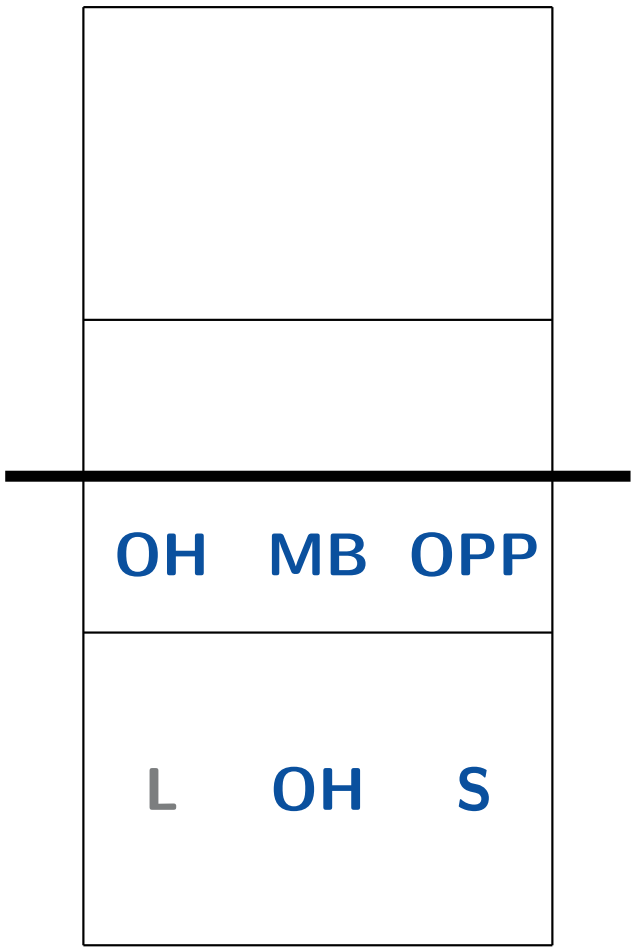
Setter (S)
Setting

Outside Hitter (OH)
Attacking, Passing

Middle Blocker (MB)
Blocking, Attacking

Opposite Hitter (OPP)
Attacking, Blocking

Introduction to Volleyball



Setter (S)
Setting

Outside Hitter (OH)
Passing, Attacking

Middle Blocker (MB)
Blocking, Attacking

Opposite Hitter (OPP)
Blocking, Attacking

Libero (L)
Passing

Existing metrics

- Standard metrics
  - Serving: Ace%, Error%
  - Receiving: Error%, Passer Rating
  - Digging: Digs / Set, Digs / Opportunity
  - Setting: Assists / Set
  - Attacking: Hitting Efficiency = (Kills – Errors) / Attempts
  - Blocking: Blocks / Set

- State of the art
  - Fellingham (JQAS 2022): PAAPS
    - Similar to regularized adjusted plus-minus in basketball
  - Gordon (volleydork.com): Value Added above Expectation
    - Very similar to the present work

Act 1: Estimating Point Win Probability

Technique: Markov Chain Model

Example: First Point of 2022 National Championship

Texas Louisville

Player	Skill	Eval	(X, Y)	Attack Code
Anna Deeber	Serve	–	(2.99, -0.13)	
Emma Halter	Reception	#	(0.93, 5.80)	
Saige K.-Torres	Set	#	(2.13, 3.13)	
Molly Phillips	Attack	–	(3.33, 3.20)	X6
Raquel Lazaro	Dig	+	(0.86, 4.98)	
Elena Scott	Set	#	(2.99, 1.65)	
Claire Chaussee	Attack	–	(0.63, 2.83)	V5
Kayla Caffey	Block	+	(3.26, 3.43)	
Phekran Kong	Dig	!	(0.89, 3.13)	
Raquel Lazaro	Set	#	(0.97, 2.61)	
Claire Chaussee	Attack	#	(0.67, 2.91)	X5

Evaluation Codes: $\# > + > ! > - > / > =$

Dataset: 4,147 matches, 600K+ points, 5M+ contacts, ∼6,000 players

Markov Chain Model: Game State

Definition: A **volley** is a sequence of consecutive contacts by the same team

The game state on each contact is given by:

- Whether the team started the point by serving or receiving
- The sequence of contacts made during the current volley (including evaluation code *except* for contacts ending a volley)

Terminal states: (S, P) and (R, P)

Example: (S, D#) $\rightarrow$ (S, D#S#) $\rightarrow$ (S, D#S#A) $\rightarrow$ (R, P)

Example: First Point of 2022 National Championship

Player	Skill	Eval	State	P(Sideout)
Anna Deeber	Serve		(S, SV)	57%
Emma Halter	Reception	#	(R, R#)	63%
Saige K.-Torres	Set	#	(R, R#S#)	64%
Molly Phillips	Attack		(R, R#S#A)	64%
Raquel Lazaro	Dig	+	(S, D+)	49%
Elena Scott	Set	#	(S, D+S#)	47%
Claire Chaussee	Attack		(S, D+S#A)	47%
Kayla Caffey	Block	+	(R, B+)	56%
Phekran Kong	Dig	!	(S, D!)	51%
Raquel Lazaro	Set	#	(S, D!S#)	51%
Claire Chaussee	Attack		(S, D!S#A)	51%
Point Louisville				**0%**

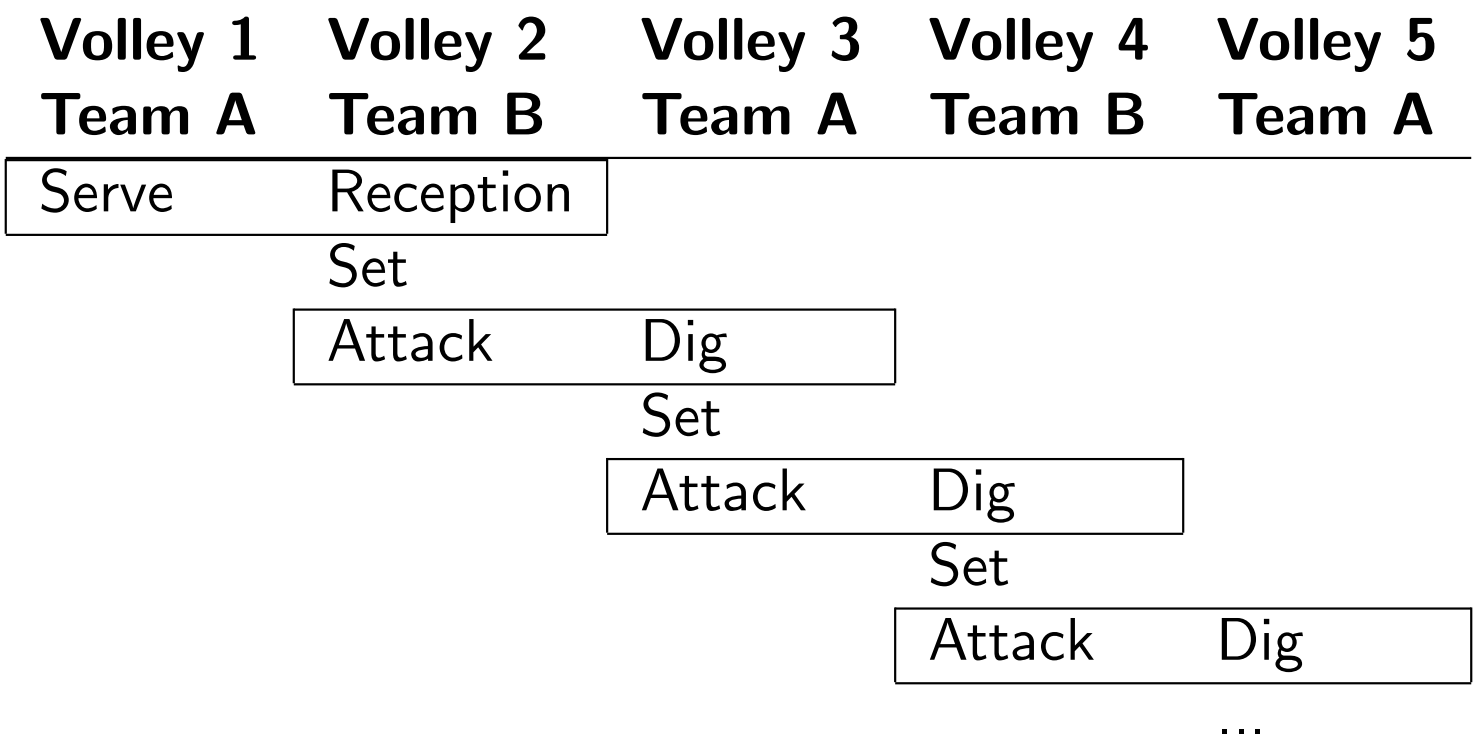
Act 2: Evaluating Individual Contributions

Technique: Domain Knowledge

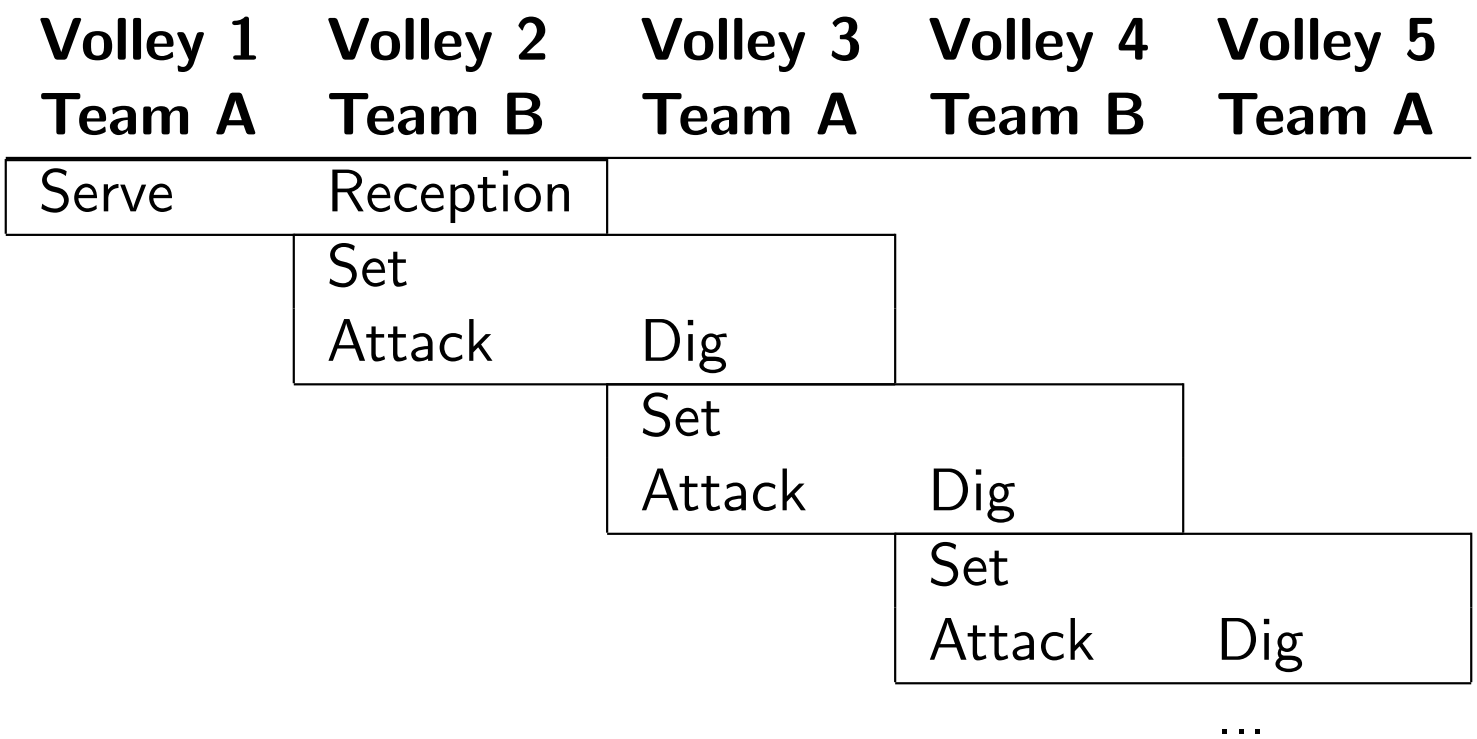
How a Point Progresses

Volley 1 Team A	**Volley 2 Team B**	**Volley 3 Team A**	**Volley 4 Team B**	**Volley 5 Team A**
Serve				
	Reception			
	Set			
	Attack			
		Dig		
		Set		
		Attack		
			Dig	
			Set	
			Attack	
				Dig
				...

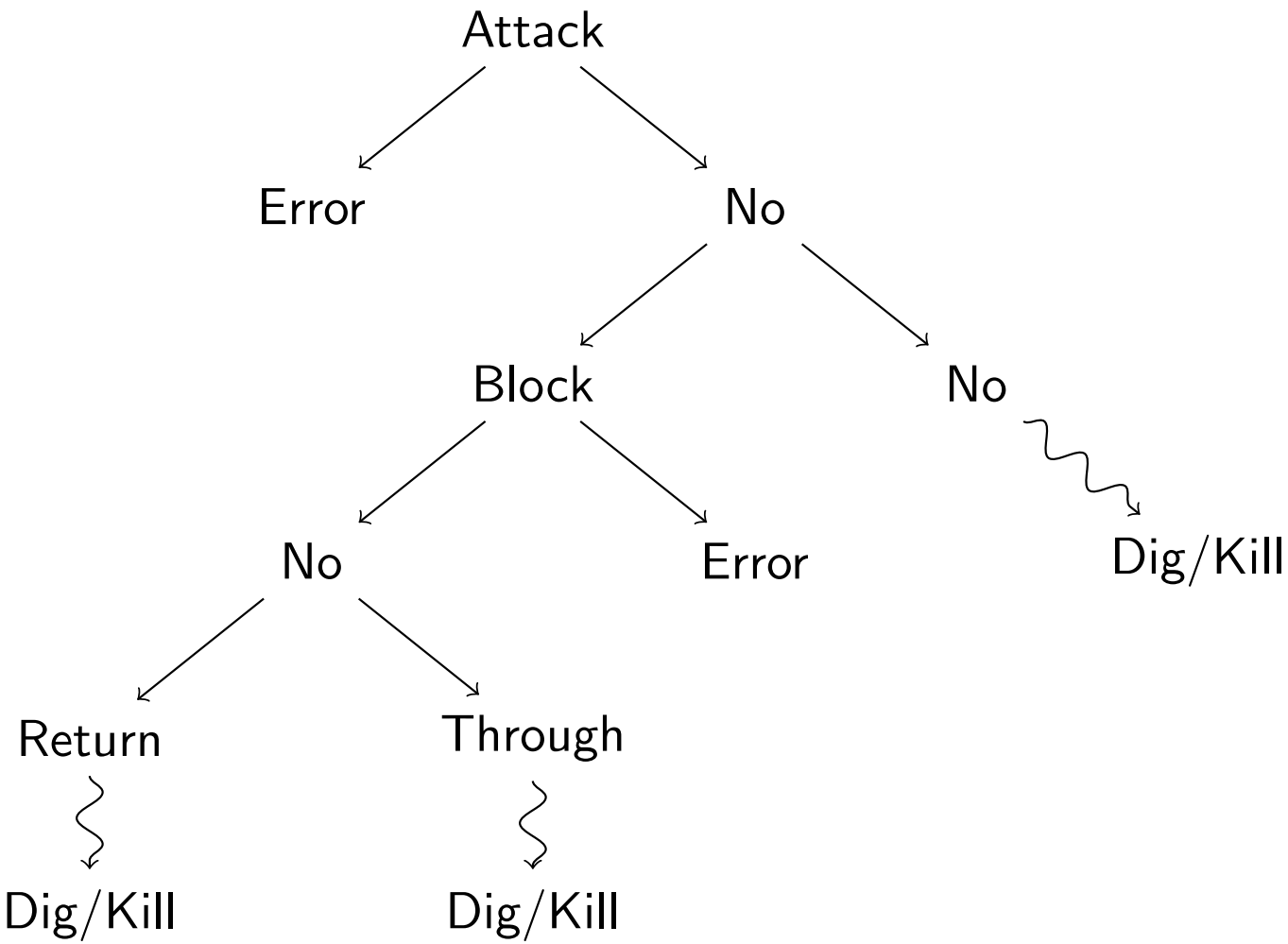
How a Point Progresses



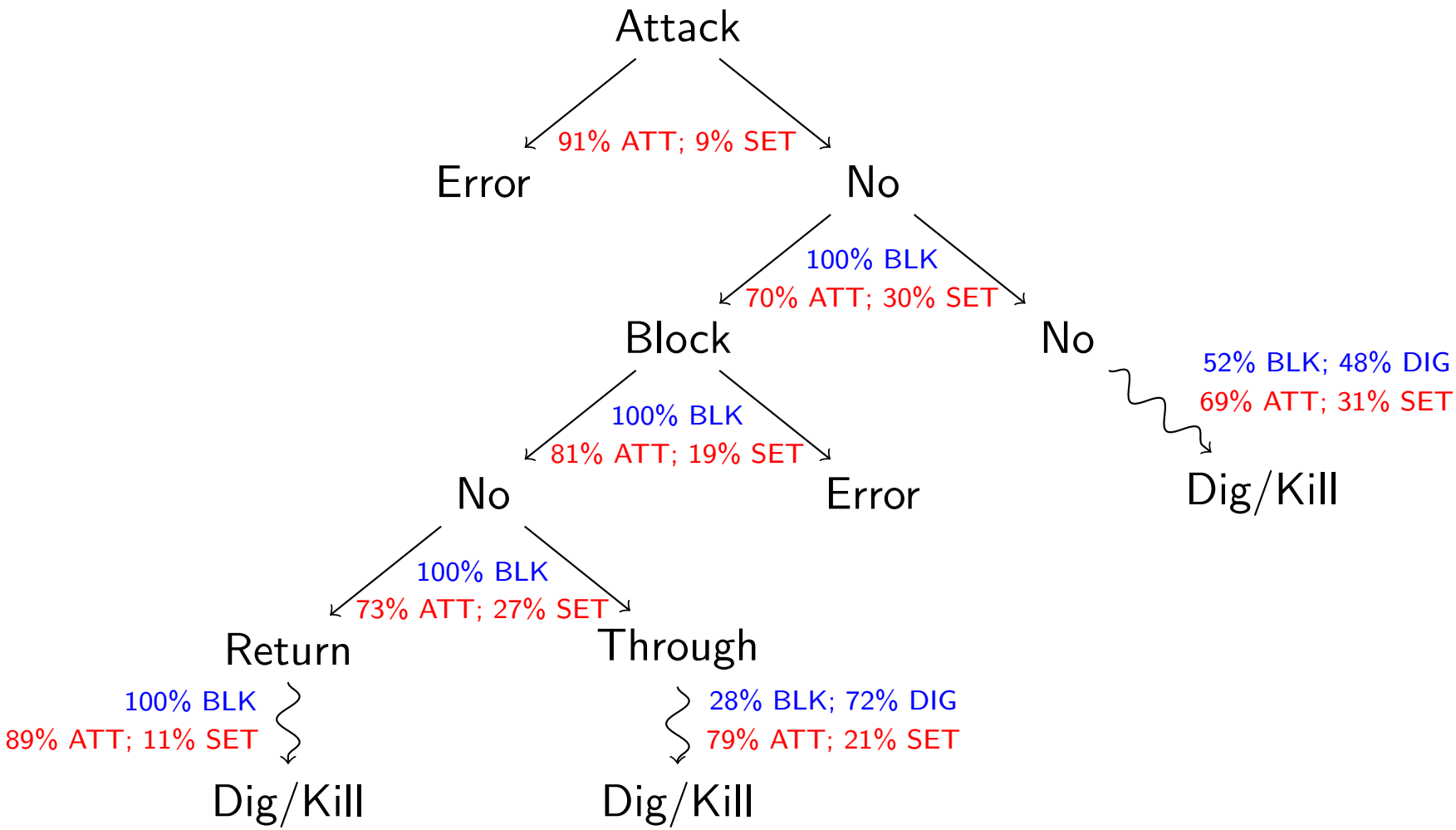
How a Point Progresses



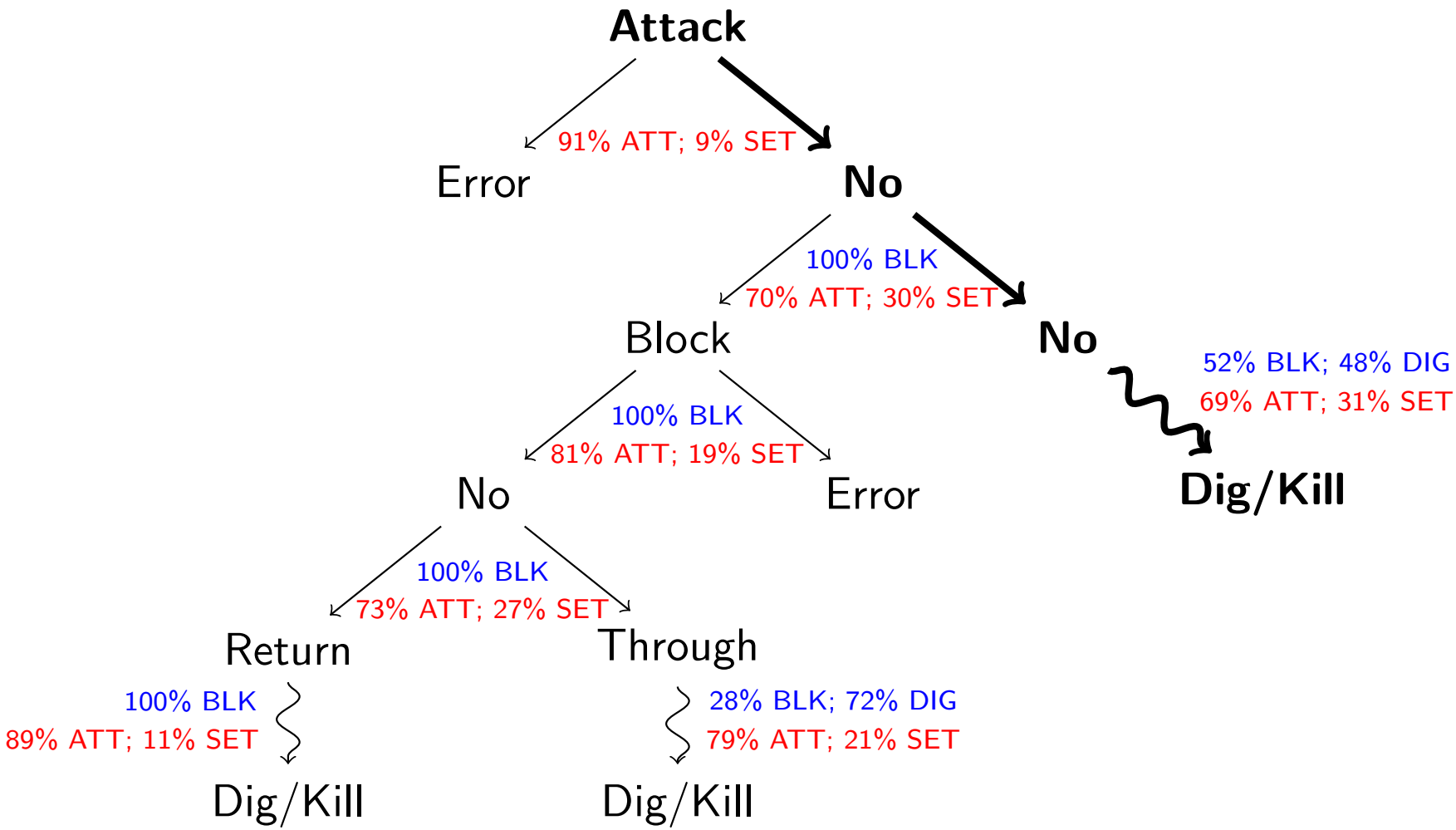
Attack Outcome Tree



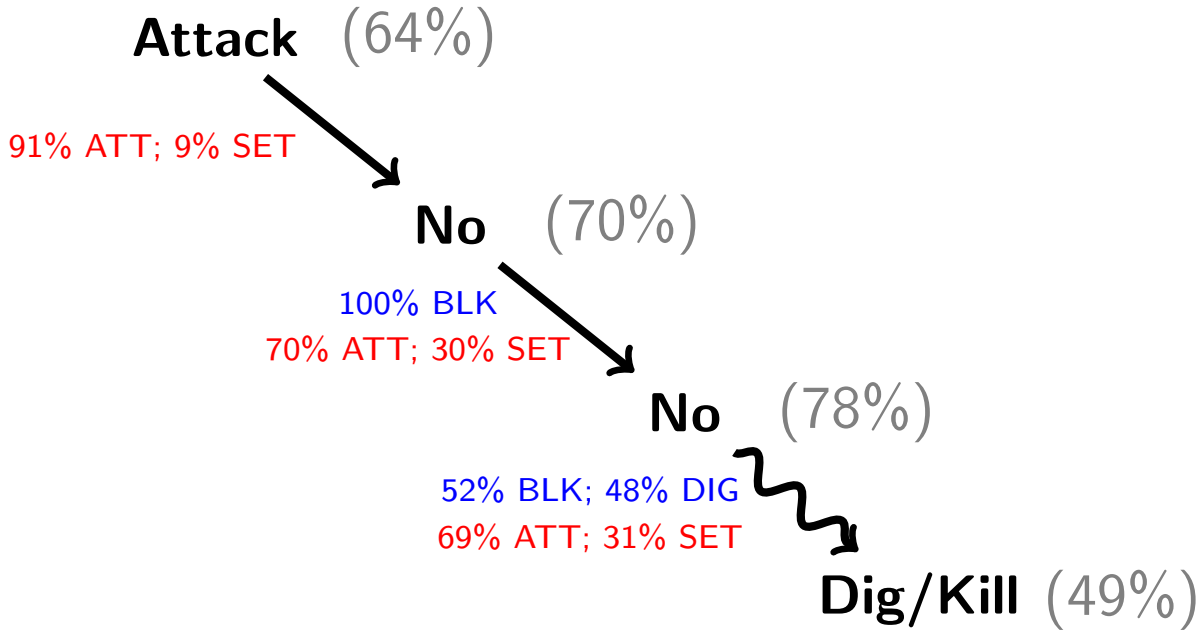
Attack Outcome Tree



Attack Outcome Tree



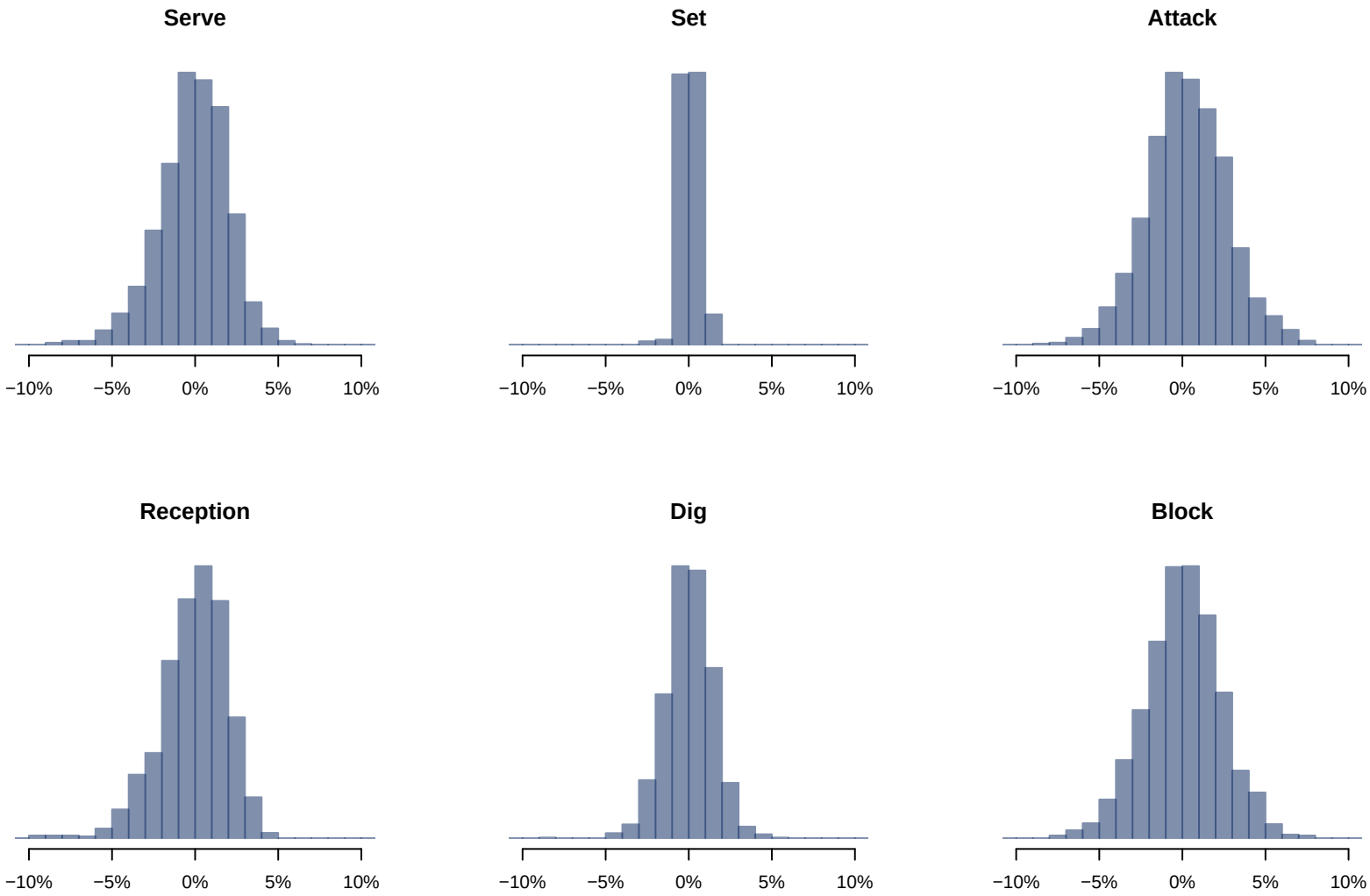
Example: First Attack of 2022 National Championship



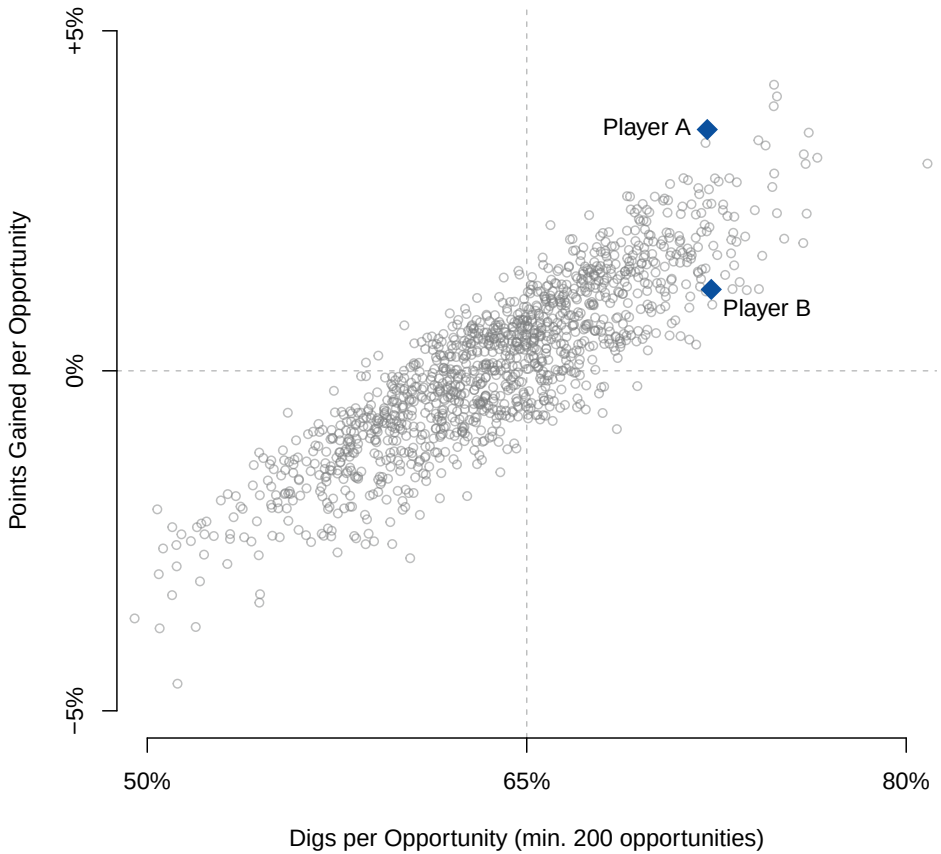
TOT	+6%	+8%	–29%	–15%	Standard Stats
ATT	+5%	+6%	–20%	–9%	1 attempt, 0 kills, 0 errors
SET	+1%	+2%	–9%	–6%	0 assists
BLK	—	–8%	+15%	+7%	—
DIG	—	—	+14%	+14%	1 opportunity, 1 dig

Caution: Blocker/digger assignment is a work in progress

Distribution of Points Gained per Opportunity



Standard vs. Advanced Metrics: Back Row Defense

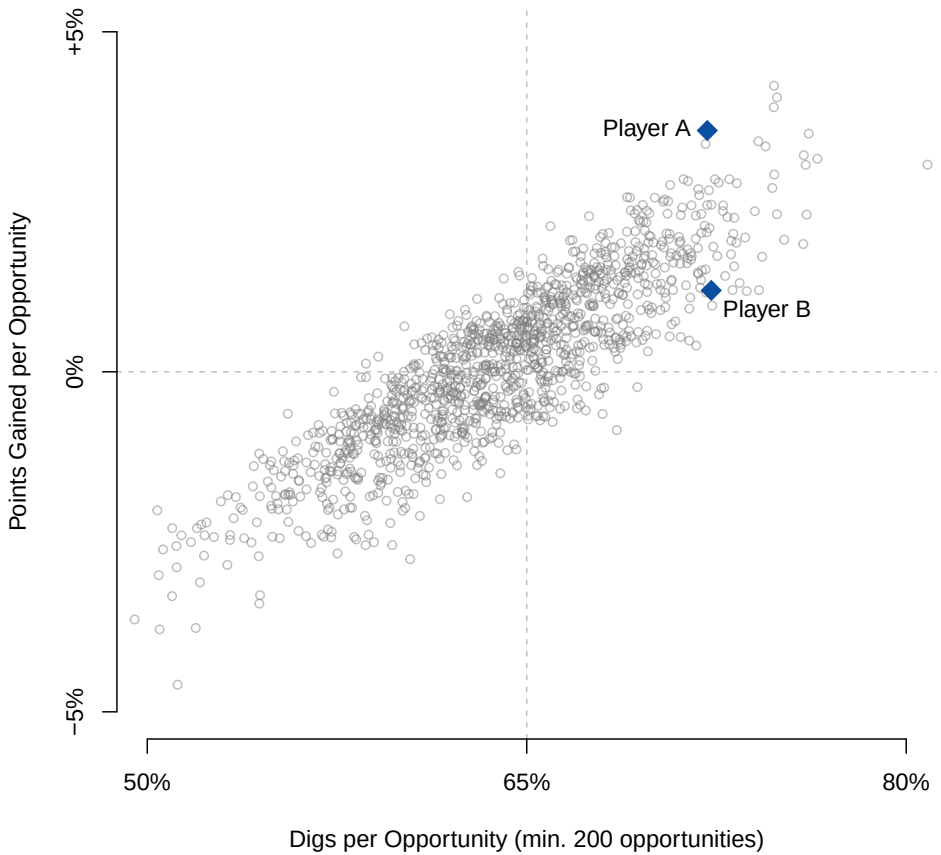


Player A:
72% digs per opportunity
+3.6% PG per opportunity

Player B:
72% digs per opportunity
+1.2% PG per opportunity

Caution: Dig evaluation codes are biased against setters by design

Standard vs. Advanced Metrics: Back Row Defense



Player A:
72% digs per opportunity
+3.6% PG per opportunity
No block touch: 85%
Perfect dig rate: 48%

Player B:
72% digs per opportunity
+1.2% PG per opportunity
No block touch: 66%
Perfect dig rate: 36%

Caution: Dig evaluation codes are biased against setters by design

Act 3: Adjusting for Strength of Schedule

Technique: Linear Mixed-Effect Models

Server vs. Receiver

$$\text{Exp. Points Gained} = \beta_{\text{Server}} + \delta_{\text{Receiver}} \quad \text{(good)}$$

$$\text{Exp. Points Gained} = (\beta_{\text{Team}} + \beta_{\text{Server}}) + (\delta_{\text{Team}} + \delta_{\text{Receiver}}) \quad \text{(better)}$$

$$\text{Exp. Points Gained} = (\beta_{\text{Conf}} + \beta_{\text{Team}} + \beta_{\text{Server}}) + (\delta_{\text{Conf}} + \delta_{\text{Team}} + \delta_{\text{Receiver}}) \quad \text{(best)}$$

- Fit random-effects model using lme4 package in R

Server:

Adj. Points Gained $\approx$ Points Gained – $(\hat{\delta}_{\text{Conf}} + \hat{\delta}_{\text{Team}} + \hat{\delta}_{\text{Receiver}})$

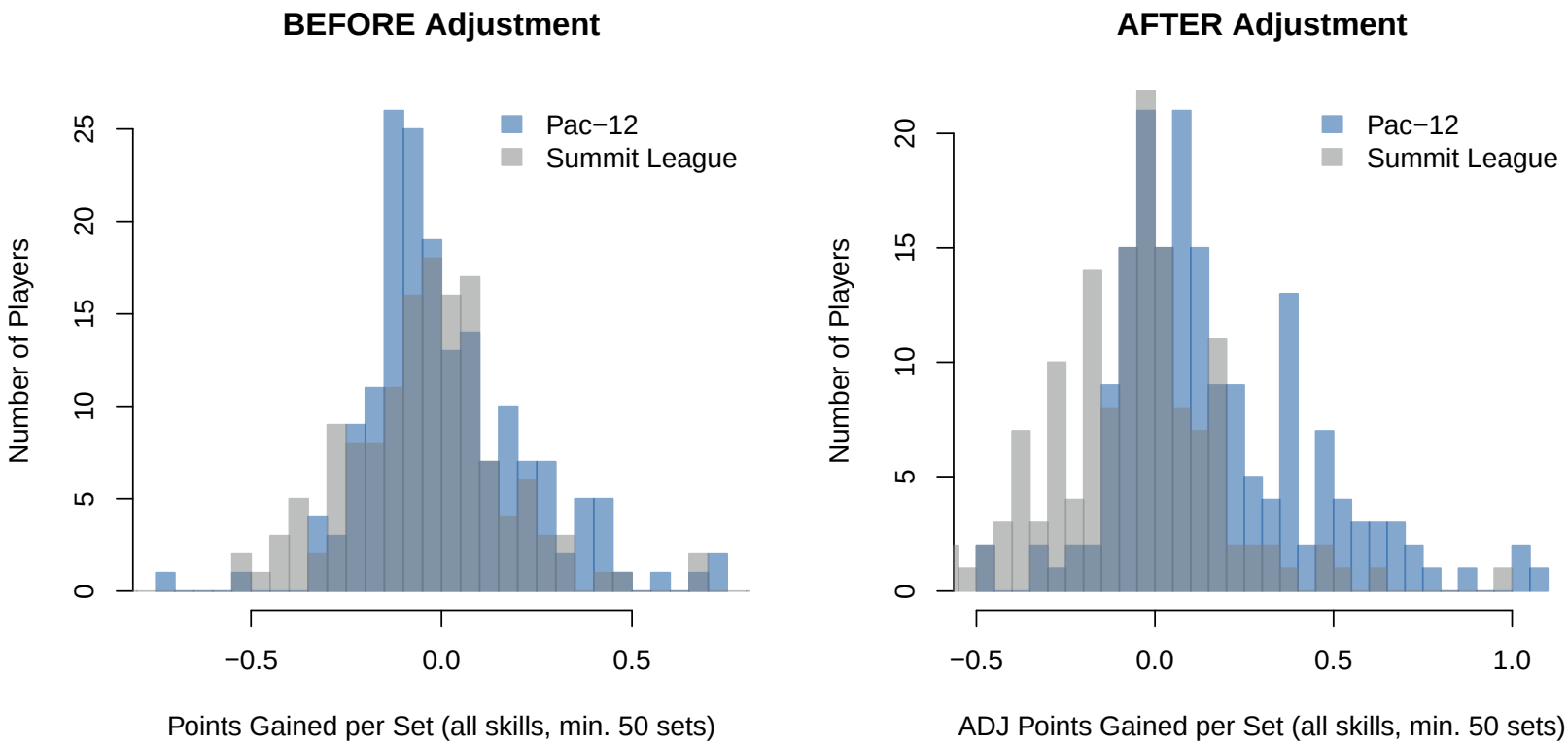
- Requires a de-biasing step
- Some generalization required for extension to other skills

Results: Top 10 Conferences (all skills)

Conference	Avg SoS
Big Ten	+0.23
Pac-12	+0.23
SEC	+0.21
Big 12	+0.20
ACC	+0.15
West Coast	+0.09
American	+0.04
Big West	+0.04
Mountain West	+0.02
Mid-American	+0.01

SoS units: points gained per set

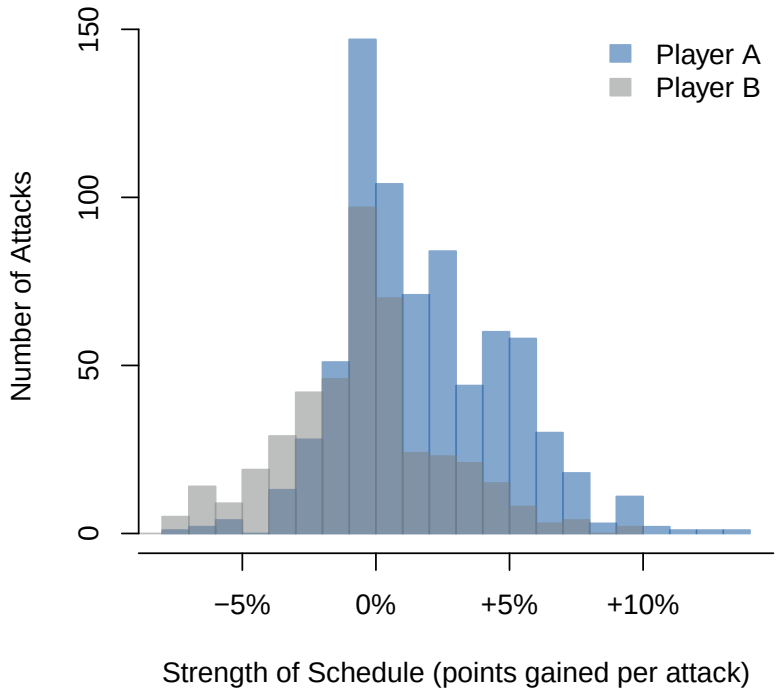
Example: Conference Comparison (all skills)



- Separation between conferences is evident
- One elite player from the Summit League still stands out

Caution: Additive assumption is not literally true in real life

Example: Teammate Comparison (outside hitters)



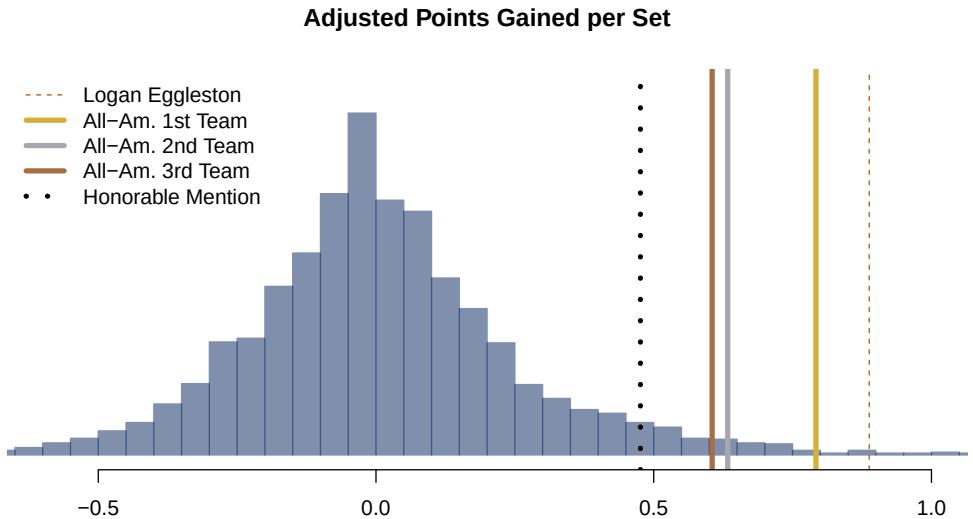
- Player A average SOS: +1.4% (PG per attack)
- Player B average SOS: –0.1% (PG per attack)

Toughest SoS: Player A vs. Nebraska, +13.0%
Kaitlyn Hord blocking, Lexi Rodriguez digging

Caution: SoS depends on which zone the attacker hits

(Act 4: Discussion)

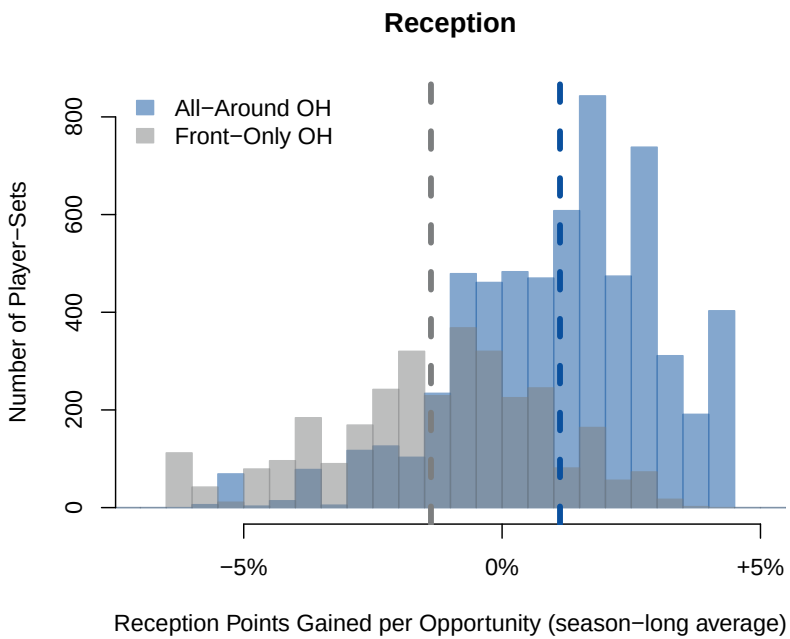
Question: Where does Logan Eggleston rank?



PLAYER	TEAM	CONF	POS	SETS PLAYED	POINTS GAINED PER SET (ADJ)	SERVE PG*/S	PASS PG*/S	SET PG*/S	ATTACK PG*/S	BLOCK PG*/S
Brooke Nuneviller	Oregon	Pac-12	OH	122	**+1.09**	+0.07	+0.41	+0.00	+0.56	+0.04
Mckenna Melville	Central Florida	AAC	OH	104	**+1.09**	-0.14	+0.23	-0.00	+0.79	+0.22
Claire Hoffman	Washington	Pac-12	OH	112	**+1.04**	+0.13	+0.23	-0.00	+0.65	+0.02
Julia Bergmann	Georgia Tech	ACC	OH	86	**+1.03**	+0.09	+0.25	-0.01	+0.64	+0.06
Kendall Kipp	Stanford	Pac-12	OPP	117	**+1.02**	+0.03	-0.02	-0.00	+0.72	+0.29
Amber Igiede	Hawaii	Big West	MB	102	**+0.98**	+0.07	+0.04	+0.01	+0.47	+0.38
Elizabeth Juhnke	South Dakota	Summit	OH	113	**+0.96**	+0.01	-0.01	-0.00	+0.69	+0.26
Madi Kubik	Nebraska	Big Ten	OH	109	**+0.94**	+0.05	+0.42	-0.01	+0.44	+0.05
Asjia Oneal	Texas	Big 12	MB	87	**+0.93**	+0.05	+0.04	+0.00	+0.35	+0.50
Logan Eggleston	Texas	Big 12	OH	91	**+0.89**	+0.09	+0.05	+0.01	+0.70	+0.05

Application: Defensive Specialist Strategy

- 63% of teams replace at least one OH with a DS in back row



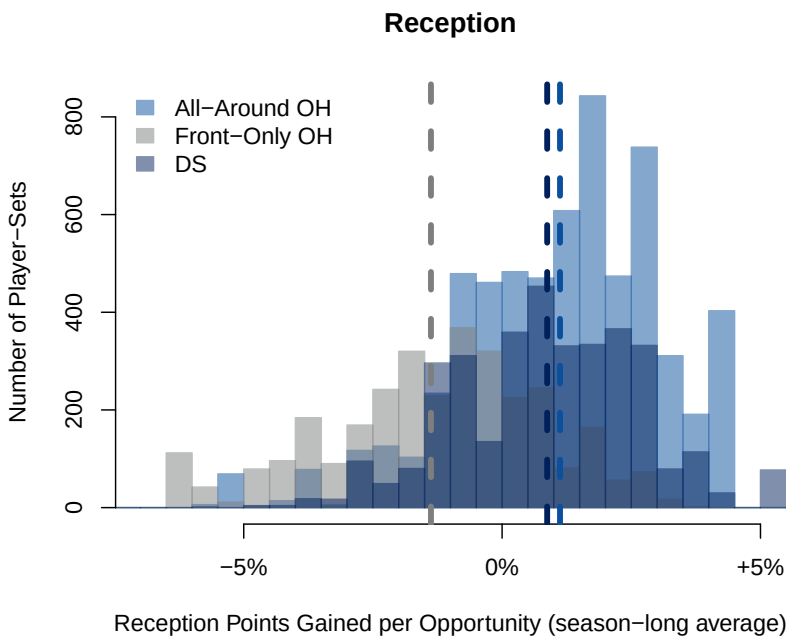
Reception PG per opportunity

All-Around OH: $+1.1\%$

Front-Only OH: -1.4%

Application: Defensive Specialist Strategy

- 63% of teams replace at least one OH with a DS in back row



Reception PG per opportunity

All-Around OH: $+1.1\%$

Front-Only OH: -1.4%

Defensive Specialist: $+0.9\%$

Substitutable Opportunities:
0.1 opportunities per point

- *Example:* Point win probability 50% $\rightarrow$ 50.2%
  - Match win probability 50% $\rightarrow$ 52%

Pythagorean formula for volleyball: $p^{10}/(p^{10}+(1-p)^{10})$

Limitations and Next Steps

- Improve blocker and digger assignments
- Correct bias for digs made by setter
- Reward good decision-making in strength of schedule
- Account for whether setter is back row or front row
- Leverage (X, Y) coordinate information

Thank You!