

# Randomness of Play Calling in American College Football

Brian Curley <sup>1,1</sup>, Gretchen Hopkirk <sup>1,2</sup>, Ryan  
Lokhorst <sup>1,3</sup> and Annette Pilkington <sup>1,4</sup>

<sup>1</sup> University of Notre Dame, Notre Dame, Indiana, U.S.A.

<sup>1</sup>[bcurley@nd.edu](mailto:bcurley@nd.edu), <sup>2</sup>[ghopkirk@nd.edu](mailto:ghopkirk@nd.edu),

<sup>3</sup>[rlokhors@nd.edu](mailto:rlokhors@nd.edu), <sup>4</sup>[Pilkington.4@nd.edu](mailto:Pilkington.4@nd.edu).

1.

# Introduction

Randomness of Play Calling in  
American College Football



# Game Theory

- › In American Football it is widely accepted that the ability to predict the next play of the offensive team gives an advantage to their opponent

# Game Theory (cont.)

- › Recent analysis of the NFL subscribes to the belief that predictability is a disadvantage for any team
- › We did not find any literature regarding the relationship of predictability to wins, losses, etc.



IS *RANDOMNESS* A KEY  
FACTOR IN WINNING?

# How we tested randomness

- › Predictability and game outcome (i.e. winning or losing) have a more complex relationship than we expected
- › We used the Wald Wolfowitz test for randomness to categorize sequences of plays



# 7,220

Division 1 College Football Games  
Played Between 2005 and 2013\*

\*Publicly available data

We categorized each play as a rushing or passing play and looked at the sequence of rushing and passing plays for each team



Rush

OR

Pass



# Random Play Calling Example



# Expectations

- › One would expect that sequences of offensive play that differ significantly from a random pattern increase the ability of the opposing team to predict plays
- › This would lead to a disadvantage for the offensive team

# How did we test this?

- > We used the Wald Wolfowitz Run Test to determine randomness

# Wald Wolfowitz Run Test

- › Given a random sequence of length  $N$ , made up of Rushes (R) and Passes (P) with  $N_R$  rushes and  $N_P$  passes
- › The number of runs ( $X$ ) (of both R's and P's) is approximately normally distributed with mean and standard deviation:

$$E(X) = \mu = \frac{2N_s N_f}{N} + 1, \quad \sigma(X) = \sqrt{\frac{(\mu - 1)(\mu - 2)}{N - 1}}.$$

# Example:

RRPPPRPRRRRPPPPRPPRRRRPPRRRRP  
PPRPRRRRPPRRRRPPRRPPRRPPRRRRRPP  
RPRPPRPPRPPRPPRPP

$$N_R = 42, N_P = 45, N = 42 + 45 = 87$$

$$X = \# \text{ of Runs} = 52$$

$$\mu = \frac{2(45)(42)}{87} + 1 = 44.44828, \quad \sigma \approx \sqrt{\frac{(43.44828)(42.44828)}{86}} \approx 4.630918$$

$$Z = (X - \mu)/\sigma = 1.630718, \quad \text{p-value} = 0.1029497.$$

# We created three categories for the sequence of plays:

- > L: z-scores too low in absolute value to reject the randomness hypothesis
- > SP: significant positive z-scores, indicating too many switches between rush and pass plays
- > SN: significant negative z-scores, indicating too few switches between play types

# Classifications of our data

Home Team(H)	Away Team(A)			Total
	ASN	AL	ASP	
HSN	N = 19	N = 384	N = 15	N = 418
HL	N = 327	N = 6073	N = 224	N = 6624
HSP	N = 11	N = 163	N = 4	N = 178
Total	N = 357	N = 6620	N = 243	N = 7220

HSN = Home team Sign't. Neg. (too few switches between play types)

ASN = Home team Sign't. Neg. (too few switches between play types)

HSP = Home team Sign't. Pos. (too many switches between play types)

ASP = Home team Sign't. Pos. (too many switches between play types)

HL = Home team not Sign't. (not enough evidence to reject randomness)

AL = Home team not Sign't. (not enough evidence to reject randomness)

# Number of Wins for Home and Away Teams in Each Category

Home Team	Away Team							
	ASN		AL		ASP		Total	
HSN	H	9	H	226	H	10	H	245
	A	10	A	158	A	5	A	173
	p(A)	0.52	p(A)	0.41	p(A)	0.33	p(A)	0.41
HL	H	193	H	3767	H	162	H	4122
	A	134	A	2306	A	62	A	2502
	p(A)	0.40	p(A)	0.38	p(A)	0.28	p(A)	0.38
HSP	H	5	H	99	H	4	H	108
	A	6	A	64	A	0	A	70
	p(A)	0.54	p(A)	0.39	p(A)	0.00	p(A)	0.39
Total	H	207	H	4092	H	176	H	4475
	A	150	A	2528	A	67	A	2745
	p(A)	0.42	p(A)	0.38	p(A)	0.27	p(A)	0.38

P(A) = percentage of wins for the away team in each category.



# What are the trends?

- › It seems to be that fewer switches between play types are related to a higher winning percentage for both home and away teams
- › It appears that predictability itself does not appear to be a disadvantage
- › However, predictability of a certain kind (too many switches between play types) may be disadvantageous

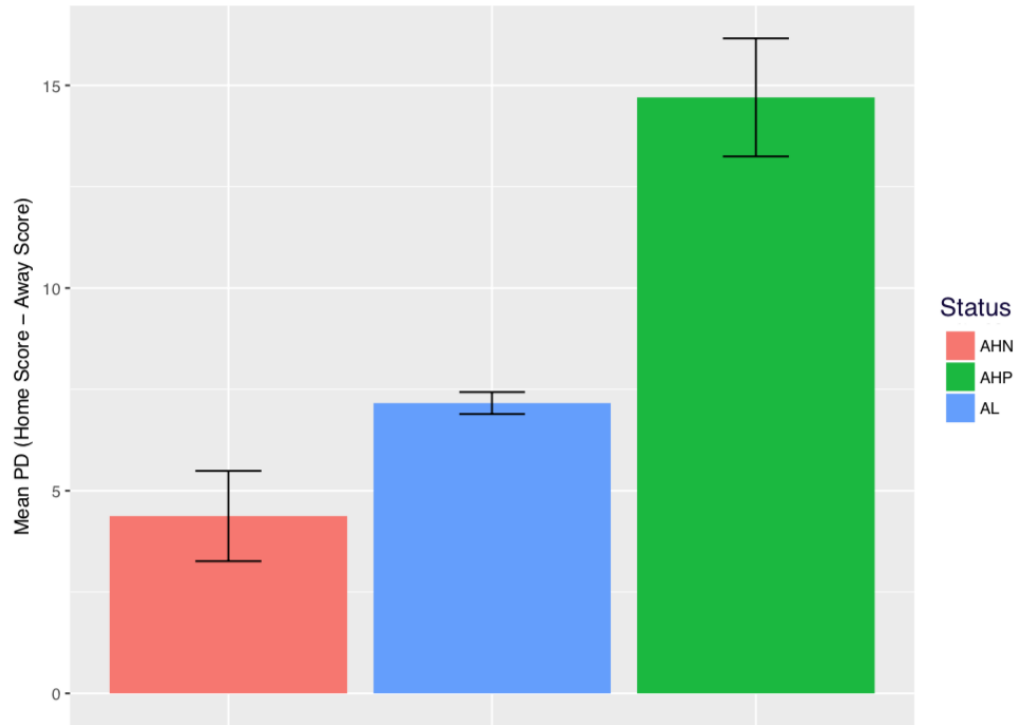
# Statistical Significance (Chi-Squared Test)

- › Marginals for away teams:
  - › The success rate of ASP teams is significantly lower than those of ASN and AL teams
- › (Conditional) Home Team Category HL:
  - › The success rate of ASP teams is significantly lower than those of ASN and AL teams

# Comparison of Average Point Differentials

- › Home team score minus away team score
  - › Shows significant differences between away teams who make relatively high amounts of switches between play types (ASP) and away teams with other play types

Means by Status of WW Z-Score for Away Team



# Other Key Game Statistics

- › These statistics showed similar trends
  - › There were statistically significant (disadvantageous) differences for the differences in rushing yards and first downs from rushing for away teams in the ASP category

# Conclusions

- › Limiting the amount of switches between play types may be helpful, especially for away teams
  - › There are benefits to staying consistent
  - › These factors may not be causal, but are still important for teams to take note of

# THANKS!

Any questions?

You can find us at:

Gretchen Hopkirk: [ghopkirk@nd.edu](mailto:ghopkirk@nd.edu)

Brian Curley: [bcurley@nd.edu](mailto:bcurley@nd.edu)