Appendix

Caleb Skinner

The appendix includes additional figures and tables. For more context, please refer to the main text. The appendix is divided by the chapter to which it applies.

Methodology

As we adjusted the point margin for the server, we assume that serve impact is additive. That is, we assume the overall impact from serving does not change as the point margin changes. In figure Figure 1, we display the server's win probability at each point margin in our data set. The dashed line is the overall win rate for the server. To reduce noise, we only include point margins where at least 1100 points have been played. The figure shows that the serve impact does decrease slightly as the point margin increases for both men and women. However, the difference is not large enough to meaningfully impact our measurement of momentum. Thus, we continue with the assumption that serve impact is additive. Moreover, it is likely that this figure is impacted by sampling bias. Fewer points are played with large point margins, and these matches tend to have lower serve win rates.

Our second assumption regarding serve impact is that the server win percentage is constant over the course of a match. In Figure 2, we visualize the server's win percentage at each point in a match. To reduce noise, we only include points with at least 500 observations. We find that the server win rate does not meaninfully change over the course of a match.

Table 1 supplements the figures in Describing Momentum. It contains the means and intervals of backward momentum in various conditions.

Table 1: Momentum Estimates for Example Sequences

Sequence	2.5%	Mean	97.5%	Sample Size
Fifteen In A Row	0.68	0.81	0.92	211.00
Ten In A Row	0.47	0.66	0.84	3,746.00
Five In A Row	0.11	0.42	0.71	91,139.00

Table 1: Momentum Estimates for Example Sequences

Sequence	2.5%	Mean	97.5%	Sample Size
Three In A Row	-0.12	0.25	0.62	378,240.00
Sixteen Out Of Twenty	0.33	0.55	0.73	14,358.00
Twelve Out Of Fifteen	0.28	0.48	0.67	40,654.00
Eight Out Of Ten	0.20	0.41	0.61	110,619.00
Four Out Of Five	-0.10	0.22	0.54	449,751.00

Results

Does Momentum Exist?

Table 2 displays the log odds for all variables in the model evaluating the relationship between backward momentum and point victor. The dependent variable is player 1's odds of winning the next point.

Table 2: Momentum Exist - All Variables

Term	Estimate	Standard Error
(Intercept)	0.3962***	0.01
Backward Momentum	0.1879***	0.01
Pre-Match Win Probability	0.0057***	0.00
Women	-0.3515***	0.01
Women & Backward Momentum	-0.0846***	0.02
Women & Pre-Match Win Probability	0.0004*	0.00
Australian Open & Men & Returner	-1.2598***	0.01
French Open & Men & Returner	-1.1938***	0.01
US Open & Men & Returner	-1.2400***	0.01
Wimbledon & Men & Returner	-1.3593***	0.01

Table 2: Momentum Exist - All Variables

Term	Estimate	Standard Error
Australian Open & Women & Returner	-0.5856***	0.01
French Open & Women & Returner	-0.5881***	0.01
US Open & Women & Returner	-0.6339***	0.01
Wimbledon & Women & Returner	-0.6857***	0.01
Australian Open & Men & Server	-0.1172***	0.01
French Open & Men & Server	-0.1718***	0.01
US Open & Men & Server	-0.1176***	0.01
Australian Open & Women & Server	-0.0789***	0.01
French Open & Women & Server	-0.1474***	0.01
US Open & Women & Server	-0.0573***	0.01

In Figure 4, we display backward momentum's effect on the point victor with varying betting odds and servers. In this situation, we describe high betting odds as a 99% implied win probability, even betting odds as a 50% implied win probability, and low betting odds as a 1% implied win probability. For all curves, the estimates are for women playing at Wimbledon.

In Figure 5, we test the assumptions of the linear model regressing **backward_momentum** on **future_momentum**. For readability, we take a random sample of 1000 points from the model. This subset appears to meet all the model's assumptions. The first plot demonstrates that the residuals appear to be normally distributed. The second plot assesses the model's linearity conditions. The third plot evaluates the homogeneity of variance. None of the covariates have any meaningful correlation.

Table 3 displays the full results for the second approach. The dependent variable is **future_momentum**.

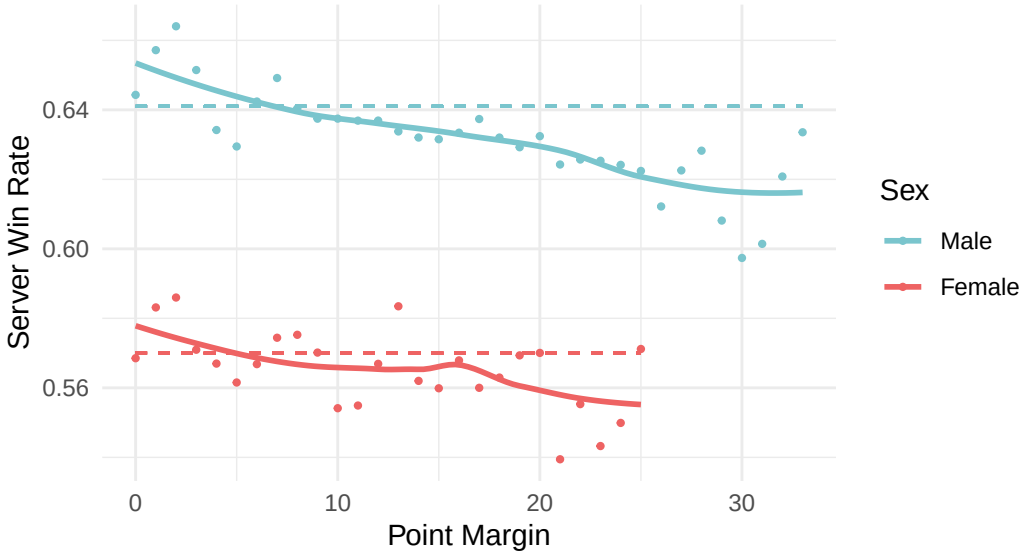

Figure 1: Serve Impact by Point Margin

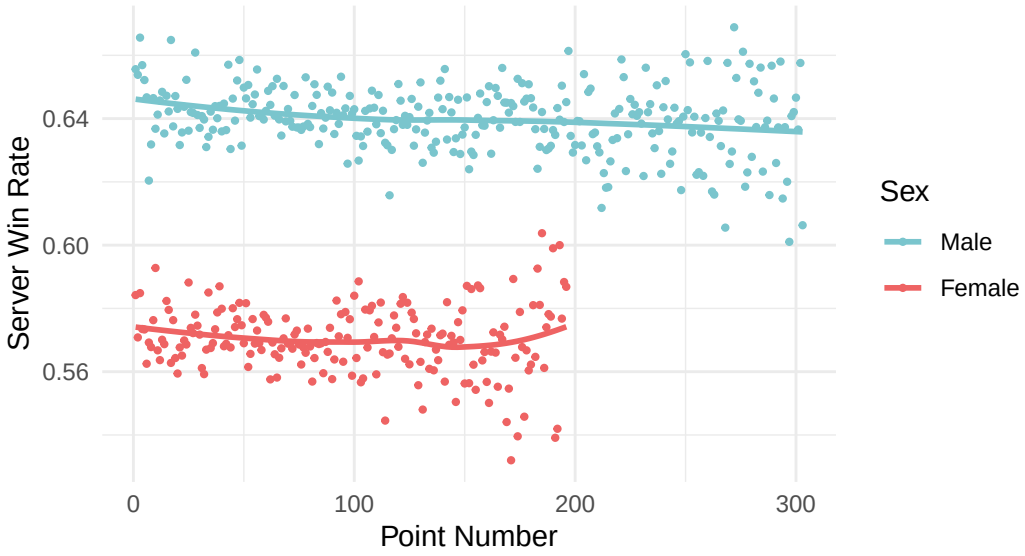

Figure 2: Serve Impact over Time

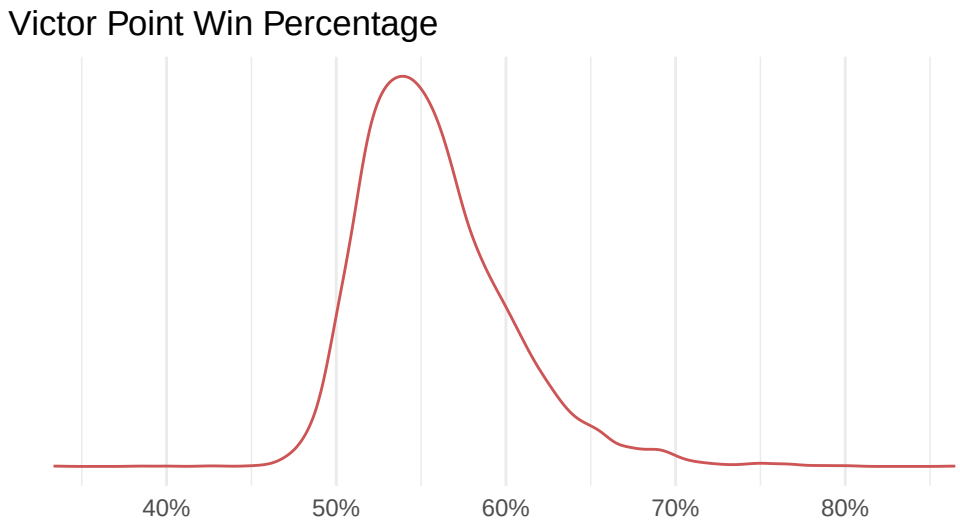

Figure 3: Victor Win Percentage

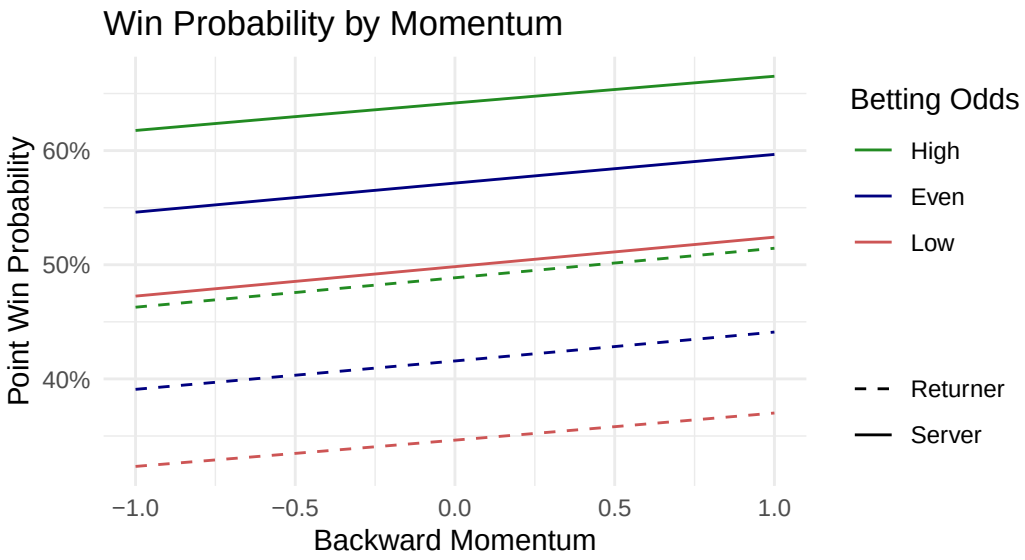

Figure 4: Backward Momentum on Win Probability with various betting odds and servers

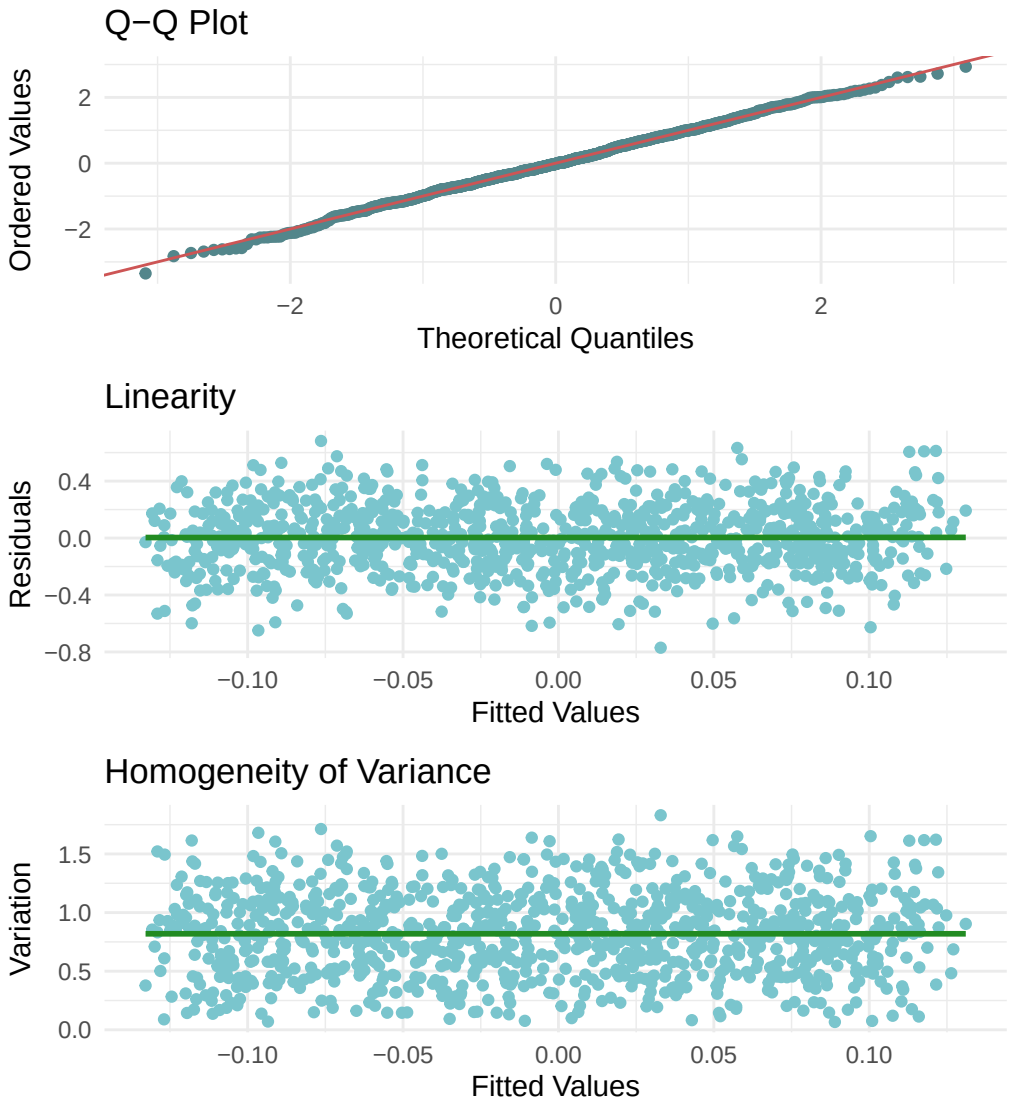

Figure 5: Assumptions for Second Approach

Table 3: Momentum Relationship Model Full Results

Term	Full	Men	Women
(Intercept)	-0.1202***	-0.1182***	-0.1239***
	(0.0009)	(0.0007)	(0.0011)
Pre-Match Win Probability	0.0024***	0.0024***	0.0025***
	(0.0000)	(0.0000)	(0.0000)
Backward Momentum	0.0169***	0.0224***	0.0081***
	(0.0004)	(0.0005)	(0.0007)
Returner & Men & Australian Open	-0.0049***	-0.0055***	
	(0.0012)	(0.0010)	
Server & Men & Australian Open	-0.0030*	-0.0029**	
	(0.0012)	(0.0010)	
Returner & Women & Australian Open	0.0053***		0.0062***
	(0.0014)		(0.0014)
Server & Women & Australian Open	0.0032*		0.0034*
	(0.0014)		(0.0014)
Returner & Men & French Open	-0.0013	-0.0019	
	(0.0012)	(0.0010)	
Server & Men & French Open	-0.0008	-0.0007	
	(0.0012)	(0.0010)	
Returner & Women & French Open	-0.0092***		-0.0086***
	(0.0014)		(0.0014)
Server & Women & French Open	-0.0082***		-0.0081***
	(0.0013)		(0.0014)
Returner & Men & US Open	-0.0005	-0.0011	
	(0.0010)	(0.0008)	
Server & Men & US Open	0.0006	0.0007	
	(0.0010)	(0.0008)	

Table 3: Momentum Relationship Model Full Results

Term	Full	Men	Women
Returner & Women & US Open	-0.0008		0.0000
	(0.0011)		(0.0011)
Server & Women & US Open	0.0000		0.0000
	(0.0011)		(0.0011)
Returner & Men & Wimbledon	-0.0015	-0.0021*	
	(0.0010)	(0.0008)	
Server & Men & Wimbledon	-0.0002		
	(0.0010)		
Returner & Women & Wimbledon	0.0003		0.0012
	(0.0012)		(0.0012)

What Factors Impact Momentum?

In Figure 6, we test the assumptions for the point-level linear regression. For readability, we take a random sample of 1000 points from the model. The first three plots appear to meet all the model's assumptions. The first plot demonstrates that the residuals appear to be normally distributed. The second plot addresses the model's linearity conditions. The third plot evaluates the homogeneity of variance. Table 4 and Figure 7 demonstrates the correlation of the covariates. **Server**, **sex**, **long_point**, **break_converted**, **break_saved**, and **bet_odds** are removed from the correlation matrix for readability. They have no large associations with any of the other covariates. The largest correlation is about .44 and between **point_victor** and **winner** and **unf_err**. The results from the assumptions tests are almost identical to the reduced model.

Table 4: Point Level Covariates Multicollinearity

Term	VIF
p1_winner	3.36
p1_double_fault	1.87
p2_double_fault	1.87
p2_unf_err	1.27

Table 4: Point Level Covariates Multicollinearity

Term	VIF
p1_unf_err	1.27
point_victor:long_point	1.16
server:sex:tournament	1.15
p1_ace	1.13
p2_ace	1.11
break_saved	1.01

In Table 5, we display the full results for the point-level linear model. The dependent variable is **future_momentum**.

Table 5: Point Level Model Full Results

	Combined		Men		Women	
Term	Full	Reduced	Full	Reduced	Full	Reduced
(Intercept)	-0.1251***	-0.1009***	-0.1271***	-0.1333***	-0.1250***	-0.0700***
	(0.0010)	(0.0042)	(0.0010)	(0.0051)	(0.0013)	(0.0064)
Pre-Match Win Probability	0.0025***	0.0021***	0.0025***	0.0025***	0.0025***	0.0015***
	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0001)
Point Victor	0.0034***	0.0033*	0.0047***	0.0045*	0.0009	0.0011
	(0.0007)	(0.0016)	(0.0009)	(0.0021)	(0.0012)	(0.0026)
P2 Break Converted	-0.0165***	-0.0144***	-0.0182***	-0.0181***	-0.0151***	-0.0116*
	(0.0015)	(0.0034)	(0.0020)	(0.0049)	(0.0021)	(0.0047)
P1 Break Converted	0.0156***	0.0151***	0.0160***	0.0195***	0.0155***	0.0118*
	(0.0015)	(0.0034)	(0.0020)	(0.0049)	(0.0021)	(0.0047)
P2 Break Saved	0.0088***	0.0119***	0.0121***	0.0182***	0.0039*	0.0041
	(0.0012)	(0.0028)	(0.0016)	(0.0038)	(0.0019)	(0.0041)
	-0.0068***	-0.0086**	-0.0092***	-0.0120**	-0.0032	-0.0040

Table 5: Point Level Model Full Results

	Combined		Men		Women	
Term	Full	Reduced	Full	Reduced	Full	Reduced
P1 Break Saved						
	(0.0012)	(0.0028)	(0.0016)	(0.0038)	(0.0019)	(0.0041)
P1 Winner	0.0002	0.0033*	-0.0004	0.0042*	0.0013	0.0021
	(0.0007)	(0.0016)	(0.0009)	(0.0021)	(0.0012)	(0.0026)
P2 Winner	-0.0007	-0.0011	0.0006	0.0015	-0.0030*	-0.0048
	(0.0007)	(0.0016)	(0.0009)	(0.0021)	(0.0012)	(0.0026)
P1 Ace	0.0068***	0.0066*	0.0077***	0.0027	0.0042	0.0156**
	(0.0012)	(0.0028)	(0.0014)	(0.0033)	(0.0025)	(0.0055)
P2 Ace	-0.0086***	-0.0091**	-0.0099***	-0.0141***	-0.0051*	0.0032
	(0.0012)	(0.0028)	(0.0014)	(0.0033)	(0.0025)	(0.0055)
P1 Double Fault	-0.0011	-0.0019	0.0030	-0.0011	-0.0068**	-0.0035
	(0.0015)	(0.0034)	(0.0019)	(0.0046)	(0.0022)	(0.0049)
P2 Double Fault	0.0006	0.0037	-0.0011	-0.0005	0.0032	0.0089
	(0.0015)	(0.0033)	(0.0019)	(0.0045)	(0.0022)	(0.0049)
P1 Unforced Error	-0.0004	0.0014	-0.0020*	-0.0015	0.0017	0.0050*
	(0.0007)	(0.0016)	(0.0009)	(0.0021)	(0.0012)	(0.0025)
P2 Unforced Error	-0.0014*	-0.0021	0.0003	-0.0004	-0.0036**	-0.0040
	(0.0007)	(0.0016)	(0.0009)	(0.0021)	(0.0011)	(0.0025)
Point Loser & Long Point	0.0002	-0.0025	0.0008	-0.0035	-0.0005	-0.0011
	(0.0009)	(0.0021)	(0.0012)	(0.0029)	(0.0014)	(0.0031)
Point Victor & Long Point	-0.0020*	-0.0035	-0.0028*	-0.0058*	-0.0008	-0.0006
	(0.0009)	(0.0021)	(0.0012)	(0.0028)	(0.0015)	(0.0031)
Returner & Men & Australian Open	-0.0032**	-0.0097***	-0.0014	0.0021		
	(0.0012)	(0.0027)	(0.0011)	(0.0025)		
Server & Men & Australian Open	-0.0043***	-0.0113***	-0.0025*	0.0011		
	(0.0012)	(0.0026)	(0.0010)	(0.0024)		

Table 5: Point Level Model Full Results

	Combined		Men		Women	
Term	Full	Reduced	Full	Reduced	Full	Reduced
Returner & Women & Australian Open	0.0056***	-0.0126***			0.0057***	-0.0120***
	(0.0014)	(0.0030)			(0.0014)	(0.0030)
Server & Women & Australian Open	0.0038**	-0.0148***			0.0039**	-0.0151***
	(0.0014)	(0.0030)			(0.0014)	(0.0030)
Returner & Men & French Open	0.0000	-0.0168***	0.0017	-0.0053*		
	(0.0012)	(0.0028)	(0.0011)	(0.0026)		
Server & Men & French Open	-0.0016	-0.0123***	0.0003	-0.0001		
	(0.0012)	(0.0027)	(0.0010)	(0.0025)		
Returner & Women & French Open	-0.0096***	-0.0099***			-0.0095***	-0.0091**
	(0.0014)	(0.0030)			(0.0014)	(0.0031)
Server & Women & French Open	-0.0077***	-0.0125***			-0.0078***	-0.0126***
	(0.0013)	(0.0029)			(0.0014)	(0.0030)
Returner & Men & US Open	0.0013	-0.0039	0.0030***	0.0079***		
	(0.0011)	(0.0023)	(0.0009)	(0.0021)		
Server & Men & US Open	-0.0005	-0.0072**	0.0013	0.0054**		
	(0.0010)	(0.0023)	(0.0008)	(0.0020)		
Returner & Women & US Open	-0.0004	-0.0061*			-0.0004	-0.0058*
	(0.0012)	(0.0025)			(0.0012)	(0.0026)
Server & Women & US Open	0.0003	-0.0045			0.0003	-0.0051*
	(0.0011)	(0.0025)			(0.0011)	(0.0025)
Returner & Men & Wimbledon	0.0007	-0.0077***	0.0025**	0.0044*		
	(0.0010)	(0.0023)	(0.0009)	(0.0021)		
Server & Men & Wimbledon	-0.0018	-0.0126***				
	(0.0010)	(0.0022)				
	0.0009	0.0039			0.0009	0.0046

Table 5: Point Level Model Full Results

	Combined		Men		Women	
Term	Full	Reduced	Full	Reduced	Full	Reduced
Returner & Women & Wimbledon	(0.0012)	(0.0025)			(0.0012)	(0.0026)

In Figure 8, we test the assumptions for the game-level linear regression. For readability, we take a random sample of 1000 points from the model. This subset appears to meet all the model's assumptions. The first plot demonstrates that the residuals appear to be normally distributed. The second plot addresses the model's linearity conditions. The third plot evaluates the homogeneity of variance. Table 6 and Figure 9 displays the correlation of the covariates. **Sex** and **bet_odds** are removed from the correlation matrix for readability. They have no large associations with any of the other covariates. The largest correlation is .805 between **interruption** and **change_ends**. There is also a large correlation between **server** and **game_victor** (.496). The results from the assumptions tests are almost identical to the reduced model.

Table 6: Game Level Covariates Multicollinearity

Term	VIF
game_victor	11.06
server	4.14
server:sex:tournament	1.80
game_victor:change_ends	1.42
set_victor	1.04

In Table 7, we show the full results for the six game level models.

Table 7: Game Level Model Full Results

	Combined		Men		Women	
Term	Full	Reduced	Full	Reduced	Full	Reduced
(Intercept)	-0.1322***	-0.1148***	-0.1351***	-0.1434***	-0.1290***	-0.0920***
	(0.0025)	(0.0106)	(0.0026)	(0.0133)	(0.0036)	(0.0168)

Table 7: Game Level Model Full Results

	Combined		Men		Women	
Term	Full	Reduced	Full	Reduced	Full	Reduced
Server	-0.0091**	-0.0062	-0.0141***	-0.0139	-0.0062	-0.0076
	(0.0033)	(0.0070)	(0.0035)	(0.0083)	(0.0044)	(0.0095)
Game Victor	0.0238***	0.0200***	0.0259***	0.0227***	0.0204***	0.0183**
	(0.0020)	(0.0046)	(0.0026)	(0.0062)	(0.0031)	(0.0069)
Pre-Match Win Probability	0.0025***	0.0023***	0.0025***	0.0027***	0.0025***	0.0017***
	(0.0000)	(0.0002)	(0.0000)	(0.0002)	(0.0000)	(0.0003)
P1 Tiebreak Victor	-0.0186**	0.0092	-0.0131	0.0207	-0.0330*	-0.0274
	(0.0069)	(0.0155)	(0.0079)	(0.0183)	(0.0148)	(0.0293)
P2 Tiebreak Victor	0.0181**	0.0198	0.0147	0.0150	0.0244	0.0286
	(0.0069)	(0.0150)	(0.0080)	(0.0177)	(0.0141)	(0.0294)
P2 Set Victor	-0.0397***	-0.0291***	-0.0350***	-0.0199*	-0.0490***	-0.0433***
	(0.0030)	(0.0071)	(0.0037)	(0.0091)	(0.0052)	(0.0116)
P1 Set Victor	0.0320***	0.0249***	0.0287***	0.0253**	0.0386***	0.0234*
	(0.0030)	(0.0072)	(0.0037)	(0.0091)	(0.0053)	(0.0117)
Game Victor & Server	0.0012	-0.0008	0.0013	-0.0050	0.0012	0.0034
	(0.0023)	(0.0053)	(0.0030)	(0.0074)	(0.0035)	(0.0077)
Game Loser & Interruption	0.0051*	0.0087	0.0005	-0.0001	0.0139***	0.0222*
	(0.0024)	(0.0055)	(0.0029)	(0.0071)	(0.0041)	(0.0089)
Game Victor & Interruption	-0.0051*	-0.0090	-0.0050	-0.0114	-0.0052	-0.0047
	(0.0024)	(0.0055)	(0.0030)	(0.0070)	(0.0041)	(0.0090)
Game Loser & Change Ends	0.0008	-0.0062	0.0055	-0.0026	-0.0079*	-0.0121
	(0.0024)	(0.0055)	(0.0029)	(0.0070)	(0.0040)	(0.0088)
Game Victor & Change Ends	-0.0004	0.0081	-0.0015	0.0123	0.0015	0.0012
	(0.0024)	(0.0055)	(0.0029)	(0.0069)	(0.0041)	(0.0089)
	-0.0030	-0.0084				

Table 7: Game Level Model Full Results

	Combined		Men		Women	
Term	Full	Reduced	Full	Reduced	Full	Reduced
Returner & Men & Australian Open	-0.0030	-0.0068				
	(0.0030)	(0.0068)				
Server & Men & Australian Open	-0.0075*	-0.0162*				
	(0.0030)	(0.0068)				
Returner & Women & Australian Open	0.0038	-0.0101				
	(0.0035)	(0.0076)				
Server & Women & Australian Open	0.0064	-0.0106				
	(0.0035)	(0.0076)				
Returner & Men & French Open	-0.0018	-0.0184**	0.0012	-0.0101		
	(0.0030)	(0.0069)	(0.0030)	(0.0073)		
Server & Men & French Open	-0.0046	-0.0203**	0.0030	-0.0042		
	(0.0030)	(0.0070)	(0.0030)	(0.0073)		
Returner & Women & French Open	-0.0089*	-0.0122			-0.0126**	-0.0016
	(0.0035)	(0.0077)			(0.0040)	(0.0088)
Server & Women & French Open	-0.0064	-0.0110			-0.0128**	-0.0003
	(0.0035)	(0.0077)			(0.0039)	(0.0088)
Returner & Men & US Open	0.0017	-0.0049	0.0047	0.0034		
	(0.0026)	(0.0058)	(0.0026)	(0.0063)		
Server & Men & US Open	-0.0036	-0.0116*	0.0040	0.0046		
	(0.0026)	(0.0058)	(0.0026)	(0.0062)		
Returner & Women & US Open	-0.0021	-0.0065			-0.0059	0.0036
	(0.0029)	(0.0064)			(0.0035)	(0.0076)
Server & Women & US Open	-0.0008	-0.0086			-0.0071*	0.0020
	(0.0029)	(0.0064)			(0.0035)	(0.0076)
Returner & Men & Wimbledon	0.0018	-0.0101	0.0049	-0.0012		
	(0.0026)	(0.0056)	(0.0026)	(0.0060)		

Table 7: Game Level Model Full Results

	Combined		Men		Women	
Term	Full	Reduced	Full	Reduced	Full	Reduced
Server & Men & Wimbledon	-0.0051*	-0.0188***	0.0023	-0.0024		
	(0.0026)	(0.0056)	(0.0026)	(0.0060)		

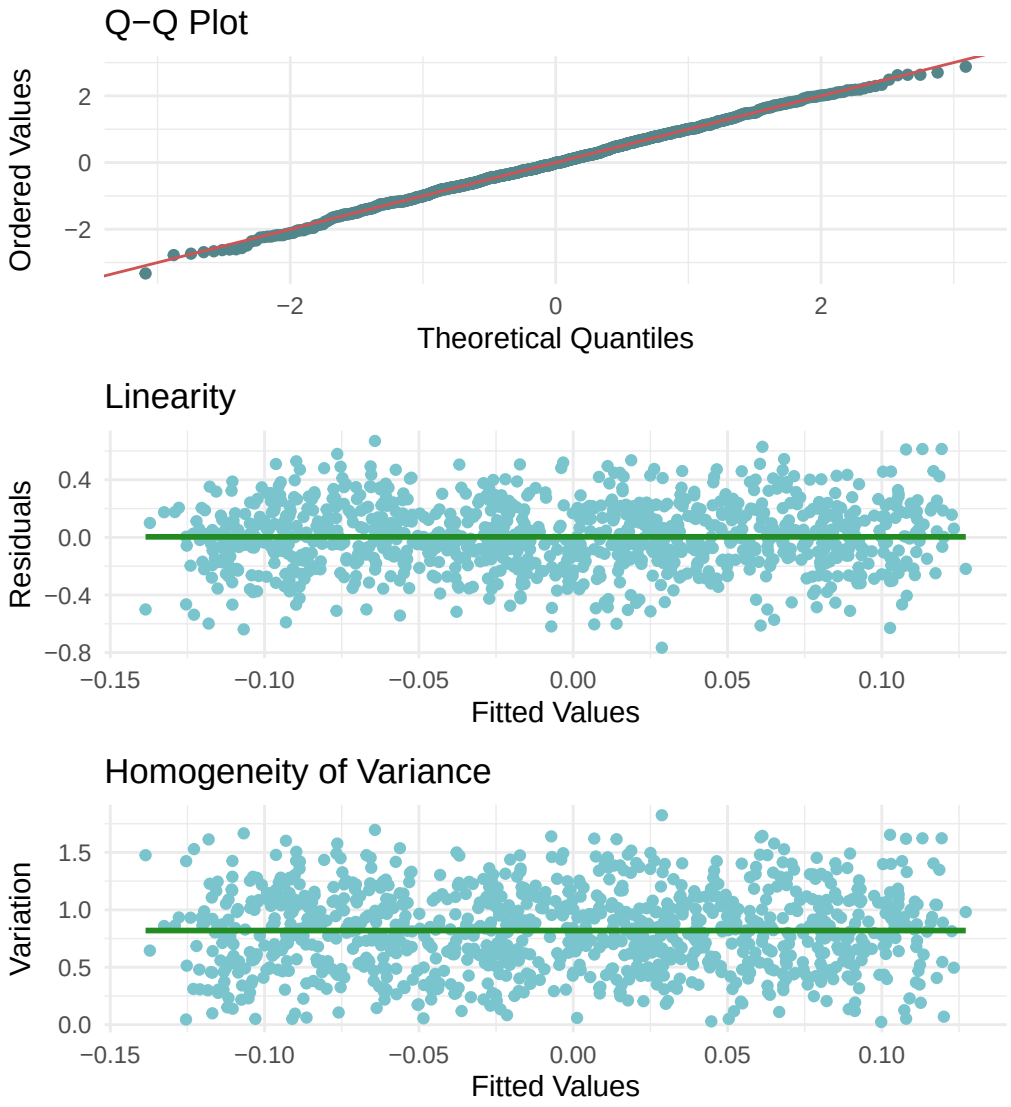

Figure 6: Assumptions for Point Level

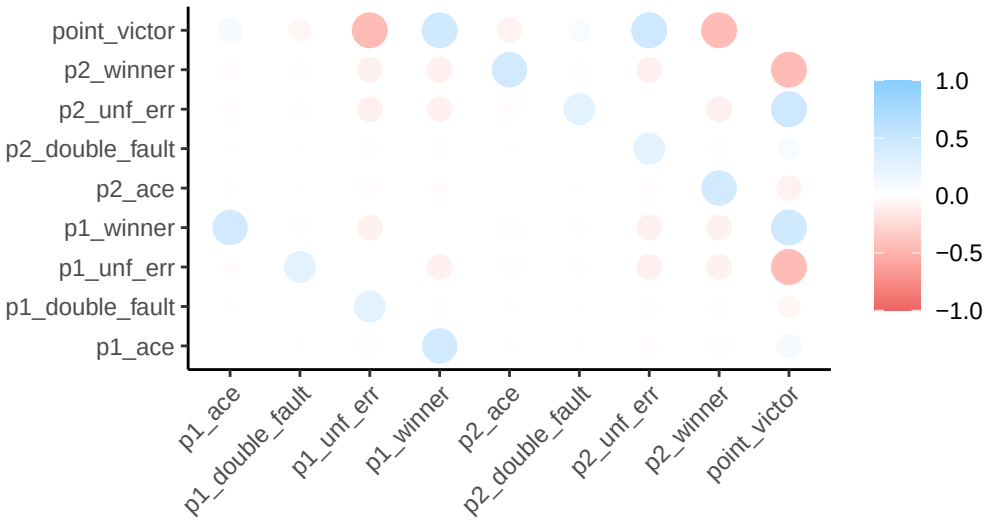

Figure 7: Point Level Covariates Correlation Matrix

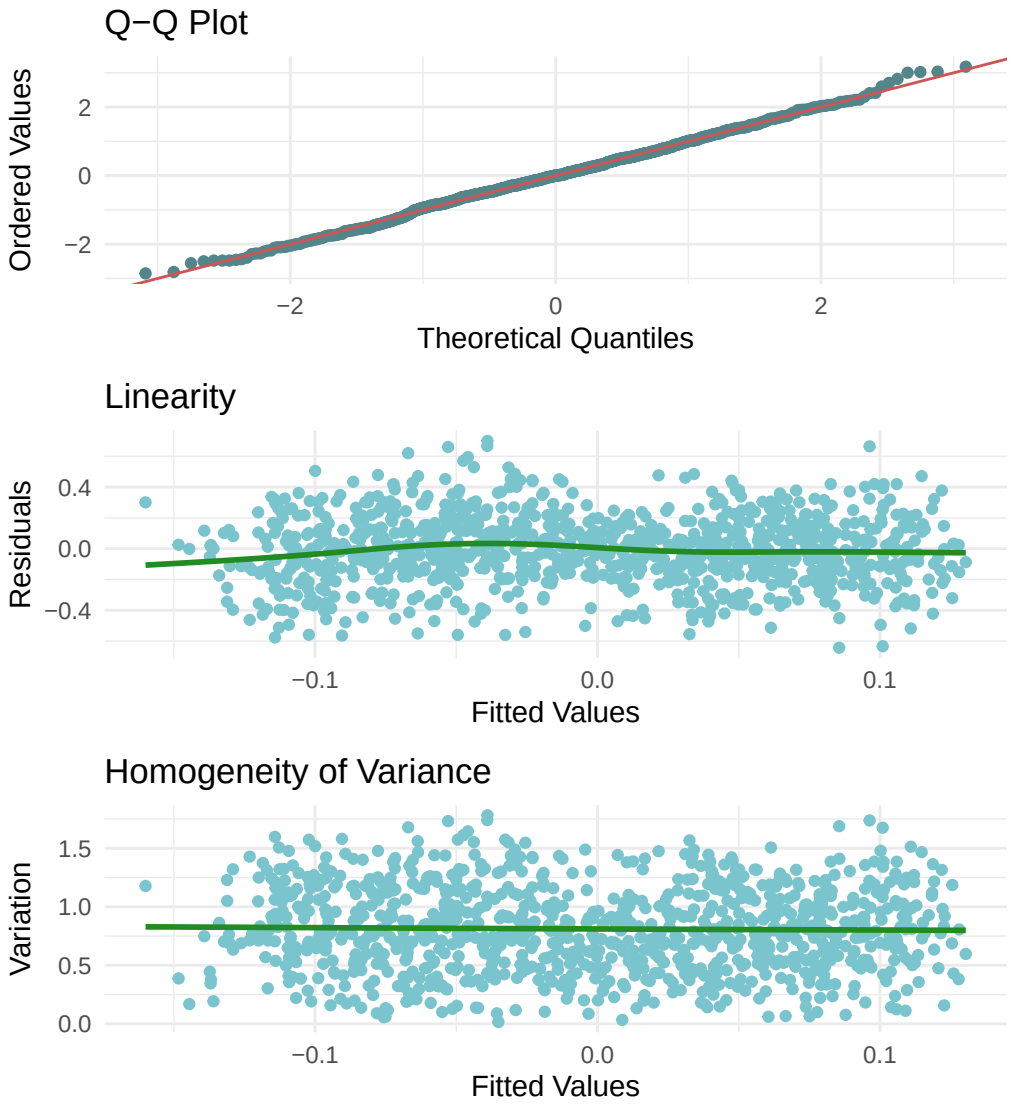

Figure 8: Assumptions for Game Level

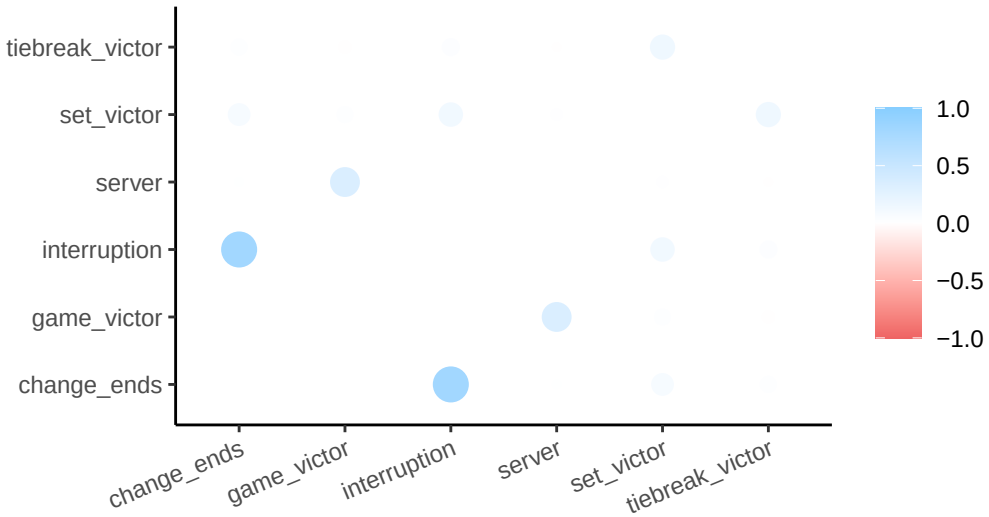

Figure 9: Game Level Covariates Correlation Matrix