

THE IMPORTANCE OF THE SERVE IN MEN'S TENNIS: A STATISTICAL ANALYSIS

*Michael R. Summers, Business Administration Division, Pepperdine University,
Malibu, CA 90263, 310-506-4536, msummers@pepperdine.edu*

ABSTRACT

Since serving and returning serve constitute a major part of a tennis match, this study seeks to identify several performance measures of each that were significantly related to a player's overall ranking and to develop a multiple regression model that explains the player's ranking. Also, the performance measures that were found to be most significant were compared for players who performed well at the four grand slam tournaments to examine whether different serving and returning characteristics were most important at each tournament.

BACKGROUND

In men's tennis a powerful serve can outweigh deficiencies in other parts of a player's game and allow him to reach the upper ranks of players. Largely because of improvements in racquet technology over the years, many of the top players are capable of serving a ball from 130 to 140 MPH. On the other side of the coin, those players who are capable of returning a high percentage of these serves also have a significant advantage. It's very helpful, when serving and facing a break point, to be able to hit a great serve and save the break or, when receiving and holding a break point, to be able to hit a great return and get the break. Two of the most successful players in the recent past were Pete Sampras, who some say had the best second serve of all time, and Andre Agassi, who didn't have a powerful serve but was one of the best service returners ever.

Among the many statistical studies of men's tennis, a few have concentrated on the probability of winning points when serving or when receiving. For example, a 2001 study by O'Donoghue [4] focused on a player's probability of winning the next point and of winning the game from every possible game score. From this it was determined which point in a game was most important. Barnett and Clarke [3] used serving and returning statistics of top players to predict the probability of two particular opponents in a match winning a point, a game, or a match. These numbers could be updated during a match in order to facilitate betting on the match.

DATA

Men's Association of Tennis Professionals rankings as of September 26, 2011, were obtained from the ATP web site [1]. Another page at the ATP web site showed the top 60 players in six serving performance categories: aces, 1st serve percentage, 1st serve points won, 2nd serve points won, service games won, and break points saved [2]. This page also showed the top 60 performers in four return of serve performance categories: points won returning 1st serve, points won returning 2nd serve, break points converted, and return games won. With a few exceptions, the top 50 ranked players were among the top 60 performers in each category.

RESULTS

Top Performers

The serving and returning performance statistics of the top 50 players were compiled (data available from author). The most effective server overall was Roger Federer, who won 78% of his first serves and 57% of his second serves, both tops among all players. The top-ranked player, Novak Djokovic, also had very high percentages, 74% and 56%, respectively. Federer's service performance also resulted in a high percentage of service games won, 89%, second only to John Isner's 90%. The other top players, Rafael Nadal and Andy Murray, had percentages that were less impressive, but they were among the leaders in return of serve. Murray tied for the highest percentage of points won returning the first serve, 37%, and Nadal tied for the highest percentage of points won returning the second serve, 58%. While Federer's return statistics were not as impressive, Djokovic's were the highest overall, 37% and 58%, respectively, as well as having the highest percentage of break points won, 48%, and by far the highest percentage of return games won, 41%. Because Djokovic had top statistics in both serving and receiving, it is unsurprising that he comfortably led the overall rankings.

Some of the players ranked lower performed well either in serving or in returning, but not both. The most extreme example was John Isner, who led overall in service games won, 90%, but was last overall in return games won, just 13%. At the other end of the spectrum was Fabio Fognini, who won just 69% of his service games, the lowest among the top 50 players, but had a very respectable winning percentage of 31% when receiving, sixth overall. However, Isner had an overall ranking of 18, while Fognini ranked 39th, perhaps indicating that having a good serve might be more helpful than having a good return!

Correlations With Players' Rankings and Between Factors

In order to see which serving performance measures and which returning performance measures were most closely correlated to a player's overall ranking points, a correlation matrix was created (Table 1). As expected, all of the serving and returning performance measures were positively correlated with a player's ranking points. We would, of course, expect the percentage of service games won and the percentage of return games won to be highly correlated with a player's ranking, since these factors fairly directly affect the outcome of matches. The percentage of service games won had a correlation of .3735, significant at the .01 level, while the percentage of return games won had the highest correlation overall, .6264, significant at the .001 level. Among the six serving performance measures, however, the factor most correlated with the ranking points was the percentage of second serves won, with a correlation of .4370.

All of the returning performance measures were correlated very significantly with ranking points, and three of the four had a p-value less than .001. None of the six serving measures had a p-value below .001. These high correlations with the return of serve might suggest that returning serve is actually more important than serving, in contrast to the Isner-Fognini example above.

The never-ending question of whether it is better to go for a big first serve or to get a high percentage of first serves in can be examined here. However, neither the number of aces nor the percentage of first serves in were significantly correlated to ranking points, both with a correlation of about .14. While it might be that expressing the number of aces as a percentage of a player's service games would be a better measure of having a big serve, using the total number of aces for the year (as the ATP does here)

might have a better chance of being correlated with a player's ranking, since successful players play more matches per tournament, giving them more chances for aces. In any case, the number of aces had the lowest correlation of the ten measures. As expected, the number of aces was negatively correlated with the percentage of first serves in, but not significantly. Interestingly, though, the percentage of first serves in was negatively correlated with the percentage of first serves won ($p < .05$) and had virtually no correlation with service games won, showing that serving carefully was not very effective.

TABLE 1
Correlation Matrix

	Points	Aces	1st Serve %	% 1st Serv. Won	% 2nd Serv. Won	% Service Gms. Won
Points	1.0000					
Aces	0.1410	1.0000				
First Serve %	0.1452	-0.1878	1.0000			
% First Serves Won	0.2569	0.8702	-0.3446	1.0000		
% Second Serves Won	0.4370	0.4822	0.1610	0.5630	1.0000	
% Service Games Won	0.3735	0.7859	0.0652	0.8809	0.8039	1.0000
% Break Points Saved	0.3256	0.5285	0.3497	0.5380	0.5514	0.7288
% Won Ret. First Serve	0.5076	-0.4152	-0.0466	-0.3455	-0.1379	-0.3519
% Won Ret. Second Serve	0.5532	-0.4100	0.0063	-0.2256	0.1206	-0.1510
% Break Points Won	0.3828	-0.1513	-0.2191	-0.0247	-0.0407	-0.1179
% Return Games Won	0.6264	-0.3987	-0.0076	-0.2828	-0.0063	-0.2450

	% Break Pts. Saved	% Won Ret. 1st Serve	% Won Ret. 2nd Serve	% Break Pts. Won	% Return Gms. Won
Points					
Aces					
First Serve %					
% First Serves Won					
% Second Serves Won					
% Service Games Won					
% Break Points Saved	1.0000				
% Won Ret. First Serve	-0.2173	1.0000			
% Won Ret. Second Serve	-0.0507	0.7110	1.0000		
% Break Points Won	0.0039	0.5982	0.6536	1.0000	
% Return Games Won	-0.0903	0.9043	0.8863	0.7294	1.0000

Correlation

0 -- .29
.29 -- .37
.37 -- .45
.45 -- 1.00

Significance

Not Significant
 $p < .05$
 $p < .01$
 $p < .001$

It is interesting to note that the common perception that most players excel in either serving or receiving, but not both, gets some support from our data. There are many negative correlations between the serving performance measures and the returning performance measures. Some of these are quite significant ($p < .01$), such as aces correlated with the percentage won returning first serves, the percentage won returning second serve, and the overall percentage of return games won. The most direct comparison, between the percentage of service games won and the percentage of return games won, had a correlation of $-.2450$, not quite significant at the $.05$ level.

Multiple Regression Model

Excluding the factors of service games won and return games won, the serving performance measure with the highest correlation to points was the percentage of second serves won, and the return performance measure with the highest performance measure was the percentage won returning second serves. Since these two factors had very little correlation to each other, a multiple regression model using these two independent variables to predict rating points was quite good. Adding the next most highly-correlated variable, the percentage won returning first serves, did not result in an improvement, as the latter variable was not significant. However, a model using just the percentage of second serves won and the percentage won returning first serves produced the best results (Table 2). The overall model and both independent variables are all highly significant.

TABLE 2
Best Multiple Regression Model

<i>Regression Statistics</i>	
Multiple R	0.72095312
R Square	0.5197734
Adjusted R Square	0.49794491
Standard Error	1888.27926
Observations	47

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	2	169805970.2	84902985	23.811706	9.81334E-08
Residual	44	156886337.1	3565598.6		
Total	46	326692307.3			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	-39519.775	6370.330421	-6.203725	1.693E-07
% 2nd Serves Won	496.61914	101.346928	4.9001894	1.34E-05
% Won Ret. 1st Serve	523.399481	95.36495045	5.4883841	1.894E-06

Performance Differences at the Grand Slams

Finally, we examined whether different serving and returning performance measures had different levels of importance at the four grand slam tournaments. Since the Australian Open and the U.S. Open are played on hard courts, the French Open on clay, and Wimbledon on grass, it might be that different factors could affect success at the different tournaments. A previous study by Summers [5] compared a variety of performance statistics between the French Open and Wimbledon, showing some significant differences.

Here we looked at players who were successful at the grand slam tournaments, defined as those who reached the round of 16, to see whether there were significant difference in several performance measures. Sample sizes were slightly below 16 in most cases, since there were a few players who made the round of sixteen but who were not among the top 60 in the performance measures. An analysis of variance was performed between the average performance measures of the top players at the four grand slam tournaments. As shown in Table 3, successful players at the French Open did have fewer aces for the year than did the successful players at the other three tournaments, as expected, but the differences were not significant. The successful players at the grand slams differed little on the percentage of second serves won and the percentage won returning first serves, the factors used in our best overall model.

CONCLUSIONS

Among the performance measures of serving, the percentage of second serves won had the highest correlation with a player's ranking. Of the performance measures of returning serve, both the percentage won returning the first serve and the percentage won returning the second serve were highly significant. Overall, the performance measures of returning serve had higher correlations with ranking points than did the measures of serving. There were some significant negative correlations between serving measures and returning measures, suggesting that most players (except the very best) were better at one than the other. We found no significant differences in serving or returning measures among successful players at the four grand slam tournaments.

REFERENCES

- [1] Association of Tennis Professionals. <http://www.atpworldtour.com/Rankings/Singles.aspx>.
- [2] Association of Tennis Professionals. <http://www.atpworldtour.com/Rankings/MatchFacts.aspx>.
- [3] Barnett, Tristan, and Stephen R. Clarke. Combining Player Statistics to Predict Outcomes of Tennis Matches. *IMA Journal of Management Mathematics* (2005) 16.
- [4] O'Donoghue, Peter G. The Most Important Points in Grand Slam Singles Tennis. *Research Quarterly for Exercise and Sport*, June 2001, Vol. 72, Issue 2.
- [5] Summers, Michael R. Clay vs. Grass: A Statistical Comparison of the French Open and Wimbledon. *American Journal of Economics and Business Administration*, 3 (2), 2011.

TABLE 3
Means of Top Players at Grand Slams

Aces

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
Australian	16	5460	341.25	15898.87
French	14	3955	282.5	10195.19
Wimbledon	14	4943	353.0714	19832.84
U.S.	14	5266	376.1429	28692.13

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	66833.96	3	22277.99	1.200796	0.318282	2.775762
Within Groups	1001845	54	18552.69			
Total	1068679	57				

% Second Serves Won

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
Australian	15	802	53.46667	6.409524
French	14	730	52.14286	11.67033
Wimbledon	14	747	53.35714	6.247253
U.S.	13	701	53.92308	7.076923

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	23.84359	3	7.947863	1.013994	0.394037	2.7826
Within Groups	407.585	52	7.838173			
Total	431.4286	55				

% Won Returning First Serves

<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>
Australian	15	479	31.93333	8.638095
French	14	467	33.35714	6.554945
Wimbledon	14	445	31.78571	11.1044
U.S.	13	418	32.15385	13.80769

ANOVA

<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
Between Groups	21.64222	3	7.214072	0.726722	0.540666	2.7826
Within Groups	516.1971	52	9.926867			
Total	537.8393	55				