

Opening up the court (surface) in tennis grand slams

September 15, 2018

Abstract

The Australian Open, French Open, Wimbledon, and US Open are the four most prestigious events (grand slams) in tennis and are played on four different surfaces: Plexicushion (hard court), clay, grass, and DecoTurf (hard court), respectively. It is often thought that some players achieve better results at some grand slams than others and that difference in results may be attributed, at least partially, to the difference in the surfaces. For example three top players of all time, Rafael Nadal, Roger Federer, and Serena Williams have achieved their best results on clay, grass, and hard courts, respectively. In this paper, we examine the effect of surface type at grand slams from 2013-2017 for top players, adjusting for confounders such as league, competitor strength, and player attributes. We answer the question of court effect on player results at first a high level using a “complete” data with few features and end using partial data with a variety of tennis-specific features. Our primary model is a hierarchical model with fixed and random effects accounting for both player and tournament features. We show that differences in surface type become more apparent with more detailed data and that more study should be conducted in this area.

1 Introduction

The four tennis grand slams (Australian Open (AO), French Open (FO), Wimbledon (Wim.), US Open (USO)) are played in different cities (Melbourne, Paris, London, New York City) at different times of year (January, May, June, August) over the course of two weeks with 7 total rounds and 128 players for both men (ATP) and women (WTA). All four slams are played on different surfaces (Plexicushion hard court, clay, grass, DecoTurf hard court). It is often thought that top players achieve better results on some surfaces than others. For example, there is the idea of “home court advantage” in which players from the country of the grand slam tournament are thought to perform better than at other grand slams. Recently, Spanish players, led by Nadal, seem to have dominated the clay court scene (Lewit 2018) making it seem like players of a certain country may perform better on some surfaces than others.

Rafael Nadal is known as the “King of Clay” in tennis, having won 11 out of his current 17 grand slams titles at the French Open, which is played on a clay surface (Jurejko 2018). In contrast, his rival Roger Federer has won the greatest proportion of his grand slam titles (8 out of 20) at Wimbledon, which is played on grass. On the women’s side, Serena Williams, winner of 24 grand slam titles, has been dominant both on hard court (7 titles at the Australian Open and 6 at US Open) and grass (7 at Wimbledon).

In this paper, we use the results of grand slam players from 2013-2017 to analyze the effect of court surface on player performance at the four grand slams. We examine court surface across all participants in grand slams, and also investigate player effects. The goals of this work are to:

1. Determine if results (in terms of wins/losses and point distributions) differ across the four surfaces
2. Examine fine grained attributes such as winners, aces, and unforced errors across the four surfaces
3. Assess how players perform across the four surfaces, with focus on Nadal, Federer, and Williams.

There has been little work on quantifying the high-level effect of court surface in tennis (issue 1). Existing literature provides methods for forecasting the outcome of tennis matches (Klaassen and Magnus 2003; Newton and Keller 2005; McHale and Morton 2011; Kovalchik 2016)] or for assessing whether points within a match are independent and identically distributed (Klaassen and Magnus 2001). Knottenbelt, Spanias, and Madurska (2012) take court surface into account in their model, but do not compare the results of one surface to another, and Barnett and Pollard (2007) conclude that players have more injuries on grass compared to other courts. Results from studies performed in other sports show that surface type does have an effect on the game, either directly or indirectly (Andersson, Ekblom, and Krustup 2008; Gains et al. 2010).

We examine the match attributes more unique to tennis to address issue (2). These attributes include winners (when a player hits a shot that the opponent is unable to return), aces (a serve an opponent is unable to return), and unforced errors (UE) (when a player makes a simple mistake resulting in a loss of the point). In general, players hope to maximize the winner to UE ratio (W/UE), which may lead to playing more aggressively on faster courts and more conservatively on slower courts such as clay (Paxinos 2007).

Finally, to assess how individual players differ in performance across surfaces, issue (3), we take two approaches. The first is to use a hierarchical model with fixed and random effects, taking into account player and tournament attributes using the complete data set. This allows us to leverage tournament information to model players with few matches played and to help account for noise. Our second approach is to model performance for specific players individually. Since Nadal, Federer, and Williams are extremely popular and have had much success in the 2013-2017 time range, we have enough data points with tennis-specific feature attributes to examine the effects individually. In both cases, we examine whether our models pass “common sense” tests like how the models in Thomas et al. (2013) show that commonly well known hockey players also have high status in the model. We also examine whether these players do have court surface effects. For the individual models, we do not split the data into training and test sets due to the relatively low number of observations for the individuals. However, when the data becomes available, we would like to test these models on the 2018 grand slam results. For model selection, we use AIC to compare models to one another (Wasserman 2004).

Few academic papers have been written about Nadal, Federer, or Williams. One paper studies Federer’s odds of winning when Nadal suddenly withdrew from Wimbledon and shows that Federer was too heavily favored by bookmakers (Leitner, Zeileis, and Hornik 2009). One analysis of Williams shows how she has gotten better with age, even past the point when other greats began to decline, but the study does not look at surface type (Morris 2015).

Readers may object that we are looking at differences between grand slams, which each have their own time period, weather conditions, play time conditions, and “home court effects” instead of differences in surfaces alone. However, (1) grand slam data is the most readily available and most complete which makes it the best choice at the moment for modelling, (2) we adjust for these confounders where we can, and (3) analyzing the difference in the grand slams is still useful as they are considered to be the most prestigious events in tennis.

The rest of this paper is organized as follows. In Section Data we describe our grand slam tennis data. In Section Early Data Analysis we examine the data at a high level and use clustering methods to determine how the courts differ from one another. In Section Methods we describe our models with (1) complete data with few features and (2) partial data with more features. In Section Results we describe the results of our modelling and also examine the play of Nadal, Federer, and Williams. Finally in Section Discussion, we summarize our results and discuss future work.

2 Data and EDA

2.1 Data

The primary data consists of 5080 matches split evenly over the four grand slams and the two leagues: ATP (men’s) and WTA (women’s). There are 5080 observations in this data set, which corresponds to a complete data set for the period of 2013-2017 for four grand slams, two leagues, and seven rounds. Each match has 80 attributes, many of which are redundant. We focus on the following attributes for both the winner and loser of the match: games won, points won, retirement, break points faced, break points saved, aces, country of origin, and player attributes. Additionally, we take into account the number of sets in a match, the court surface, and round of the tournament. A subset of the data is shown in Table 1.

The secondary data is partial point by point data for grand slam matches. In this data set, each row is a point in a match with details on who won the point, serve speed, and whether a player had a forced or unforced error, winner, ace, or net point win. There is also tournament info such as court surface, year, and

Table 1: Example of the complete grand slam data. It includes winner and loser attributes, match attributes, and tournament attributes. Not all attributes are shown here.

Winner	Tournament	Year	W. IOC	W. Points	W. Rank	L. Points	L. Rank
Serena Williams	Australian Open	2013	USA	52	3	18	110
Roger Federer	Australian Open	2013	SUI	95	2	63	46
Rafael Nadal	Roland Garros	2013	ESP	140	4	115	59

Table 2: # of matches played for Nadal, Federer, and Williams from 2013-2017 at each of the grand slams.

Tournament	Nadal	Federer	Williams
Australian Open	20	28	30
French Open	29	14	23
US Open	21	22	26
Wimbledon	11	29	21
Total	81	93	100

time start. Additional attributes include rally length, winner and final score of the match, retirement, and minutes played. However, these additional attributes are only available for about 1/8 of the data.

We aggregate the partial point by point data into partial match data, summing the total number of errors, winners, aces, net points, and taking the average service speed for each player in the match. We then join the partial match data with the complete data to find the score and outcome of the match. We do this by joining the features of tournament, year, player one, and player two. This final data set consists of 3066 observations (compared to the 5080 in the original grand slams). The median rank of the winners for the original data is 28, and the median rank for the winners of the partial data is 21. The same trend is true for the loser rank. This means that it is more likely for better players to have court by court data, which may effects our analysis for models when using this more detailed yet partial data. Additionally, we found that no women’s point by point data was recorded for 2015.

In Table 2, we display the number of matches Nadal, Federer, and Williams have played from 2013-2017. Over that time span, Nadal won 6 grand slams, Federer won 1, Williams won 8. Despite this, Federer played more total matches in Nadal. All three players were absent for exactly three slams during this time period due to external factors. Not unexpectedly, Nadal has the most wins at FO (29), Federer has the most wins at Wim. (29), and Williams the most on hardcourt (30 at AO and 21 at US Open).

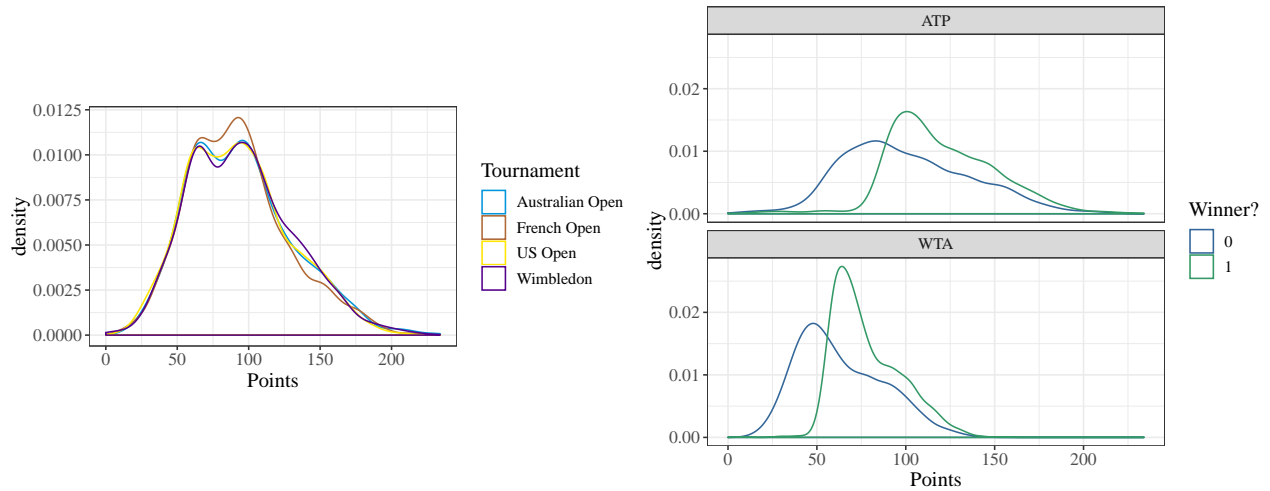
All data is obtained from Jeff Sackmann’s open website via the R package `deuce` (Sackmann 2018; Kovalchik 2017). All steps of our analysis from collection to dissemination are freely available online.

2.2 Early Data Analysis

2.2.1 Examining the distribution of points earned

We first examine the distribution of points earned per match to assess normality. The distribution of points earned is approximately normal. This distribution is similar across tournament, with Wimbledon differing slightly from the other grand slams. As expected, there are more points earned in the WTA than the ATP due to the differing numbers of games played. Also unsurprisingly, the winners of the match tended to earn more points than the losers.

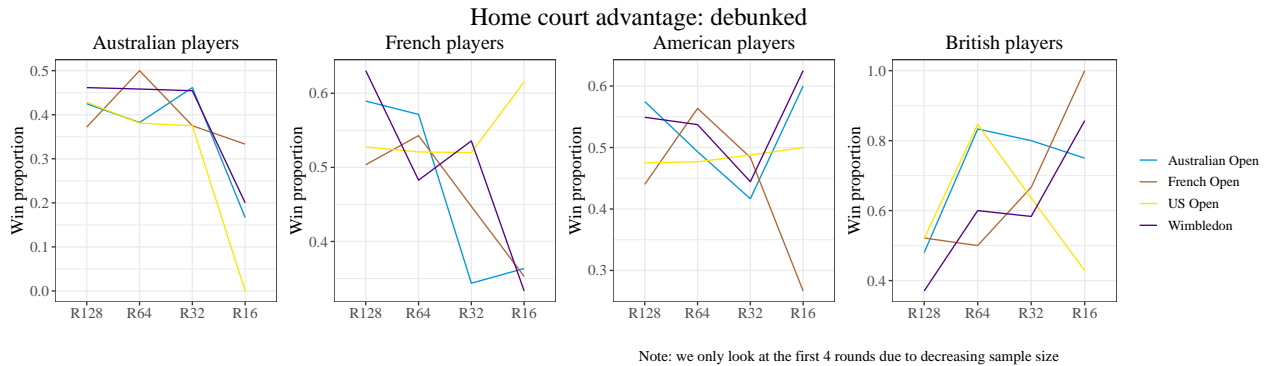
Distribution of points per match



2.2.2 Home court advantage

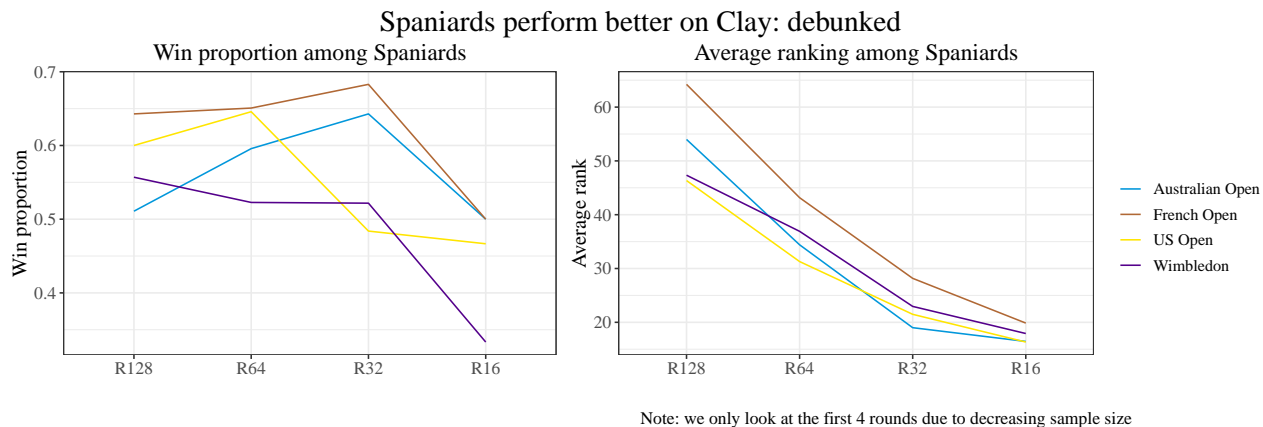
It is commonly thought that there is a home court advantage in grand slam games (SOURCE). In our data we find this to be true (i.e. French players win the French open more than French players win other slams). But, we also know that the home team is given preference for wild card bids (SOURCE) so potentially citizens of a particular country play in “their” tournament more often than they play in other tournaments. We also find this to be true in our data for France, The United States, and Australia (i.e. the proportion of French players in the French Open is greater than the proportion of French players in other slams).

Therefore, we want to see how the proportion of wins for the home country changes across the different tournaments. If there was really a home court advantage, the proportion of French wins each round would be higher at the French Open than the Australian Open, the US Open, and Wimbledon. The same would be true for Australia and the US. But, we see that this isn’t the case. After accounting for the number of players from each country, we don’t find a home court advantage in the grand slam.



2.3 Spaniards on clay

It is also commonly thought that Spaniards play better on clay. We are interested in whether Spaniards win the French Open more than they win in other tournaments. It does appear that Spaniards are winning the French Open more than they are winning other tournaments. But, this result is not significant. In addition, the ranking of Spaniards in the French Open is, on average, higher than other tournaments, which may help explain this common misconception, although this result is not significant either.

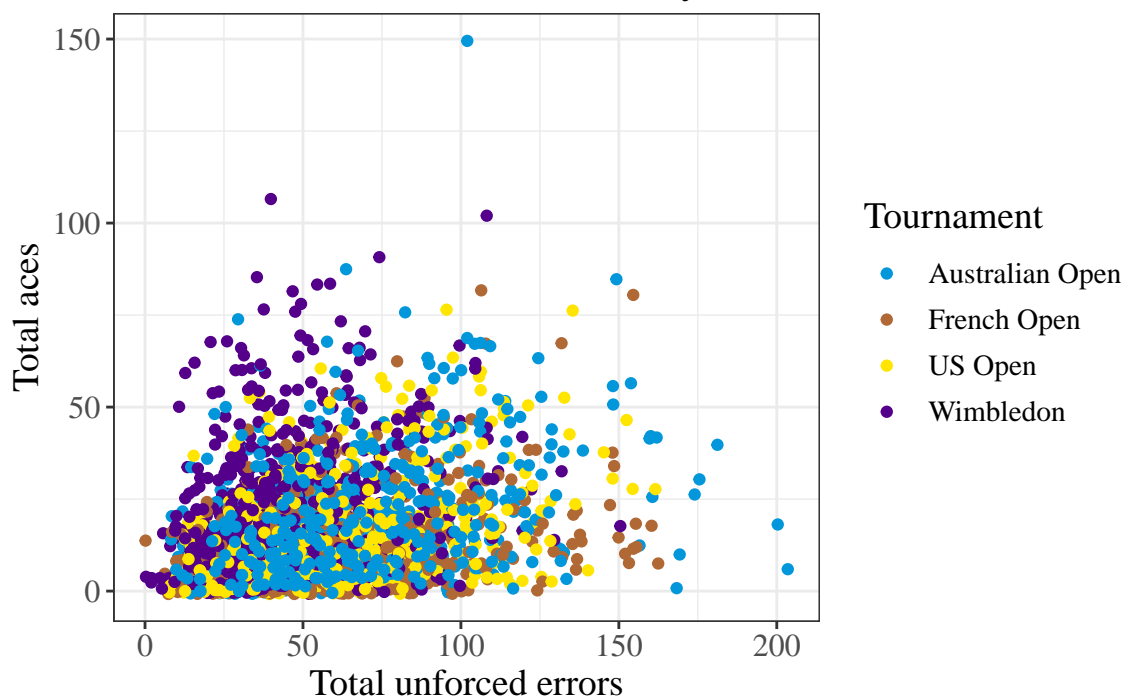


3 dont even know

Why the actual EFF won't fig.width work?

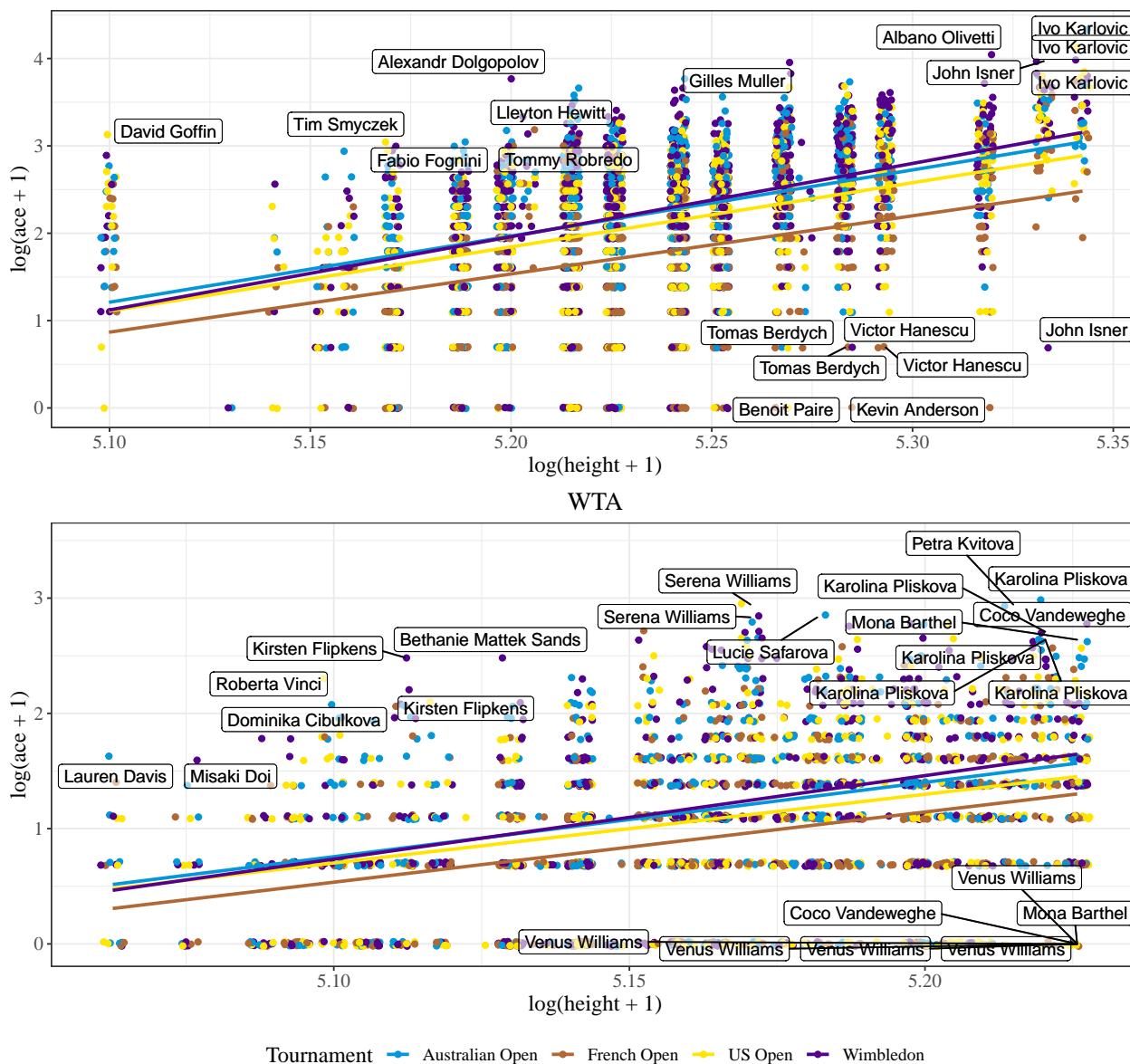
We find that the distribution of errors and aces differ by tournament. Specifically, GAMES? in Wimbledon follow different ace and error patterns than the other tournaments.

Distribution of errors and aces differ by tournmanet



3.1 Tall players and aces

Taller players are better servers: confirmed ATP



4 Methods

4.1 Hierarchical models

We use a hierarchical modeling approach to examine the grand slam results from all players from 2013-2017. This allows us to include fixed effects, which remain the same for each player, along with player-level individual effects. Since players appear multiple times in the data, the observations are not independent and including a player-level effect allows us to account for this dependence. It also provides a way to assess individual player effects.

Table 3: AIC summary of the seven logistic regression models fitted. Including individual effects for each grand slam, without including any country effects, leads to the best model fit according to AIC.

Model	Fixed Effects	Random Effects	AIC	EDF
no_court	No country, no tournament	none	11398.51	69
no_random_ef	all	none	11404.48	72
country_ef	no country, no tournament	tournament (by country)	11374.86	19
ind_ef	no tournament	tournament (by individual)	11395.69	79
ind_no_ioc	no country, no tournament	tournament (by individual)	11354.20	19
ind_intercept	no country	intercept (by individual)	11365.07	13
ind_year_ef	no country, no year	year (by individual)	11366.68	23

Table 4: Correlation matrix for random effects

Australian Open	French Open	US Open	Wimbledon
1.00	0.00	0.94	0.93
0.00	1.00	-0.05	-0.36
0.94	-0.05	1.00	0.88
0.93	-0.36	0.88	1.00

4.1.1 Modeling wins with logistic regression

We initially planned to use points won by an individual as our outcome variable, but found that it may not be an ideal measure of player performance. For instance, players may score few points in a match due to a poor performance, but they may also score few (relative) points if they win a short match. Modeling the number of points won is also complicated by the difference between length of matches for men (best of 5 sets) and women (best of 3 sets). For these reasons, we chose to model whether a player won the match using a hierarchical logistic regression approach, with possible predictors including country of origin, tournament that the match was played at, whether the match occurred in Round of 16 or later, rank of player, rank of opposing player, year of the match, and whether it was a men’s or women’s match.

We considered a variety of models incorporating effects as either fixed or random. That is $\pi_i = \beta X_i + \delta_i$, where $\pi_i = P(Y_i = 1)$, X_i are the fixed effects included in the model, and δ_i is a random effect across either players or years. Based on AIC, we found the model that included a player-level term for each tournament, but excluded fixed terms for country, was the best fit for the data. The models are summarized in table (REF TABLE).

If we examine the fixed effects for all seven models, we see that the parameter estimates do not substantially change across models (REF FIGURE). Rank and opponent rank appear to best explain the probability of winning the match. Other fixed effects in the models (whether it was a late round in the grand slam, the year, or ATP/WTF) were not found to be significantly different than zero. If we look at the random effects for Williams, Federer, and Nadal, however, the individual effects are not informative (especially after taking uncertainty into consideration).

More broadly, the effects for the Australian Open, US Open, and Wimbledon are all quite correlated. This suggests that, after accounting for additional variables such as rank and opponent rank, differences in win probability are detectable between clay and the other court surfaces, but not between the two types of hard courts and grass.

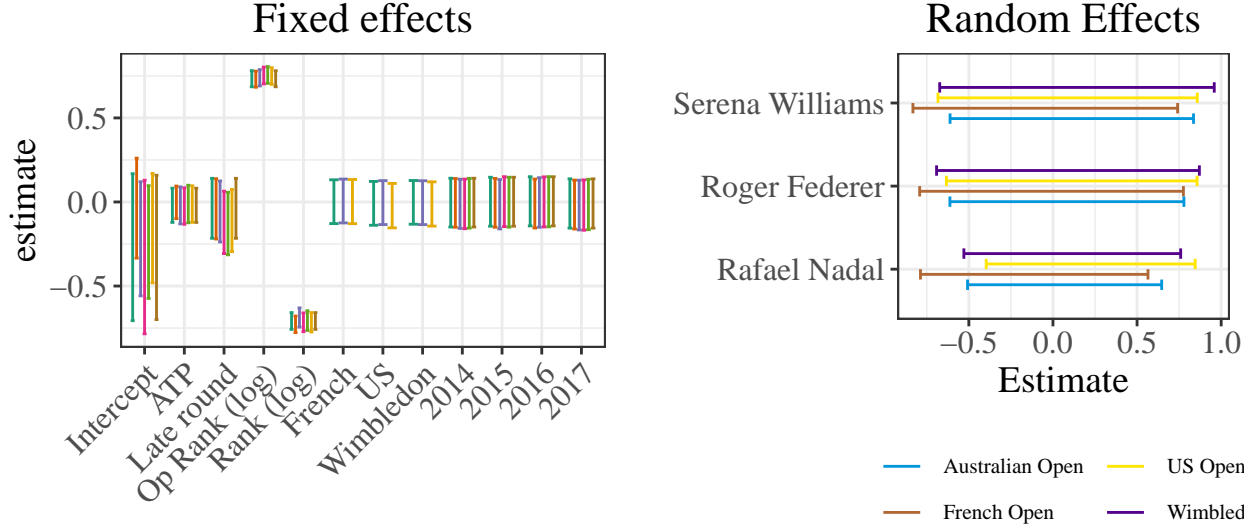


Figure 1: Estimated fixed effects for all seven models and player-level effects for Williams, Federer and Nadal under the `ind_no_ioc` model.

4.1.2 Linear regression for other outcomes

Instead of using ‘win’ as the outcome variable, we’ll now look at game statistics that may be more sensitive to court surface – aces, net points won, and unforced errors – and see if individual differences among players are detectable. We fit three additional models, each of the form $Y_i = \beta X_i + \delta_i$, where $\delta_i = \alpha_i T_i$, T_i is the four grand slam tournaments, and X_i includes the same variables as the model above.

If we look at the random effects for Williams, Federer, and Nadal (FIGURE REF), differences are much more noticeable than in the logistic regression model. For instance, we would expect Williams to have more aces at the Australian Open and Wimbledon than at the French Open. We also see that Federer’s performance at Wimbledon is consistently better than other tournaments in both aces and unforced errors. Interestingly, an increased performance on clay is not observed for Nadal in aces, net winners, or unforced errors.

4.2 Modelling players individually

We also fit individual linear models for Williams, Federer, and Nadal (subsetting the partial data to their matches only), estimating the percent of points won in a match using forward-backwards stepwise regression. The lower model is the percent of points regressed on the opponent rank and indicator variables for surface with one court as the reference variable (in the below equation, the reference surface as FO),

$$E[Y_{\%pts.won}|X] = \beta_0 + \beta_1 X_{opp.rank} + \beta_2 AO \mathcal{I}(X_{sur.} = AO) + \beta_2 Wim. \mathcal{I}(X_{sur.} = Wim.) + \beta_2 USO \mathcal{I}(X_{sur.} = USO).$$

The upper model is the lower model with the additional variables of winner to unforced error ratio (W/UE), average serve speed, percent aces, percent break points, percent net points won, and their interaction effects with court type. We do not display the full best fit models here, but they are available online. We do report the effects (± 2 . Std. Err.) of significant variables and their interaction terms, adjusting for the other covariates. Since we postulated that Nadal is best on clay, Federer on grass, and Williams on hard court, specifically at the Australian Open, we use the French Open, Wimbledon, and Australian Open, respectively, as the reference variable in the model. For Nadal and Federer we use forward-backwards stepwise regression, using AIC as the criterion, beginning with the full model. For Williams, we also use forwards-backwards stepwise regression but begin at the low model since she has more missing data.

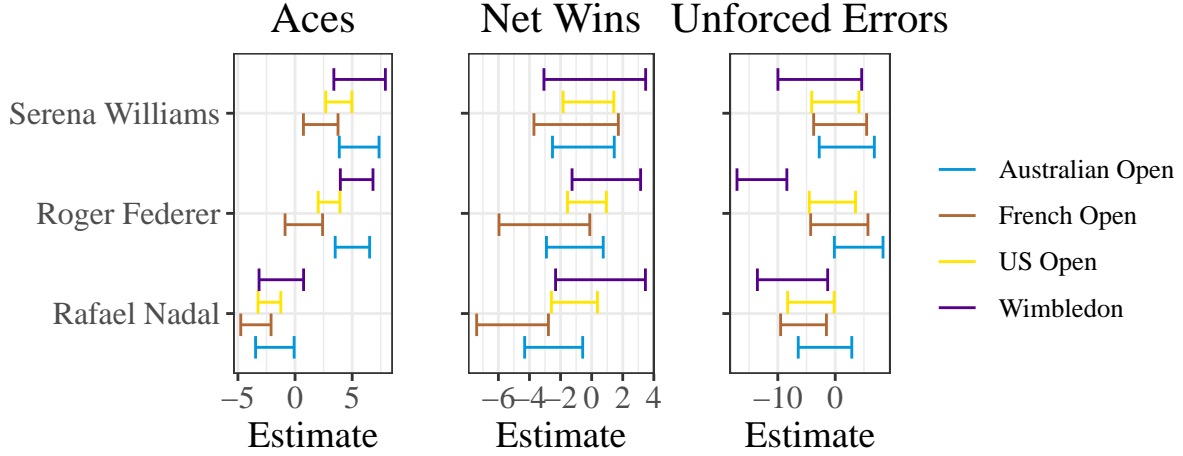


Figure 2: Estimated player-level effects for Williams, Federer and Nadal when modeling aces, net winners, and unforced errors.

4.3 Nadal: King of Clay: confirmed

Looking at Nadal individually, the model with the lowest AIC has positive coefficients for opponent rank, percent of aces, and percent of break points won. It has negative coefficients for Australian Open, US Open, and Wimbledon compared to French Open and for percent of net points won. There are no interaction effects in this model.

The standardized residuals plot shown in Figure 3 (along with the other linear regression diagnostics plots not shown here) shows that the model is a good fit for predicting the percent of points won. For the significant variables ($\alpha = 95\%$), after adjusting for the other variables, we expect percent of points won to be 0.034 ± 0.028 less at the Australian Open and 0.046 ± 0.033 less at Wimbledon compared to the French Open; to increase by $1.5 \times 10^{-4} \pm 1.3 \times 10^{-4}$ for a one unit increase in opponent rank; a $0.0012 \pm 5.6 \times 10^{-4}$ increase for a .01 increase in percent of break points won; and to decrease by $8.8 \times 10^{-4} \pm 5.5 \times 10^{-4}$ for a .01 increase in percent of net points won. As such, we see clear evidence Nadal performs better at the French Open compared to the other grand slams.

4.4 Federer: Wimbledon extraordinaire: confirmed

For Federer, the model with the lowest AIC has negative coefficients for Australian Open, French Open, and US Open vs Wimbledon, and W/UE. It has positive coefficients for opponent rank, percent of break points won, and all the interaction terms: court and W/UE.

The standardized residuals plot shown in Figure 3 shows that the model is a fairly good fit for predicting the percent of points won, but almost seem to see two clusters of residuals splitting with predicted point percentage of about 0.63. For the significant variables ($\alpha = 95\%$), while adjusting for the other variables, we expect percent of points won to be 0.04 ± 0.034 less at Australian Open compared to Wimbledon; 0.058 ± 0.046 less at French Open compared to the Wimbledon; 0.007 ± 0.093 less at US Open compared to Wimbledon; to decrease by 0.022 ± 0.019 for a one unit increase in W/UE at Australian Open, 0.073 ± 0.066 for a one unit increase in W/UE at French Open, and 0.041 ± 0.033 for a one unit increase in W/UE at US Open. This is strong evidence that Federer performs best at Wimbledon but possibly does better at the USO, depending on his W/UE ratio.

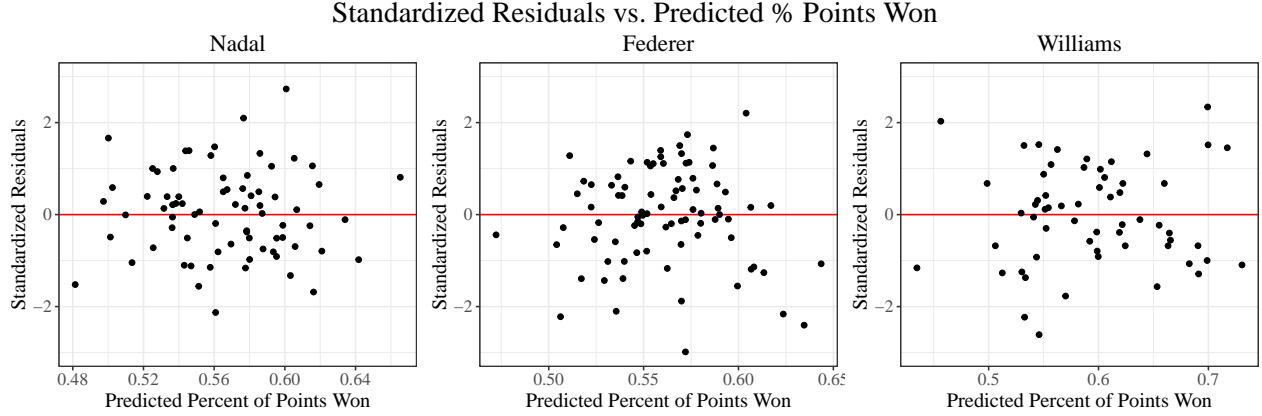


Figure 3: Standardized residuals vs percent of points won best fit model from forward-backwards stepwise regression for Nadal’s, Federer’s, and William’s, respectively.}

4.5 Williams: hard court magician only: debunked

For Williams, the model with the lowest AIC selected 22 coefficients, which makes it the largest model of the three. Since Williams only has 59 observations, we believe this model is overfitting and so caution the reader to be wary of any inference from this model. It should also be noted that no women’s point by point data was recorded for 2015 (in which she won three grand slams) so that year is excluded from this model. The coefficients are surface, opponent rank, W/UE, average serve speed, percent aces, percent break points won, percent net points won, and interaction effects with court and W/UE, court and average service speed, court and percent aces, and court and break points won.

The standardized residuals plot shown in Figure 3 shows that the model is a fair fit for predicting the percent of points won, but we tend to overestimate the percent of points won in comparison with Nadal’s model. For the significant variables ($\alpha = 95\%$), while adjusting for the other variables, we expect percent of points won to increase by $4.8 \times 10^{-4} \pm 3.2 \times 10^{-4}$ for a 1 unit increase in opponent rank; to increase by 0.034 ± 0.02 for a one unit increase in W/UE at French Open, to increase by 0.092 ± 0.045 for a one unit increase in W/UE at US Open, and to increase by 0.0038 ± 0.066 for a one unit increase in W/UE at Wimbledon compared to Australian Open; to increase by $2 \times 10^{-4} \pm 1.6 \times 10^{-4}$ for a .01 increase in percent of aces at the French Open; and to increase by 0.0024 ± 0.0011 for a .01 increase in percent of break points won at the French Open.

Based on the model results, we do not see that Williams is dominant on the hard courts compared to the other courts. Rather, she has great results on all the courts (and has the titles to prove it). Her significant interaction effects show that if she focuses on aces and capitalizing on breakpoints, then she has a better chance at winning more points at the French Open than at the Australian Open. *Is this too speculative?* (AL)

5 Discussion

At the highest level, we find that the general point distribution seems to be the same for all court surfaces. There may be some differences in player performance on the different court surfaces, but these differences are not definitive (e.g. that Spaniards perform better on clay). We did not find evidence for a home court advantage across all players in the grand slams. We also found that the French Open seems to be the most different out of the four grand slams.

Using a hierarchical model approach, we did not see strong individual effects for Williams, Federer, and Nadal in predicting wins. We did, however, see differences for these players in hierarchical models to predict

aces, net winners, and unforced errors, and that these differences were largely consistent with common tennis knowledge.

Individual linear models predicting percent of points won showed that Nadal was best on clay, Federer was best on grass, and Williams was strong in all grand slams.

Throughout our analyses, we often found that we were restricted by available data. Due to the large amount of grand slam data available, we were able to fit individual models for Federer, Nadal, and Williams that provided substantially more information about their performance on different court surfaces than the more general models which included all players. Ideally, these types of models would be possible for all players, not just the legends of the game. Point by point data from matches among lower-profile players is more likely to be missing or contain errors.

- HIGH LEVEL
 - in general point distribution seems to be the same,
 - at the high level there are some surface differences that we have evidence for but cannot be sure either way: spainiards,
 - there is no home court advantage
 - French Open seems most different of the tournaments
 - country does not seem to have an effect but individual does with court
- INTERMEDIATE LEVEL
 - Aces, Winners, and UEs are all important
 - Taller people have more aces
 - Wimbledon seems to have more winners and fewer UEs
- LOW LEVEL
 - Hierarchical model doesnt show much effect for Nadal, Federer, and Williams
 - Individual models predicting % points won does show Nadal good on clay, fed on grass, and williams more jack of all trades
 - we have less info about higher ranked (worse) players -> we need more data to pick out these effects
- FUTURE WORK:
 - closer look at point by point data
 - include other masters tournaments
 - pay more attention to year and players over time

6 References

- Andersson, Helena, Björn Eklblom, and Peter Krstrup. 2008. "Elite Football on Artificial Turf Versus Natural Grass: Movement Patterns, Technical Standards, and Player Impressions" 26 (February): 113–22.
- Barnett, Tristan, and Graham Pollard. 2007. "How the Tennis Court Surface Affects Player Performance and Injuries." *Medicine and Science in Tennis* 12 (1). Society for Tennis Medicine; Science: 34–37.
- Gains, Graydon L, Andy N Swedenhjelm, Jerry L Mayhew, H Michael Bird, and Jeremy J Houser. 2010. "Comparison of Speed and Agility Performance of College Football Players on Field Turf and Natural Grass." *The Journal of Strength & Conditioning Research* 24 (10). LWW: 2613–7.
- Jurejko, Jonathan. 2018. "French Open 2018: Why does 'King of Clay' Rafael Nadal reign supreme?" Edited by BBC Sport at Roland Garros. <https://www.bbc.com/sport/tennis/44385223>.
- Klaassen, Franc J.G.M., and Jan R. Magnus. 2003. "Forecasting the Winner of a Tennis Match." *European Journal of Operational Research* 148 (2): 257–67. [https://doi.org/https://doi.org/10.1016/S0377-2217\(02\)00682-3](https://doi.org/https://doi.org/10.1016/S0377-2217(02)00682-3).
- Klaassen, Franc J.G.M., and Jan R Magnus. 2001. "Are Points in Tennis Independent and Identically Distributed? Evidence from a Dynamic Binary Panel Data Model." *Journal of the American Statistical Association* 96 (454). Taylor & Francis: 500–509. <https://doi.org/10.1198/016214501753168217>.
- Knottenbelt, William J., Demetris Spanias, and Agnieszka M. Madurska. 2012. "A Common-Opponent Stochastic Model for Predicting the Outcome of Professional Tennis Matches." *Computers & Mathematics with Applications* 64 (12): 3820–7. <https://doi.org/https://doi.org/10.1016/j.camwa.2012.03.005>.

- Kovalchik, Stephanie. 2017. *Deuce: Resources for Analysis of Professional Tennis Data*.
- Kovalchik, Stephanie Ann. 2016. "Searching for the Goat of Tennis Win Prediction." *Journal of Quantitative Analysis in Sports* 12 (3). De Gruyter: 127–38.
- Leitner, Christoph, Achim Zeileis, and Kurt Hornik. 2009. "Is Federer Stronger in a Tournament without Nadal? An Evaluation of Odds and Seedings for Wimbledon 2009." *Austrian Journal of Statistics* 38 (4): 277–86.
- Lewit, Chris. 2018. "Secrets of Spanish Tennis: Paradigms and Superstructure." Edited by Tennisplayer.net. https://www.tennisplayer.net/public/classiclessons/chris_lewit/secrets_of_spanish_tennis_culture/?format=print.
- McHale, Ian, and Alex Morton. 2011. "A Bradley-Terry Type Model for Forecasting Tennis Match Results." *International Journal of Forecasting* 27 (2): 619–30. <https://doi.org/https://doi.org/10.1016/j.ijforecast.2010.04.004>.
- Morris, Benjamin. 2015. "Serena Williams and the Difference Between All-Time Great and Greatest of All Time." Edited by FiveThirtyEight. <https://fivethirtyeight.com/features/serena-williams-and-the-difference-between-all-time-great-and-greatest-of-all-time/>.
- Newton, Paul K, and Joseph B Keller. 2005. "Probability of Winning at Tennis I. Theory and Data." *Studies in Applied Mathematics* 114 (3). Wiley Online Library: 241–69.
- Paxinos, Stathi. 2007. "Australian Open court surface is speeding up." Edited by The Age. <https://www.theage.com.au/sport/tennis/australian-open-court-surface-is-speeding-up-20071120-ge6chj.html>.
- Sackmann, Jeff. 2018. "Tennis Data Repositories." <https://github.com/JeffSackmann>.
- Thomas, A. C., Samuel L. Ventura, Shane T. Jensen, and Stephen Ma. 2013. "COMPETING Process Hazard Function Models for Player Ratings in Ice Hockey." *The Annals of Applied Statistics* 7 (3). Institute of Mathematical Statistics: 1497–1524. <http://www.jstor.org/stable/23566482>.
- Wasserman, Larry, ed. 2004. *All of Statistics: A Concise Course in Statistical Inference*. Springer.