In tennis, do smashes win matches?

It is summer, and Wimbledon and the French and US Opens are with us. In tennis, does the man with the fastest serve win? And can you make too few unforced errors for success? **Sarah Budrus**, **Susan Vander Plas** and **Dianne Cook** analyse what gets the star players to match point.

How to win a Grand Slam

Have you ever wondered what it takes not only to be a great tennis player but also to win a Grand Slam title? To be a great tennis player, do you have to be able to smash the ball harder than your opponent? Andy Murray won the 2012 US Open, and he can hit the ball at 145 miles per hour. Is serving aces the key to success? Roger Federer hit 50 aces against Andy Roddick to win the 2009 Wimbledon final. (This is the record for a Grand Slam final, but not a Wimbledon record: in 2010, in the first round, John Isner hit 113 aces against Nicolas Mahut – and Mahut himself hit 103 aces against Isner. The totals are explained by the fact that theirs is the longest match on record, lasting 11 hours and 5 minutes over three days of play and 183 games.) We wanted to find the qualities that make for – and ideally that predict – Grand Slam winners. And to test our conclusions, we wanted to see if they were worth betting on. Were they good enough to win us money?

Four tournaments make up the Grand Slam: the Australian Open, the French Open, Wimbledon and the US Open. We scraped data from their websites (REF: ……..) and examined the match statistics for the 2012 Grand Slam for each player. We investigated the characteristics that make it more likely for him or her to reach the quarter-finals. We chose the quarter-finals as our benchmark because as players advance in a

tournament, they win more prize money. A first round win in the 2013 Australian Open paid out A\$27 600 (US\$28 400), but winners' earnings roughly double for each round they reach. Making the quarter-finals brings in A\$250 000 (US\$ 260 000), which is a substantial reward.

The tournament websites contain up-to-date measurements on the percentage of first serves in, first and second serve points won, the number of aces, double faults, unforced errors, winners, receiving points won, break point conversions, sets played, tie breaks played, server points won, total games, return games won, fastest serve speed, and average first and second serve speeds. (A 'winner' is defined as a shot that is not reached by the opponent and wins the point. See Wikipedia (http://en.wikipedia.org/wiki/Tennis_terms) for a good explanation of the other terms.) These statistics were collected for the first three rounds, and averaged over these rounds for analysis. In addition, the number of matches played was recorded (as rounds 1–7, where round 5 is the quarter-finals), so that we could examine which match statistics are most related to winning.

An obvious first question is: how different are men's and women's games? Figure 1 displays side-by-side boxplots comparing men and women on selected match statistics for all four Grand Slams of 2012. Men, unsurprisingly, serve faster and hit more aces, but

The authors can make money betting on tennis – but they don't bet on the man with the fastest serve

women make fewer unforced errors. Return points won are roughly equal for men and women. Differences between tournaments can be seen as well. Median serve speeds increase over the year for both men and women. The Australian Open has more variance in unforced errors, probably reflecting the hiatus in competition just prior to this tournament: it is the first of the season, and follows a long break, which may leave players unprepared for the tournament.

Figures 2 and 3 examine whether smashes win matches using scatterplots. Round reached is plotted against winners (i.e. smashes) and unforced errors (top, middle rows), and separately the number of winners is plotted against unforced errors (bottom), for the women's and men's matches of all four Grand Slams. The relationship is interesting! To get to the quarter-finals it is important to hit winners, but, surprisingly, not too many of them (or too few). Amazingly, the player who hit the most winners in a single match – more than 80 of them in fact – lost that match. She is Bethanie Mattek-Sands, in the Australian Open, and is represented by the lone dot on the extreme right in the first box of Figure 2.

A similar pattern is seen with unforced errors. You need to make some errors (those who make none do not get to the quarter-finals) but not too many. What these plots say is that to win you need to make more winners

Bethanie Mattek-Sands hit 80 winners in the Australian open – and lost the match

than unforced errors – in other words, keep your ratio of winners to unforced errors below 1 (the white diagonal line).

Figures 4 and 5 display scatterplots of round reached by speed of fastest serve (top), percentage of second serve points won (middle), and return points won (bottom), in the women's and men's matches for all four Grand Slams. Surprisingly, serve speed does not matter – even for men – as much as winning your second serve, and returning your opponents serve to win the point! One player, Sara Errani, made the French Open finals and the US
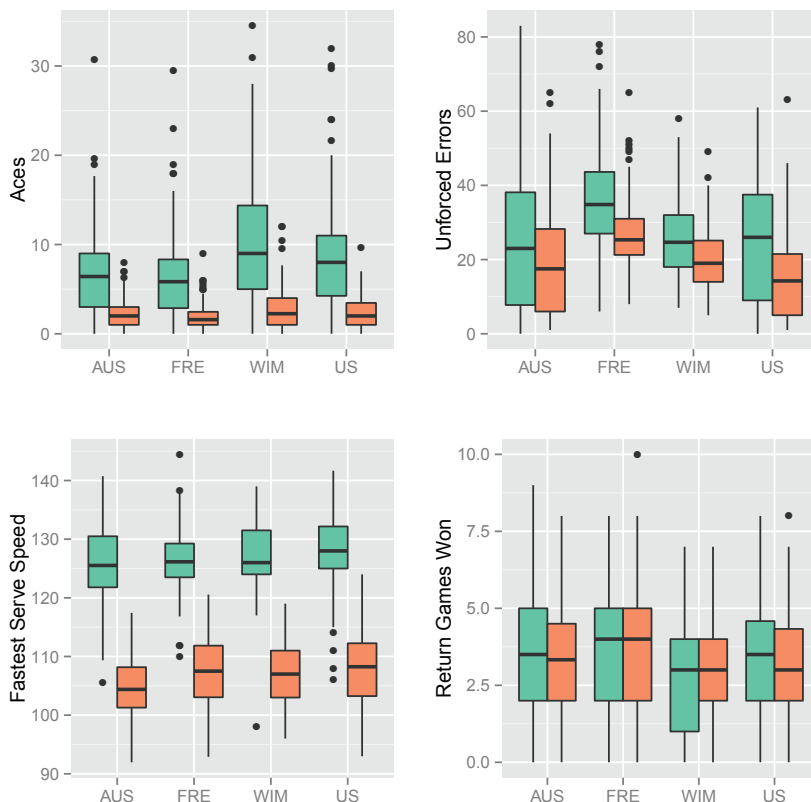

Figure 1. Battle of the sexes! Plots of aces, unforced errors, fastest serve speeds, and return games won, for men (green) and women (orange) for the four Grand Slams of 2012
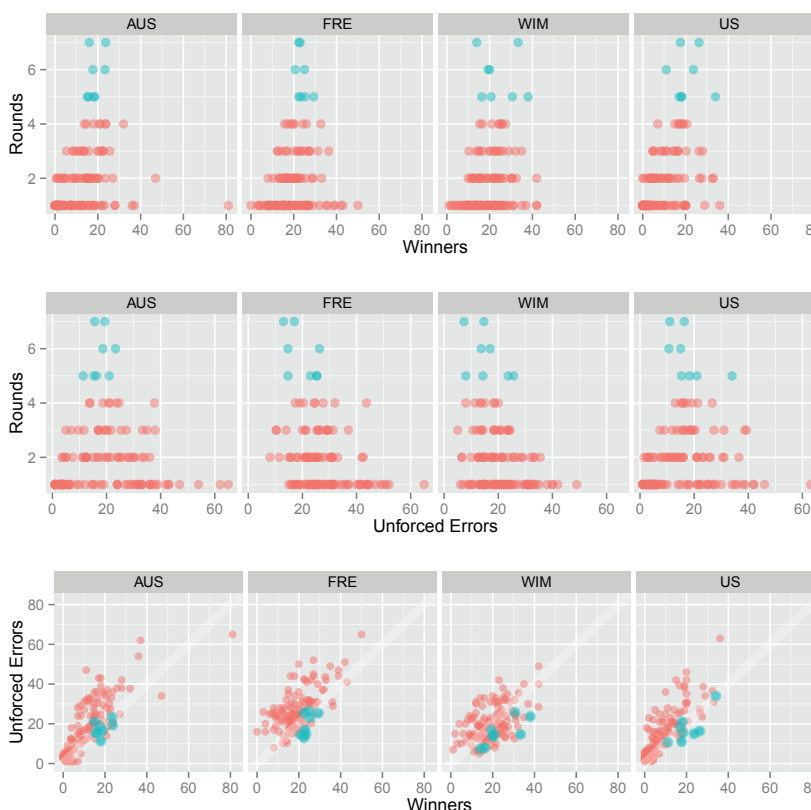

Figure 2. Do smashes win matches? Plots examining winners and errors and success in making it further in the tournament (blue), for women in all four Grand Slams
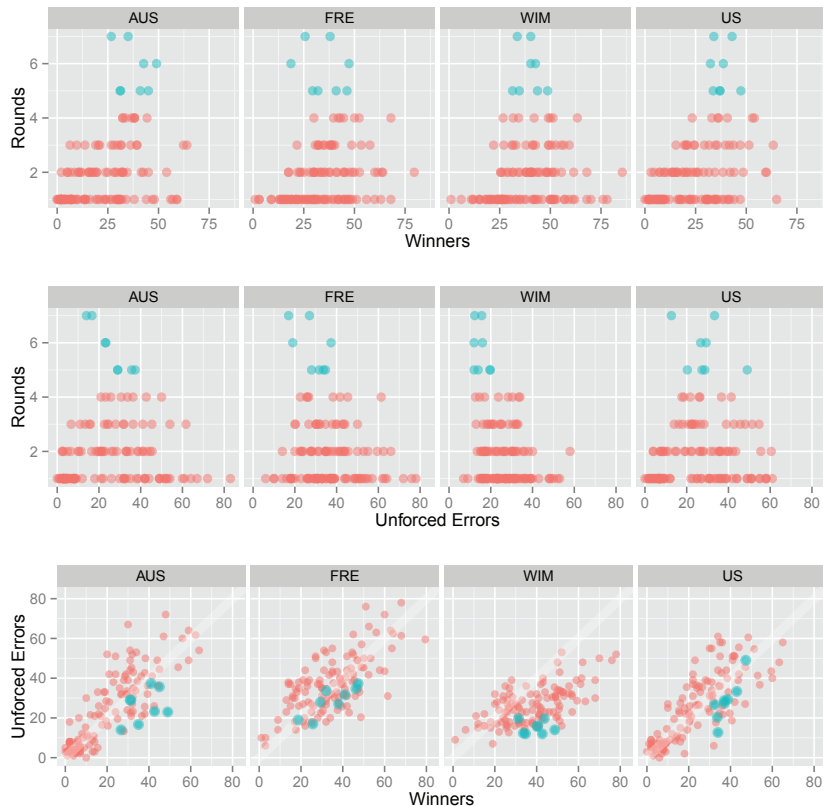
Figure 3. Do smashes win matches? Plots examining winners and errors and success in making it further in the tournament, for men in all four Grand Slams
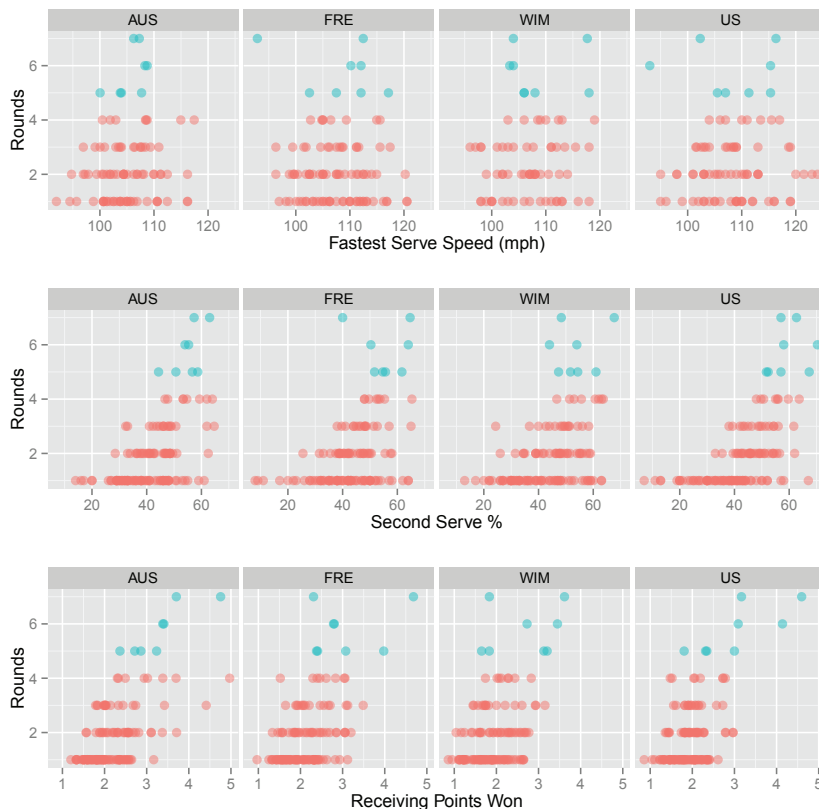

Figure 4. How important is serving? Plots examining serve speed, and defensive statistics, for women for all four Grand Slams

Open semi-finals with a first serve speed less than 100 mph. She did this by virtue of impeccable placement of the serve. Again she can be identified on the diagram.

Second serve percentage and receiving points won have the strongest relationship with the number of rounds reached. For all the women who got to the quarter-finals, their second serve percentage was above 50, roughly, and they all won 40 or more receiving points

Serve speed does not matter even for men as much as winning your second serve

per match. Men's second serve percentage needed to be above 45 and receiving points won percentage needed to be above 35. This suggests that it is harder for men to win return of serve than it is for women, which again makes sense because men have faster serve speeds. In all tournaments but the men's French Open where Florent Serra won 56% of his receiving points but went out in the second round, the player with the highest percentage of receiving points won made the finals.

Table 1 shows the top few statistics and cut-offs in determining whether a player makes the quarter-finals, for both the men and women. If you do not beat the cut-offs, you do not reach the quarter-finals. The cut-offs were calculated based on the statistics that players reaching the quarter-finals achieved, and averaged over the tournaments.

Returning the opponent's serve well turned out to be the most important player characteristic. First serve winning percentage and second serve winning percentage both need to meet a bare minimum, and unforced errors cannot be too low or too high. These were the top four variables that predicted making the quarter-finals.

And fans of Samuel Groth, who delivered the fastest serve ever recorded, of 164.3 mph at the 2012 ATP Challenger in Busan, might care to know that being able to serve really fast did not make it as one of the most important characteristics of Grand Slam success. In fact it did not even win him his ATP Challenger match.
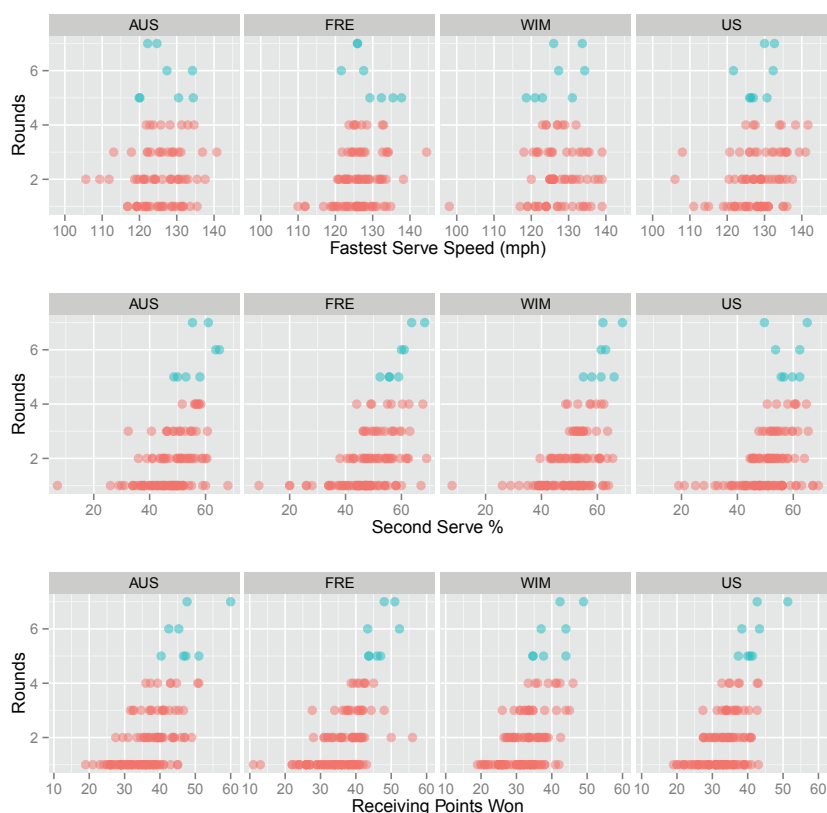
Figure 5. How important is serving? Plots examining serve speed, and defensive statistics, for men for all four Grand Slams

The bottom line

Smashes are important, but only up to a point! The players who are successful are those who force the pace of the game with smashes, but who do not overdo it. Defensive play is probably more important: being able to win points on the opponent's serve, and winning points on one's own second serve, correlate best with getting through to the quarter-finals and the big money prizes.

Can we make money on our predictions?

How good were our conclusions – in other words, how useful for predicting future winners? Were they strong enough to bet on, and make money on? There was only one way to find out. Our data above are from the 2012 Grand Slams. The 2013 Australian Open began in January. We asked people to bet against us on it. We set the odds; would we come out winners or losers?

To set our odds we developed scores for each player based on our criteria, on their performances in the 2012 Australian and US Open tournaments (both are played on hard surfaces, so players should perform similarly) and their play in the first two rounds of the 2013 Australian Open. Odds were made by calculating a weighted sum of the scores for the components we have described above. The actual weights were made with some trial and error, to get odds that made more sense based on how we had seen players perform in the past year. For men, history needed to be weighted more than for women, which surprised us initially. Table 2 shows a sample of the odds that we offered.

The bet was on sets of eight players who the punters thought would make the quarter-finals. People interested in betting were asked to watch the first two rounds of the 2013 Australian Open tournament and then to select the eight male and eight female players they believed would get through. They paid 10c per player. For each of their chosen players, if they won, the better would win the decimal portion of the odds value, otherwise they would lose their 10c. For example, Serena Williams' odds were 1.1, so if she made the quarter-finals the better would win 11c, but if she lost they would lose 10c.

The experiment was conducted in the week of January 14th, 2013, and we earned 7c on the dollar! (It did not, alas, make us millionaires. All proceeds went to the department's graduate student organization.) But it looks like we can make money!

Sarah Budrus and Susan Vander Plas are graduate students at Iowa State University, where Dianne Cook is Professor of Statistics.

Table 1. The most important statistics and values to match for men and women aiming to make the quarter-finals

Statistic	Men's rule	Women's rule
Return points won	> 39.0	> 45.6
First serve winning %	> 74.3	> 61.8
Second serve winning %	> 52.9	> 48.8
Unforced errors	between 22.3, 35.2	between 12.3, 26.5

Table 2. Example odds for the women's and men's 2013 Australian Open

Player	Odds	Player	Odds
Serena Williams	1.1	Novak Djokovic	1.1
Maria Sharapova	1.2	Andy Murray	1.1
Sloane Stephens	1.2	Roger Federer	1.2
Caroline Wozniacki	1.2	Marin Cilic	1.3
Venus Williams	1.3	Juan Martin Del Potro	1.3
Victoria Azarenka	1.3	Jo-Wilfried Tsonga	1.4
Svetlana Kuznetsova	1.4	Tomas Berdych	1.7
Jamie Hampton	1.4	Stanislas Wawrinka	1.7