COACHING & SPORT SCIENCE REVIEW

International Tennis Federation
www.itfcoachingreview.com
April 2011. 19th Year. Issue 53. 4-8
ISSN 2225-4757
https://doi.org/10.52383/itfcoaching.v19i53.542

# Grand Slam men's singles tennis 1995-2009 Part 2: Points, games and sets.

Rod Cross & Graham Pollard.

University of Sydney.

ABSTRACT

Data from Grand Slam men's singles matches is presented for the period 1995 to 2009 showing how set score distributions have varied with time and across the four events. The relative number of tie-break sets provides a good indication of court speed, and so does the number of 6-0 or 6-4 sets. We show (a) how serve point probabilities vary between match winners and losers, (b) that match winners win about 9 out of 10 service games on average, while losers win about 7 out of 10 service games, (c) about 1/3 of all points are won by a player hitting a winner, (d) about 78% of match winners win the first set and (e) a player who wins the first set 7-6 is 2.5 times more likely to win the match than his opponent.

**Key words:** Statistics, Points, Games, Set scores.
**Received**: 16 de November de 2010
**Acepted**: 1 de April de 2011
**Corresponding author:** Rod Cross, University of Sydney.
Email: cross@physics.usyd.edu.au

## INTRODUCTION

In Cross and Pollard (2009), we presented data on each of the four Grand Slam events over the period 1991 to 2009. The data were related primarily to serve speeds, aces, double faults and tiebreak sets. In the present paper we present additional data for the same events concerning the manner in which points, games and sets are won, and the probabilities of winning points including the all-important break point conversions.

## SERVE POINT PROBABILITIES

A fundamental parameter in the statistical analysis of tennis matches is the probability, p, of a player winning a point on serve, commonly known as the serve point probability. Some of those points are won on the first serve and some on the second serve, but the combined result is easily calculated for both match winners (pA) and match losers (pB) from data published on the web during each tournament. Results can be calculated for each set, but we focus attention on the total number of serve points won by each opponent during a complete 3, 4 or 5-set match. Results are presented only for completed matches. During a whole tournament, it is common for five or six of the nominal 127 matches to terminate before completion due to an injury to one of the players.

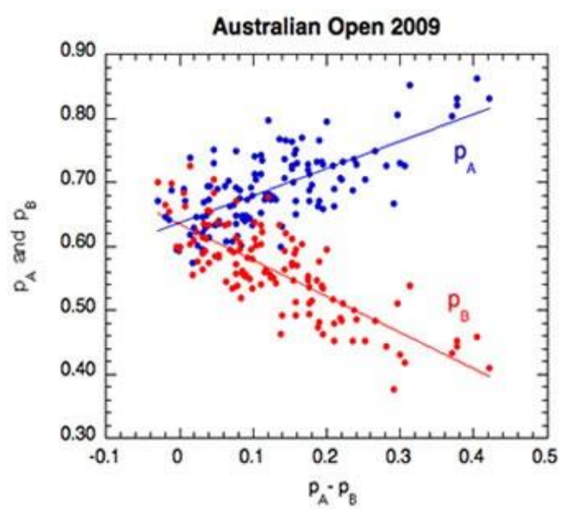


***Figure 1. Serve point probabilities pA (blue dots) and pB (red dots) for each of the 123 completed matches at the 2009 Australian Open men's singles event, as a function of the pA - pB difference for each***

In Fig. 1. we show the point probabilities for each player (one winner, one loser) for each of the 123 completed matches at the 2009 Australian Open, as a function of the pA - pB difference in each match. The usual result is that the match winner has a larger p value than the loser. If pA is close to pB then the match commonly takes 4 or 5 sets to determine a winner. Point probabilities depend on both the serving ability of the server and the returning ability of his opponent, and do not remain constant for any given player during a tournament or even from one set to the next in any given match. It is clear from Fig. 1 that pA depends on pB and vice-versa, with the result that pA can exceed 0.8 and pB can be as low as 0.4 when a strong player is pitted against a weak opponent.

Results similar to those in Fig. 1 were found for all four Grand Slam events, with small differences as summarised in Table 1. The main difference is that pA and pB (averaged over all players) are slightly higher at Wimbledon than at the other three events. No significant difference in average pA or pB values was found from the first round to later rounds in any of the four events. One might expect an increase in pA and pB from the first round to the later rounds, but the better players in the later rounds face stronger opponents, with the result that there is no significant change in pA or pB.

| EVENT | AUST 2009 | | FRENCH 2009 | | WIMB 2009 | | US 2008 | |
|---|---|---|---|---|---|---|---|---|
| | PA | PB | PA | PB | PA | PB | PA | PB |
| N | 123 | 123 | 121 | 121 | 120 | 120 | 120 | 120 |
| Mean | 0.691 | 0.563 | 0.700 | 0.570 | 0.721 | 0.611 | 0.695 | 0.591 |
| SD | 0.059 | 0.069 | 0.062 | 0.063 | 0.063 | 0.069 | 0.059 | 0.058 |

***Table 1. Serve point probabilities averaged over N match winners and N match losers in each Grand Slam event in 2008 or 2009.***

Serve point probabilities are closely related to the total number of points won during a match, as shown in Fig. 2. The probability of the match winner winning a point on serve is pA, so the probability that his opponent wins the point is 1 - pA. When his opponent is serving, the probability that the match winner wins a point is 1 - pB. Let R = the total number of points won by the winner of the match divided by the total number of points won by the loser. Each data point in Fig. 2 represents a single match, where the pA - pB difference is plotted on the horizontal axis, and the corresponding value of R is plotted on the vertical axis. When the players are evenly matched, with a small pA - pB difference, R is close to 1.

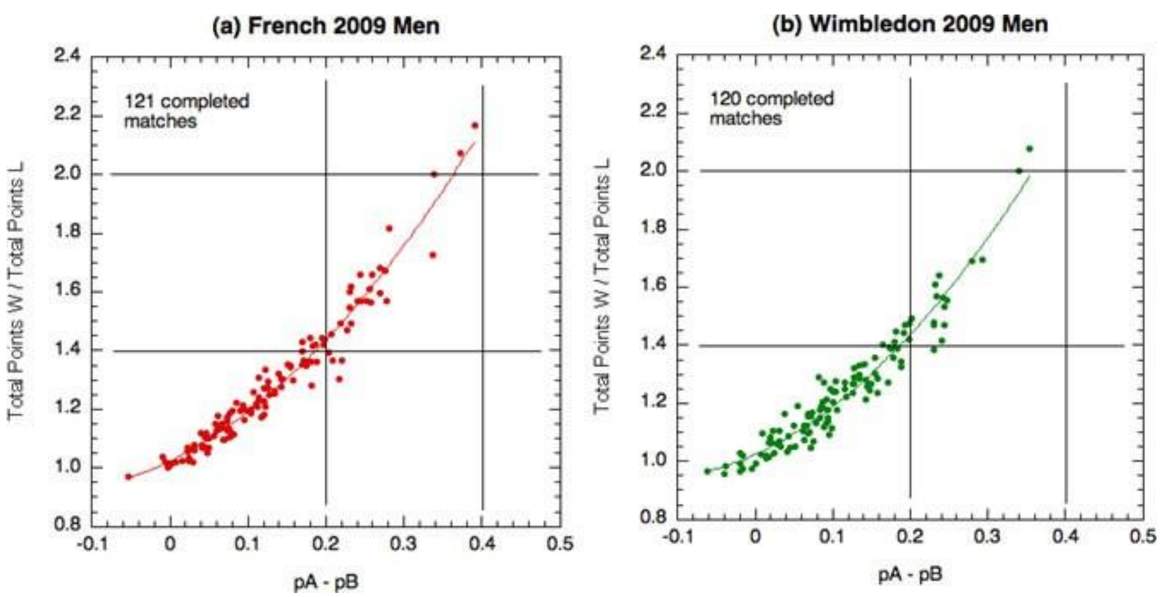


***Figure 2. R vs (pA - pB) for all completed men's singles matches played at (a) the French Open and (b) Wimbledon, in 2009. The curved lines are quadratic fits to the data points, with R = 1.025 + 1.301x + 3.799x2 for the French data and R = 1.025 + 1.226x + 4.167x2 for the Wimbledon data, where x = pA - pB.***

If the winner and loser each served the same number of times during a match, then R = 1 + (pA - pB) / 1 - (pA - pB)

In practice, winners and losers serve a different number of times, which accounts for the scatter in the data points shown in Fig. 2. Nevertheless, it is clear from the data in Fig. 2 that in a match where pA - pB = 0.2, for example, the match winner will win about 40% more points in total than the match loser, regardless of the actual values of pA and pB, and regardless of the particular court surface. Similarly, if pA - pB = 0.4 then the match winner will win more than twice as many points as the match loser. It is clearly important for a player to win more points than his opponent if he wants to win the match. The data used to construct Fig. 2 showed that only one match out of 121 was won at the French Open when R < 1, and only 8 matches out of 120 were won at Wimbledon when R < 1.

Despite the increase in serve speed, the increase in the number of aces, and the decrease in the number of double faults, serve point probabilities have not changed much over the years, as shown in Table 2. No data were collected for the 2000 Australian Open or the 2001 US Open. The main change has been an increase in pA at the French Open, associated with the increase in first serve speed, with the result that pA at the French Open is now about the same as that at the other three events. These results imply that the return of serve has improved over the years.

| YEAR | AUST OPEN | | FRENCH OPEN | | WIMBLEDON | | US OPEN | |
|---|---|---|---|---|---|---|---|---|
| | pA | pB | pA | pB | pA | pB | pA | pB |
| 1999 | 0.685 | 0.580 | 0.641 | 0.554 | 0.717 | 0.581 | 0.680 | 0.561 |
| 2000 | | | 0.657 | 0.552 | 0.702 | 0.605 | 0.699 | 0.606 |
| 2001 | 0.678 | 0.575 | 0.658 | 0.550 | 0.713 | 0.612 | | |
| 2008 | 0.690 | 0.571 | 0.685 | 0.556 | 0.716 | 0.618 | 0.695 | 0.591 |
| 2009 | 0.691 | 0.563 | 0.700 | 0.570 | 0.721 | 0.611 | | |

***Table 2. Serve point probabilities averaged over match winners and match losers in each men's singles Grand Slam event from 1999 to 2009.***

## BREAK POINT CONVERSIONS

Given the dominance of the men's serve in modern tennis, or the difficulty of breaking an opponent's serve, the opportunity of breaking serve is a significant event in the men's game. Analysis of break point conversion data for the 508 matches played at the four most recent events during 2008-2009 shows that if a match winner has a serve point probability pA > 0.82 then that player will not lose a single serve during the whole match, and if pA < 0.68 then the match winner will lose at least one of his service games during the match. The same pA limits (within ± 0.01) apply to all four events.

Table 3 shows the number of break point opportunities at each of the four most recent men's singles events (2008-2009), and the number of those opportunities converted to a service break. Data are shown only for completed matches. A count of the number of games won by each player (excluding tiebreak games) and the number of such games won as a result of break point conversions yields the fraction fW of service games won by match winners, and the fraction fL of service games won by match losers, averaged over all completed matches. Averaged over all four events, match winners win about 9 out of 10 of their service games, while match losers win about 7 out of 10 service games, both fractions being highest at Wimbledon and lowest at the Australian Open (in 2009). On average, match winners gain about twice as many opportunities to break serve as their opponents, and convert about 2.5 times as many games.

| | AUST 2009 | FRENCH 2009 | WIMB 2009 | US 2008 |
|---|---|---|---|---|
| Matches | 123 | 121 | 120 | 120 |
| Games | 4318 | 4235 | 4599 | 4444 |
| BPO(W) | 1600 | 1492 | 1244 | 1428 |
| BPO(L) | 815 | 789 | 703 | 805 |
| CBP(W) | 705 | 660 | 549 | 615 |
| CBP(L) | 269 | 247 | 216 | 255 |
| fW | 0.885 | 0.892 | 0.912 | 0.894 |
| fL | 0.636 | 0.654 | 0.729 | 0.700 |

***Table 3. Data on total number of break point opportunities and conversions at each men's singles Grand Slam event in 2008 or 2009. BPO = break point opportunities, CBP = converted break points, W = match winner, L = match loser***

## METHODS OF WINNING A POINT

There are five main methods of winning a point in tennis. A player can win the point himself by serving an ace or by hitting a clean winner. Alternatively, a player wins the point if his opponent serves a double fault, or makes an unforced or a forced error. The first four methods are listed as event statistics during each tournament. The number of forced errors for each player can be deduced from the total number of points won by each player during the match. The difference between a forced and an unforced error is a somewhat subjective judgement, but clear guidelines are given to those involved in recording the data. The number of clean winners struck by each player includes "service winners" which are defined as unplayable serves, in the sense that the player gets his racquet to the ball but the ball does not reach the net. An alternative description of a service winner would be a forced error, or even an ace, but that is not the way that it is recorded.

There is a wide variation between players in the methods or tactics used to win a point. One of the interesting facts that emerge from the statistics is that some players adopt safe tactics, hitting fewer winners and making fewer unforced errors than their opponents, while others adopt more risky tactics, hitting many more winners and making many more errors than their opponents. There is no guarantee that either tactic works best. Both methods have an approximately equal chance of success or failure. About 75% of matches are won by the winner hitting more clean winners than the loser. In those cases where the winner hits fewer clean winners than his opponent, the winner normally has fewer unforced errors.

The five methods of winning a point are shown in Table 4, as a percentage of all points won by match winners and losers, averaged over completed matches at each tournament. The sample size, N, at each event was limited by the fact that unforced errors and clean winners were not recorded for every match. Aces count for about 9% of points won, on average, although this figure can be around 25% for some players and less than 4% for others. About 1/3 of all points are won by a player hitting a clean winner, and about 1/3 are won by the opponent making an unforced error. Some players win more than half of their total points by hitting winners, while some can still win matches by winning less than 20% of their points as clean winners.

| EVENT | AUST 2009 | | FRENCH 2009 | | WIMB 2009 | | US 2008 | |
|---|---|---|---|---|---|---|---|---|
| | W | L | W | L | W | L | W | L |
| N | 82 | 82 | 118 | 118 | 119 | 119 | 25 | 25 |
| A% | 8.7 | 6.4 | 7.1 | 5.2 | 11.1 | 8.4 | 9.1 | 8.5 |
| W% | 35.6 | 33.2 | 35.1 | 32.0 | 37.0 | 33.8 | 35.9 | 36.3 |
| D% | 3.3 | 3.2 | 2.7 | 2.4 | 3.7 | 3.6 | 3.6 | 3.0 |
| U% | 33.7 | 33.5 | 28.0 | 29.6 | 23.3 | 22.8 | 31.5 | 30.7 |
| F% | 18.7 | 23.7 | 27.1 | 30.8 | 24.9 | 31.4 | 19.9 | 21.5 |
| Total% | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

***Table 4. Percentage of total points won by match winners (W) and losers (L) averaged over N matches in each Grand Slam event in 2008 or 2009. A% = Aces, W% = winners, D% = double faults, U% = unforced errors, F% = forced errors.***

## SET SCORE DISTRIBUTIONS

The set score distribution for all completed matches at each event, over the period 1995 to 2009, is summarised in Table 5. In order to compare all 3, 4 and 5-set matches on a similar basis, only the first 3 sets of each match were included in the summary, and only completed matches were included. The total number of sets for each possible set score was calculated for two separate periods, 1995 to 2004 and 2005 to 2009. The results were then normalised to a total of 1000 sets in each of the two periods. That is, the total of each column in Table 5 is 1000. The scores are listed in the usual manner, with the match winner listed first. A 3-6 score, for example, indicates that the match winner lost at least one of the first three sets 3-6.

The most common set score at each event, in each period, is 6-4. The next most common is 6-3. The third most common score is 6-2, except at Wimbledon where 7-6 is the third most common score. Tiebreak sets or 6-6 results are commonly used to monitor the speed of the court and the speed of the game itself. In that respect, we note from the set score distributions that (a) Wimbledon provides the fastest court, while the clay courts at the French Open are the slowest, and (b) the number of 6-0 results also provides an indication of court speed, as do the number of 6-4 results. The number of 6-1 or 6-2 sets also differs on each surface, but the differences or the trends are not entirely consistent with those indicated by the number of 7-6 sets.

The results in Table 5 include all 3, 4 and 5-set matches and therefore represent the combined set score distributions of the three separate outcomes. Each outcome has a separate distribution of set scores. For example, there are no 3-6 or 4-6 results in a 3-set match and there is a larger proportion of 6-3 or 6-4 results in a 3-set match than in a 4-set or a 5-set match.

Further details of the 1995-2004 data are described in Pollard, Cross & Meyer (2006).

| SET SCORE | AUST OPEN | | FRENCH OPEN | | WIMBLEDON | | US OPEN | |
|---|---|---|---|---|---|---|---|---|
| | 95-04 | 05-09 | 95-04 | 05-09 | 95-04 | 05-09 | 95-04 | 05-08 |
| 6-0 | 31.2 | 43.6 | 34.2 | 34.6 | 13.5 | 15.9 | 22.8 | 28.7 |
| 6-1 | 81.3 | 86.1 | 86.8 | 107.3 | 71.1 | 59.3 | 76.5 | 85.4 |
| 6-2 | 142.7 | 161.2 | 147.9 | 143.6 | 114.0 | 117.0 | 132.6 | 132.5 |
| 6-3 | 176.4 | 162.3 | 171.7 | 175.5 | 188.6 | 180.1 | 192.0 | 170.1 |
| 6-4 | 183.0 | 169.9 | 182.1 | 183.2 | 201.8 | 218.6 | 195.3 | 187.8 |
| 7-5 | 61.3 | 73.0 | 68.5 | 69.3 | 62.2 | 71.9 | 65.5 | 71.7 |
| 7-6 | 112.0 | 105.1 | 86.5 | 101.8 | 144.2 | 147.7 | 112.2 | 122.3 |
| 6-7 | 40.5 | 38.1 | 40.3 | 39.0 | 50.4 | 54.4 | 39.9 | 56.0 |
| 5-7 | 20.3 | 31.6 | 23.0 | 19.2 | 19.9 | 19.2 | 17.3 | 19.1 |
| 4-6 | 54.5 | 47.4 | 58.3 | 42.3 | 54.2 | 50.0 | 52.3 | 49.9 |
| 3-6 | 52.9 | 45.7 | 54.0 | 47.3 | 48.2 | 47.8 | 48.9 | 50.5 |
| 2-6 | 24.9 | 20.2 | 25.7 | 20.3 | 19.7 | 11.5 | 23.4 | 19.1 |
| 1-6 | 17.0 | 13.6 | 16.2 | 14.3 | 10.2 | 4.9 | 18.1 | 6.8 |
| 0-6 | 1.9 | 2.2 | 4.6 | 2.2 | 1.9 | 1.6 | 3.3 | 0.0 |

***Table 5. Set score distribution at each men's singles event for the periods 1995-2004 and 2005-2009, normalised to 1000 sets at each event and in each time period.***

The results in Table 5 provide useful guidelines in terms of analysing the progress of any given match. For example, suppose that a player loses the first set 0-6. What chance does he then have of winning the match? In Table 5, we see that there are about two 0-6 matches (in 1000) for every thirty 6-0 matches. A player who wins the first set 6-0 is therefore about 15 times more likely to win the match than his opponent.

Suppose that a player loses the first set 6-7. Such a result indicates that the two opponents are fairly evenly matched and suggests that both players might have about the same chance of winning the match. In fact, Table 5 shows that a player who wins the first set 7-6 is about 2.5 times more likely to win the match than his opponent. Why is this? The player who wins the first set needs to win only two more sets to win the match, but the player who loses the first set needs to win three more sets. If the players are evenly matched, then it is more likely that one player will win two of the next three or four sets rather than three of them.

About 22% of match winners lose the first set. The other 78% win the first set. During the 2005-2009 period, 63.6% of match winners won the first two sets, 3.9% of match winners lost the first two sets, 14.5% of match winners won the first set then lost the second set, and 17.9% of match winners lost the first set then won the second set. A similar result was found for the 1995-2004 period. Even though a match can be evenly balanced at the end of two sets, each player winning one set, the most likely match winner is the player who won the second set. The explanation can be found by analysing the set score distributions in Table 5 in finer detail. It was found that match winners tend to perform better as the match progresses (Pollard, Cross & Meyer, 2006). The effect is summarised in Table 6, for the 2005-2009 period, where we show the percentage of sets lost by match winners in sets 1, 2 and 3 for all completed matches. At each of the four events, match winners are much more likely to win the first set than to lose it, but if they do lose a set during the match then it is more likely that they will lose the first set rather than the second or the third set. At the Australian Open for example, 23% of match winners lost the first set, but only 17.8% of match winners lost the third set.

| | AUST OPEN | FRENCH OPEN | WIMBLEDON | USOPEN |
|---|---|---|---|---|
| Set1 | 23.0% | 20.1% | 22.1% | 22.3% |
| Set2 | 18.8% | 18.1% | 17.6% | 19.3% |
| Set3 | 17.8% | 17.2% | 17.1% | 18.8% |

***Table 6. Percentage of sets lost by match winners in sets 1, 2 and 3, for the 2005-2009 period.***

## CONCLUSIONS

In this study, and Cross & Pollard (2009), a considerable amount of data on the Grand Slam tournaments has been put together, summarized and interrelated. Some of the main conclusions of this paper are that;

(a) With the exception of the French Open, service probabilities have not changed much since 1999. For the French Open, the service probabilities are now much closer to those for the other tournaments.

(b) Regarding break point conversions, at present match winners average about twice as many opportunities as their opponents, and convert two and a half times as many games.

(c) At present about one-third of points are won by one player hitting a 'clean winner', and about one-third are won by the opponent making an unforced error. Aces account for about 9% of all points won.

(d) The most common set score is 6-4, and the next most common is 6-3. The third most common is 6-2, except at Wimbledon, where it is the score 7-6. Thus, the score 7-6 is commonly used to monitor the speed of the court and indeed the game itself.

## REFERENCES

Cross, R. and Pollard, G. (2009). Grand Slam men's singles tennis 1991-2009. Part 1: Serve speeds and other related data. ITF Coaching and Sport Science Review, 49, 8-10.

Pollard, G. Cross, R. and Meyer, D. (2006). An analysis of ten years of the four grand slam men's singles data for lack of independence of set outcomes, Journal of Sports Science and Medicine, 5, 561-566.