Journal of Quantitative Analysis in Sports

Volume 5, *Issue* 3 | 2009 | *Article* 7

Monte Carlo Tennis: A Stochastic Markov Chain Model

Paul K. Newton, *University of Southern California*
Kamran Aslam, *University of Southern California*

Recommended Citation:
Newton, Paul K. and Aslam, Kamran (2009) "Monte Carlo Tennis: A Stochastic Markov Chain Model," *Journal of Quantitative Analysis in Sports*: Vol. 5: Iss. 3, Article 7.

DOI: 10.2202/1559-0410.1169

Monte Carlo Tennis: A Stochastic Markov Chain Model

Paul K. Newton and Kamran Aslam

Abstract

We develop a stochastic Markov chain model to obtain the probability density function (pdf) for a player to win a match in tennis. By analyzing both individual player and 'field' data (all players lumped together) obtained from the 2007 Men's Association of Tennis Professionals (ATP) circuit, we show that a player's probability of winning a point on serve and while receiving serve varies from match to match and can be modeled as Gaussian distributed random variables. Hence, our model uses four input parameters for each player. The first two are the sample means associated with each player's probability of winning a point on serve and while receiving serve. The third and fourth parameter for each player are the standard deviations around the mean, which measure a player's consistency from match to match and from one surface to another (e.g. grass, hard courts, clay). Based on these Gaussian distributed input variables, we use Monte Carlo simulations to determine the probability density functions for each of the players to win a match. By using input data for each of the players vs. the entire field, we describe the outcome of simulations based on head-to-head matches focusing on four top players currently on the men's ATP circuit. We also run full tournament simulations of the four Grand Slam events and gather statistics for each of these four player's frequency of winning each of the events and we describe how to use the results as the basis for ranking methods with natural probabilistic interpretations.

KEYWORDS: tennis, Markov chains, Monte Carlo methods

Author Notes: P. K. Newton would like to acknowledge helpful comments on a draft of the manuscript by J.B. Keller and R. Stefani. Financial support is acknowledged through NSF-DMS0504308 and NSF-DMS0804629.

1 Introduction

A Markov chain model for tennis was developed in Newton and Keller (2005). As in previous models of Carter and Crews (1974) and G.H. Pollard (1983), it was based on a single input parameter for each player — the player's probability of winning a point on serve. This parameter was chosen as the total number of points won on serve divided by total points served for a player taken over many matches against a field of opponents. By solving a hierarchical system of recursion equations which links points to games, games to sets, and sets to matches, analytical formulas were worked out for each player's probability of winning a game, set, match, and tournament under all of the various scoring systems currently used on the men's (ATP) and women's (WTA) professional circuits. Because the input parameter is held constant throughout each match (interpreted as an assumption that points in tennis are independent, identically distributed random variables), the results, in a sense, model the way the scoring system in tennis effects outcomes as much as modeling individual player characteristics. More recently, Newton and Aslam (2006) explored non-iid effects by analyzing ensembles of Monte Carlo simulations of tournaments, with player parameters varying from point to point based on the notion of the point's 'importance', as introduced in Morris (1977). There have been several recent attempts at predicting outcomes of tennis matches based on player data. For example, the work of Barnett and Clarke (2005) combines player data in an attempt to predict outcomes, while earlier work of Clarke and Dyte (2000) uses the official rating system as input towards a simulation of tournaments. Walker and Wooders (2001) use minimax theory from econometrics to analyze tournament data. Most recently, the work of O'Malley (2008) builds on that of Newton and Keller (2005).[1]

In this paper, we develop a *stochastic* Markov chain model (see an introduction to these techniques in Asmussen and Glynn (2007)) based on the realization that a player's probability of winning a point on serve is not constant throughout a tournament, but varies from match to match and is better modeled as a random variable whose probability density function closely resembles a Gaussian (around the sample mean). Part of the reason a player's probability of winning a point on serve varies from match to match is because it depends on his opponent's probability of winning a point on receiving serve (the two must sum to one), and return of serve ability varies from player to player. Thus, we develop a new model, which uses four input parameters for

[1]Current state-of-the-art will be highlighted in the upcoming 2nd International Conference on Mathematics in Sport IMA Sport 2009 (Groningen, June 17–19, 2009 http://www.ima.org.uk/Conferences/maths_sport/index.html).

each player. The first two are the sample means associated with each player's probability of winning a point on serve and while receiving serve. The third and fourth parameter for each player are the standard deviations around the mean, which measure the player's consistency on serve and on return of serve from match to match and on various surfaces, such as grass (Wimbledon), hard courts (Australian Open, US Open) and clay (French Open). Physical characteristics of these surfaces are discussed in Cross (2008). Suffice it to say that it is widely believed (and supported by the data shown in figures 2 - 5) that the fastest surface (grass) favors dominant servers (offensive play), whereas the slowest surface (clay) favors receivers and consistency (defensive play). Using the analytical formulas obtained by solving the Markov chain system, in conjunction with Monte Carlo simulations using Gaussian distributed inputs, we obtain the probability density function (pdf) for a player to win a game on serve which, in the parameter range of interest, has an approximate normal distribution. We then run targeted Monte Carlo simulations of head-to-head player matches to obtain the pdf's for a player to win a match. A nice introduction to the properties and use of probability density functions can be found in Bendat and Piersol (1986). We also run full tournament simulations, using Gaussian distributed input data for the 128 players in the tournament, to obtain the statistical frequency of each player winning each of the four Grand Slam events, as well as their statistical frequency of winning n rounds. Then, the sample means obtained from these simulations are used as entries to a preference matrix (see Keener (1993)) whose dominant eigenvector (eigenvector corresponding to largest eigenvalue) provides a ranking of the players. To our knowledge, this is the first stochastic player based model for tennis that has the ability to yield rankings with a natural probabilistic interpretation.

Section §2 describes our model in detail. First, we analyze data obtained from the 2007 men's Association of Tennis Professionals (ATP) circuit for the 'field' (all player data lumped together), the four Grand Slam events, and four top players in the Men's ATP circuit: Roger Federer, Rafael Nadal, Andy Roddick, and James Blake. This sets the stage for viewing a player's probability of winning a point on serve or while receiving serve as (truncated) Gaussian distributed random variables. We describe the Markov chain equations whose solution links our Gaussian distributed inputs to the probability density function (pdf) for each player winning a match. We detail in this section how a player's return of serve ability is used in the model, and how the outputs depend on each player's standard deviation around the mean. Section §3 details the results of Monte Carlo simulations based on the full stochastic Markov chain model. We focus on individual player match-ups between four top players currently on the ATP tour and the results of full Grand Slam simulations using extensive player data from the 2007 ATP season. We also describe how to use the results of the Monte Carlo simulations as inputs to a matrix based ranking system.

2 The model

The starting point for our model are the probability density functions (pdf's) for each player to win a point on serve, denoted $p^s_A(x)$ (for player A), and for each to win a point while receiving serve, denoted $p^r_A(x)$. We interpret these variables as the player's probability as measured against the 'field' of players, as opposed to his probability as measured against any one particular opponent. When we refer later to player A's probability of winning a point on serve against a specific opponent, say player B, we will use the notation $p^s_{A|B}(x)$ and $p^r_{A|B}(x)$ to indicate conditional probabilities. We show in this section that the pdf's can be taken as truncated Gaussian distributions, with support in the interval $[0, 1]$ (their tails are sufficiently far from 0 and 1), which are completely characterized by their respective means, μ^s_A, μ^r_A, and standard deviations σ^s_A, σ^r_A.

2.1 Data

For each match played on the men's ATP circuit in the 2007 season, we obtain the percentage of points won on serve for each player and the percentage of points won while receiving serve. Figure 1 shows all of the combined player data, which we call 'the field', obtained by lumping together the data for 330 players over 59 tournaments, on all three surfaces (grass, clay, hard courts) over the full 2007 season. We denote the pdf's associated with 'field' data by $p^s_f(x)$ (pdf for the field to win a point on serve), and $p^r_f(x)$ (pdf for the field to win a point receiving serve). The corresponding means are denoted μ^s_f, μ^r_f, with standard deviations denoted σ^s_f, σ^r_f. The figure shows histogramed data for points won on serve, and points won while receiving serve (scaled to have unit area), together with the associated truncated Gaussian distributions with sample mean $\mu^s_f = 0.63316$ and sample standard deviation $\sigma^s_f = 0.094779$, and $\mu^r_f = 0.36684$ and $\sigma^r_f = 0.094779$. Note that $\mu^s_f + \mu^r_f = 1$, and $\sigma^s_f = \sigma^r_f$. We performed a chi-square goodness-of-fit test for normality with a sample size $N = 5245$, and found $\chi^2 = 11.45$, using values $\chi^2 \leq \chi^2_{9;0.10} = 14.68$ from Table A.3 in Bendat and Piersol (1986) that the hypothesis of normality is accepted at the $\alpha = 0.10$ level of significance. Our conclusion is that the probability density functions for the field can effectively be viewed as truncated normally

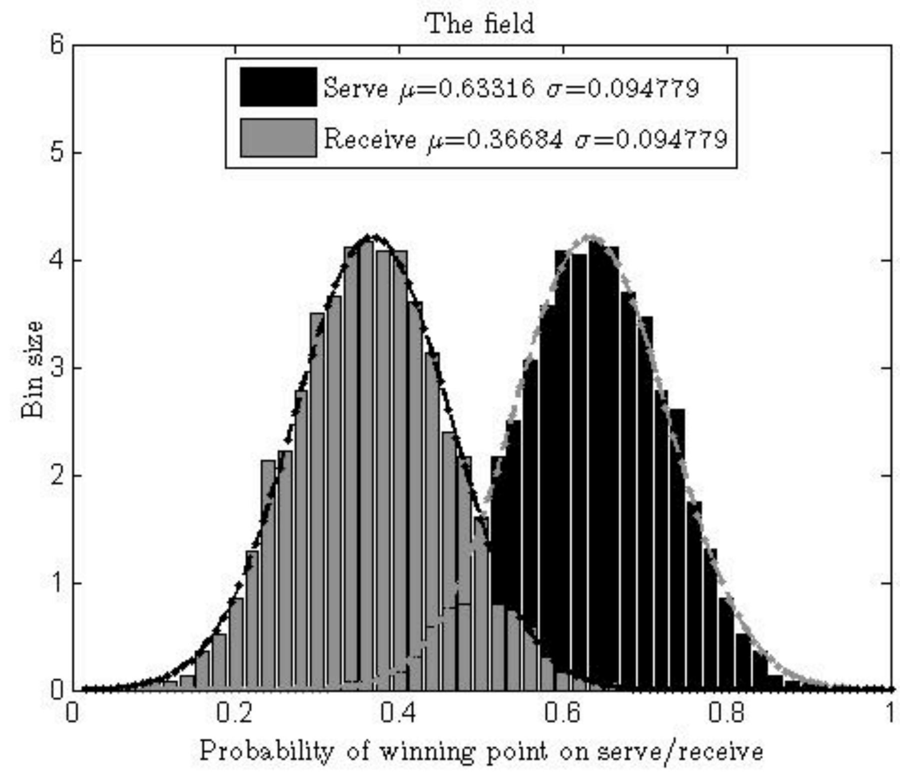

Figure 1: *The 'field' data are obtained by lumping together 330 players over 59 tournaments in the 2007 season. Plots show the histograms and Gaussian fits of the probability of winning a point on serve and receiving serve. Sample mean* $\mu_f^s = 0.63316$. *Sample standard deviation* $\sigma_f^s = 0.094779$. *Sample mean* $\mu_f^r = 0.36684$. *Sample standard deviation* $\sigma_f^r = 0.094779$.

distributed random variables, where :

$$\begin{aligned} \tilde{p}_f^s(x) &= (\sigma_f^s\sqrt{2\pi})^{-1}\exp\left(-\frac{(x-\mu_f^s)^2}{2(\sigma_f^s)^2}\right), \quad (-\infty < x < \infty) & (1) \\ p_f^s(x) &= C\tilde{p}_f^s(x), \quad (0 \le x \le 1), \quad \text{otherwise} \quad 0 & (2) \\ \int_{-\infty}^{\infty} p_f^s(x)dx &= C\int_0^1 \tilde{p}_f^s(x)dx = 1, \end{aligned}$$

DOI: 10.2202/1559-0410.1169

$$
\begin{aligned}
\tilde{p}^r_f(x) &= (\sigma^r_f\sqrt{2\pi})^{-1}\exp\left(-\frac{(x-\mu^r_f)^2}{2(\sigma^r_f)^2}\right), \quad (-\infty < x < \infty) \quad (3)\\
p^r_f(x) &= C\tilde{p}^r_f(x), \quad (0 \le x \le 1), \quad \text{otherwise} \quad 0 \quad (4)\\
\int_{-\infty}^{\infty} p^r_f(x)dx &= C\int_0^1 \tilde{p}^r_f(x)dx = 1.
\end{aligned}
$$

The four defining parameters for the probability density functions associated with the field are $\mu^s_f, \mu^r_f, \sigma^s_f, \sigma^r_f$.

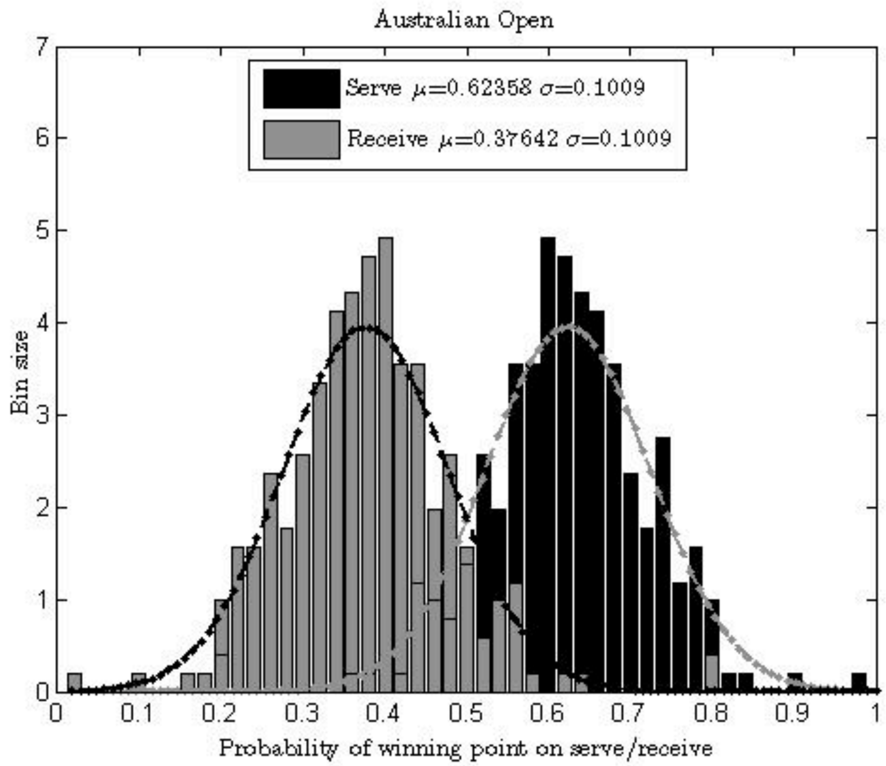

Figure 2: *Field data for the 2007 Australian Open (hard court) obtained by lumping together 128 players. Shown are histograms and Gaussian fits to data using sample means and standard deviations.*

Breaking the data down further, we show in figures 2 - 5 the field data for each of the four Grand Slam tournaments in the order in which they are played. Figure 2 shows data from the 2007 Australian Open (hard courts), histogramed together with the Gaussian fit, figure 3 shows data from the 2007 French Open (clay), figure 4 shows data from the 2007 Wimbledon (grass), and

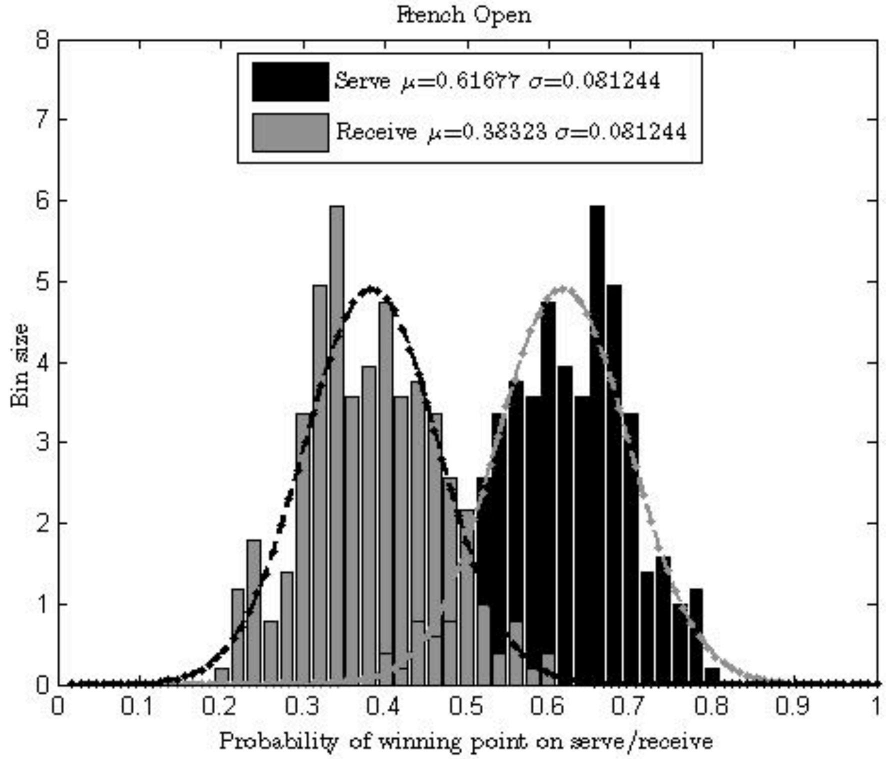

Figure 3: *Field data for the 2007 French Open (clay court) obtained by lumping together 128 players. Shown are histograms and Gaussian fits to data using sample means and standard deviations.*

figure 5 shows data from the 2007 US Open (hard courts). Because the data for each tournament is much more sparse than for the full season, the histograms are not as filled out as those in figure 1, yet the Gaussian distributions still model the density functions quite accurately. More specifically, for figures 2 - 5, using the chi-square goodness-of-fit test for normality, we found acceptance at the $\alpha = 0.10$ level of significance.

The data for the four Grand Slam events are summarized in Table 1. The sample mean for points won on serve (receiving serve) are highest (lowest) on grass, which is the fastest surface and favors players with dominant serves. The opposite is true for the slowest surface (clay). These statistical conclusions corroborate the physical characteristics of the respective surfaces with respect to the way the ball bounces, as discussed in a recent article by Cross (2008). Figures 6 - 9 show the individual player data taken over the full 2007 season for

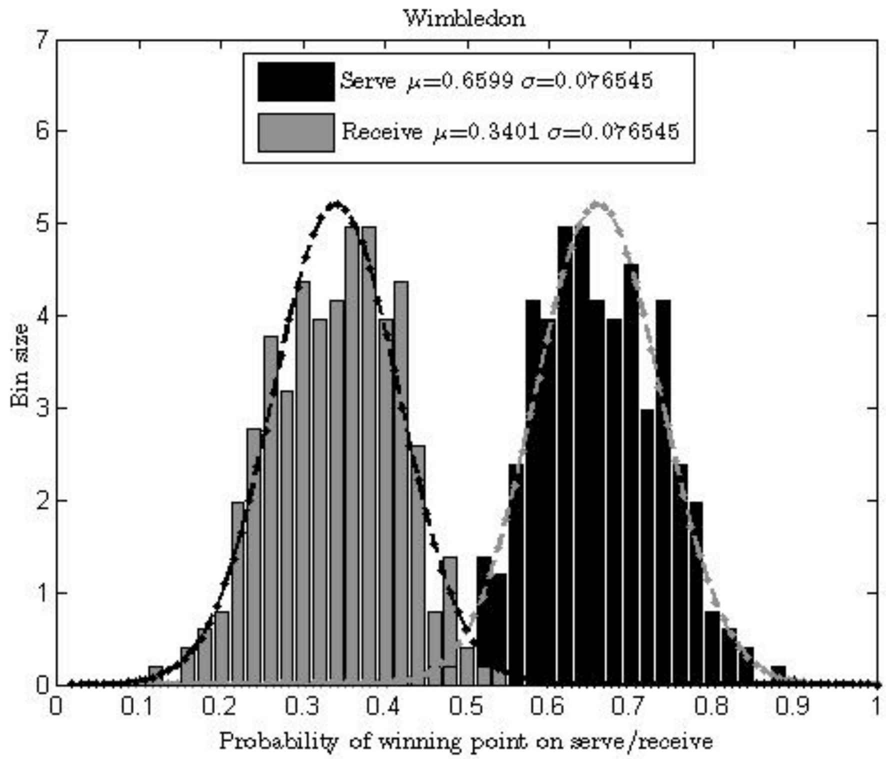

Figure 4: *Field data for the 2007 Wimbledon tournament (grass) obtained by lumping together 128 players. Shown are histograms and Gaussian fits to data using sample means and standard deviations.*

four top players of interest: Roger Federer (#1 2007 year end ranking), Rafael Nadal (#2 2007 year end ranking), Andy Roddick (#6 2007 year end ranking), and James Blake (#13 2007 year end ranking). Our notation is to underscore the variable using the initials of the player, hence Roger Federer's mean value for points won on serve is denoted μ^s_{RF} (as measured against the 'field'). Once again, although the data is much more sparse than that used in figure 1, we still conclude that the histograms would fill out Gaussian distributions if more data were available. Specifically, for these figures, we performed a chi-square goodness-of-fit test for normality, and found that the hypothesis of normality is accepted at (at least) the $\alpha = 0.10$ level of significance. For figure 7 (receive) and figure 9 (serve), acceptance was at the $\alpha = 0.05$ level of significance, while for figure 9 (receive) acceptance was at the $\alpha = 0.01$ level of significance. We note that Andy Roddick has the highest sample mean for points won on

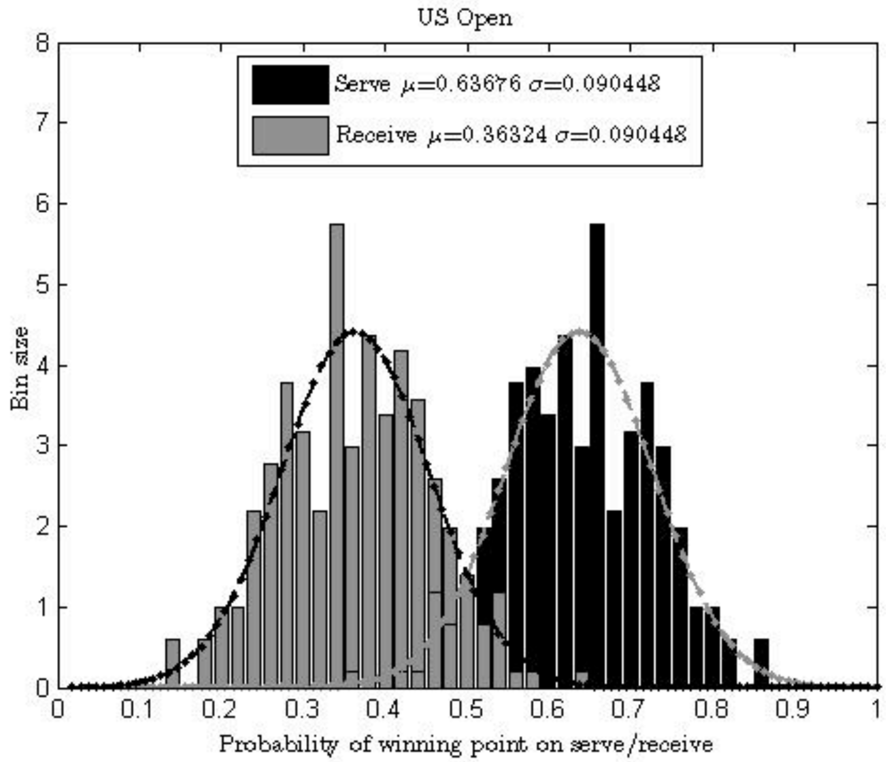

Figure 5: *Field data for the 2007 US Open (hard court) obtained by lumping together 128 players. Shown are histograms and Gaussian fits to data using sample means and standard deviations.*

serve of the four players, with value $\mu^s_{AR} = 0.73089$ (significantly higher than the field value $\mu^s_f = 0.63316$), while Roger Federer's sample mean for serve is lower, with value $\mu^s_{RF} = 0.70714$. This fact alone would lead to a prediction that Roddick would win 63% of his matches against Federer using the theory developed in Newton & Keller (2005) and Pollard (1983) all of whom use the percentage of points won on serve as the single input parameter. However, Federer won all three of their head-to-head matches in 2007, has a career record of $15 - 2$ against Roddick, and ended the year with the #1 ranking. There are two main reasons for this. First, although Roddick had the highest sample mean for points won on serve, his standard deviation for serve is also the highest, with value $\sigma^s_{AR} = 0.084164$, which we interpret as a measure of lack of consistency on serve. Also, he had the lowest sample mean for points won while receiving serve, with value $\mu^r_{AR} = 0.34411$ (compared with the field

Event	μ^s	σ^s	μ^r	σ^r
Australian Open (hard)	0.62358	0.1009	0.37642	0.1009
French Open (clay)	0.61677	0.081244	0.38323	0.081244
Wimbledon (grass)	0.6599	0.076545	0.3401	0.076545
US Open (hard)	0.63676	0.090448	0.36324	0.090448

Table 1: *Summary of field data (all 128 player data lumped together over all rounds) broken down for each of the four 2007 Grand Slam events.*

value $\mu_f^r = 0.36684$). Both of these additional parameters play an important role in our model.

2.2 The Markov system

The equation which links player A's probability of winning a point on serve, denoted p_A, to his probability of winning a game on serve, p_A^G, is the formula (first obtained by Carter & Crews (1974)):

$$\begin{aligned} p_A^G &= (p_A)^4[1+4q_A+10(q_A)^2]+20(p_Aq_A)^3(p_A)^2[1-2p_Aq_A]^{-1}; \qquad (5) \\ q_A &= 1-p_A. \end{aligned}$$

In the original non-stochastic model developed in Newton and Keller (2005), $p_A \in [0,1]$ is taken as a constant and was taken to be the sample mean of points won on serve divided by the total points served over a sufficiently large number of matches.

To obtain corresponding formulas for the probability of winning a set and a match, let p_A^S denote the probability that player A wins a set against player B, $q_A^S = 1-p_A^S$. To calculate p_A^S in terms of p_A^G and p_B^G, we define $p_A^S(i,j)$ as the probability that in a set, the score becomes i games for A, j games for B, with A serving initially. Then

$$p_A^S = \sum_{j=0}^{4} p_A^S(6,j) + p_A^S(7,5) + p_A^S(6,6)p_A^T. \qquad (6)$$

Here, p_A^T is the probability that A wins a 13-point tiebreaker with A serving initially, and $q_A^T = 1-p_A^T$.

To calculate the terms $p_A^S(i,j)$ needed in (6), we solve the following system of recursion equations:

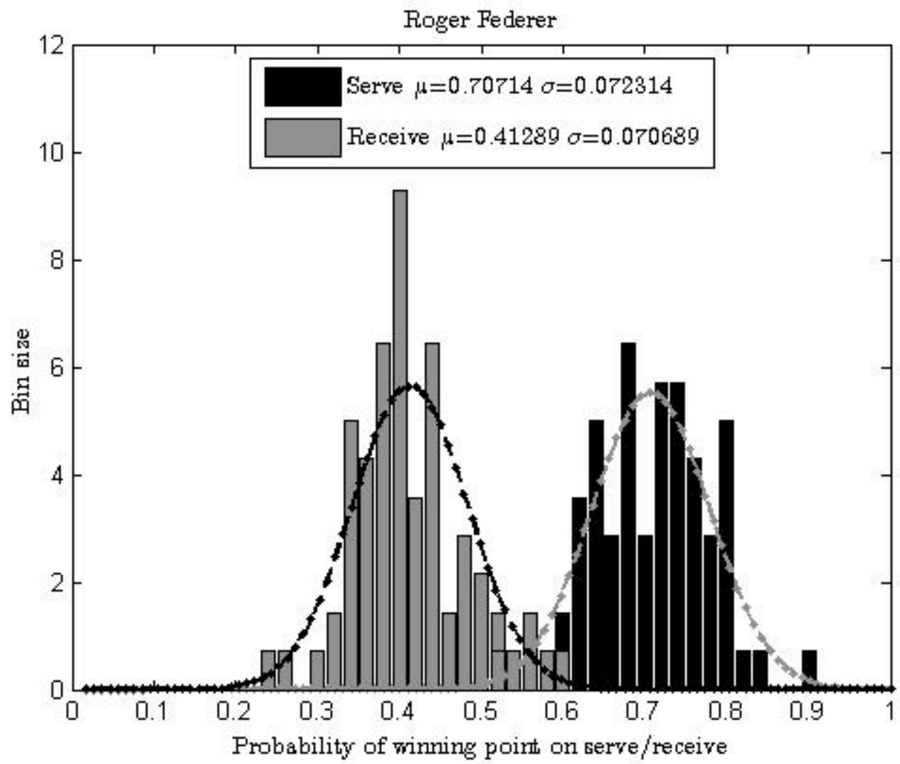

Figure 6: *Data for Roger Federer on serve and receive of serve for full 2007 season. Shown are histograms and Gaussian distributions using sample means and standard deviations.*

For $0 \leq i, j \leq 6$:

$$\textbf{if } i-1+j \textbf{ is even:} \qquad p_A^S(i,j) = p_A^S(i-1,j)p_A^G + p_A^S(i,j-1)q_A^G \tag{7}$$
$$\textbf{omit } i-1 \textbf{ term if } j=6\textbf{, } i \leq 5;$$
$$\textbf{omit } j-1 \textbf{ term if } i=6\textbf{, } j \leq 5$$

$$\textbf{if } i-1+j \textbf{ is odd:} \qquad p_A^S(i,j) = p_A^S(i-1,j)q_B^G + p_A^S(i,j-1)p_B^G \tag{8}$$
$$\textbf{omit } i-1 \textbf{ term if } j=6\textbf{, } i \leq 5;$$
$$\textbf{omit } j-1 \textbf{ term if } i=6\textbf{, } j \leq 5$$

along with the initial conditions:

$$p_A^S(0,0) = 1; \quad p_A^S(i,j) = 0 \qquad \text{if } i < 0, \text{ or } j < 0. \tag{9}$$

DOI: 10.2202/1559-0410.1169

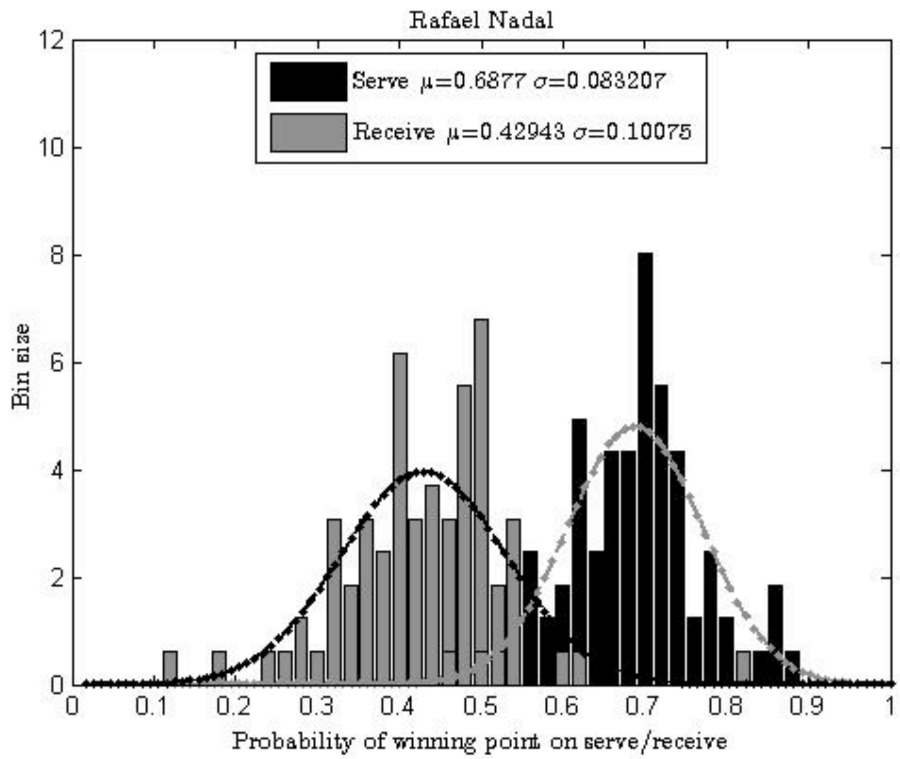

Figure 7: *Data for Raphael Nadal on serve and receive of serve for full 2007 season. Shown are histograms and Gaussian distributions using sample means and standard deviations.*

In terms of $p_A^S(6,5)$ and $p_A^S(5,6)$, we have

$$p_A^S(7,5) = p_A^S(6,5) q_B^G; \quad p_A^S(5,7) = p_A^S(5,6) p_B^G. \tag{10}$$

To calculate the probability of winning a tiebreaker, p_A^T, in terms of p_A and p_B, we define $p_A^T(i,j)$ to be the probability that the score becomes i for A, j for B in a tiebreaker with A serving initially. Then

$$\begin{aligned} p_A^T &= \sum_{j=0}^{5} p_A^T(7,j) + p_A^T(6,6) \sum_{n=0}^{\infty} p_A^T(n+2,n) \\ &= \sum_{j=0}^{5} p_A^T(7,j) + p_A^T(6,6) p_A q_B \left[1 - p_A p_B - q_A q_B\right]^{-1} \end{aligned} \tag{11}$$

To calculate $p_A^T(i,j)$, we solve:

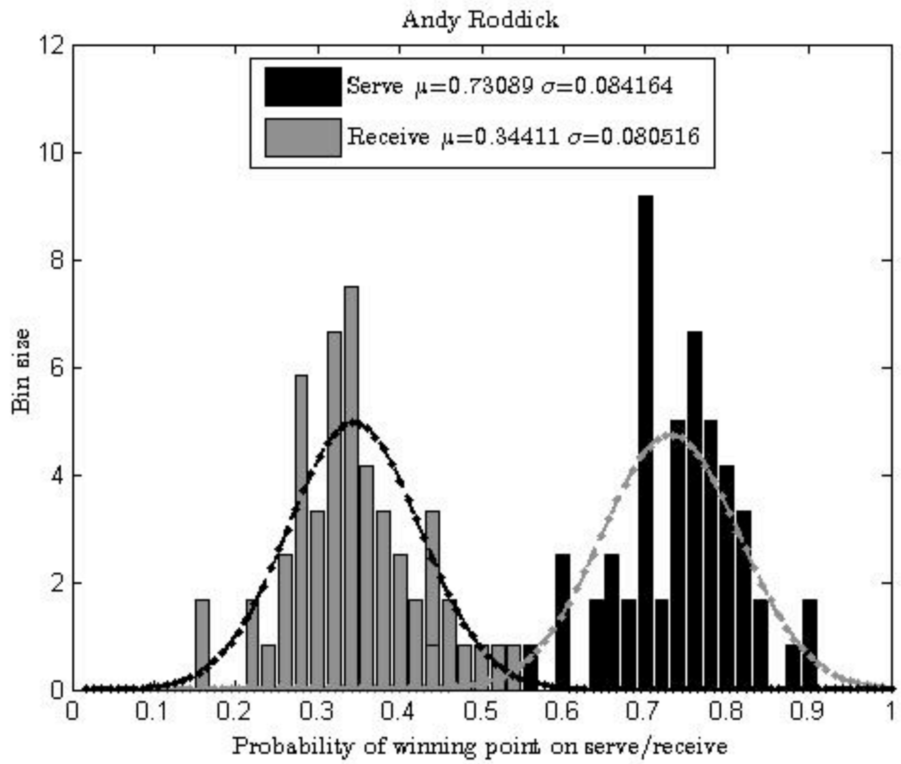

Figure 8: *Data for Andy Roddick on serve and receive of serve for full 2007 season. Shown are histograms and Gaussian distributions using sample means and standard deviations.*

For $0 \leq i, j \leq 7$:

$$\begin{aligned} &\textbf{if } i-1+j = 0, 3, 4, ..., 4n-1, 4n, ... \\ &p_A^T(i,j) = p_A^T(i-1,j)p_A + p_A^T(i,j-1)q_A \\ &\textbf{omit } j-1 \textbf{ term if } i=7\textbf{, } j \leq 6 \\ &\textbf{omit } i-1 \textbf{ term if } j=7\textbf{, } i \leq 6 \end{aligned} \tag{12}$$

$$\begin{aligned} &\textbf{if } i-1+j = 1, 2, 5, 6, ..., 4n+1, 4n+2, ... \\ &p_A^T(i,j) = p_A^T(i-1,j)q_B + p_A^T(i,j-1)p_B \\ &\textbf{omit } j-1 \textbf{ term if } i=7\textbf{, } j \leq 6 \\ &\textbf{omit } i-1 \textbf{ term if } j=7\textbf{, } i \leq 6 \end{aligned} \tag{13}$$

with the initial conditions:

$$p_A^T(0,0) = 1; \quad p_A^T(i,j) = 0 \qquad \text{if } i < 0, \text{ or } j < 0. \tag{14}$$

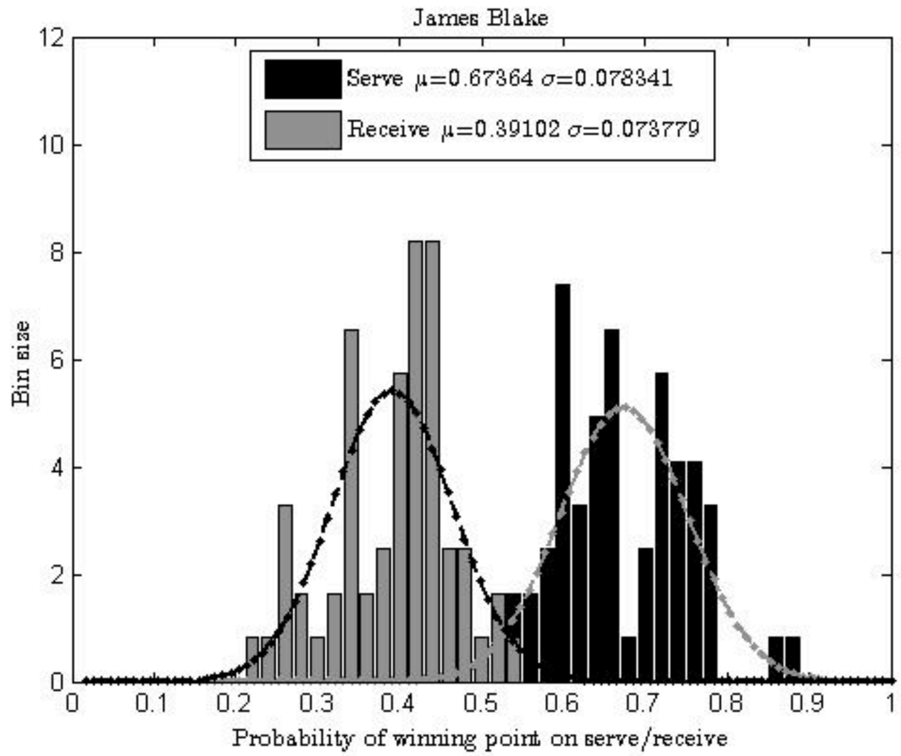

Figure 9: *Data for James Blake on serve and receive of serve for full 2007 season. Shown are histograms and Gaussian distributions using sample means and standard deviations.*

We then calculate p_A^T by using the solution of (12)–(14) in (11). This allows us to calculate p_A^S by using the solution of (7)–(9), and (10), with the result for p_A^T, in (6). Finally, using the formulas obtained for p_A^S and p_B^S, we obtain player A's probability of winning his match against player B in the three out of five set format:

$$p_A^M = (p_A^S)^3 + 3(p_A^S)^3 p_B^S + 6(p_A^S)^3 (p_B^S)^2. \tag{15}$$

More details along with all the solutions of the recursion formulas can be found in Newton & Keller (2005).

2.3 Probability density functions

In the context of the current model, p_A should be interpreted as the sample mean μ_A^s associated with the truncated Gaussian distribution (2) for winning

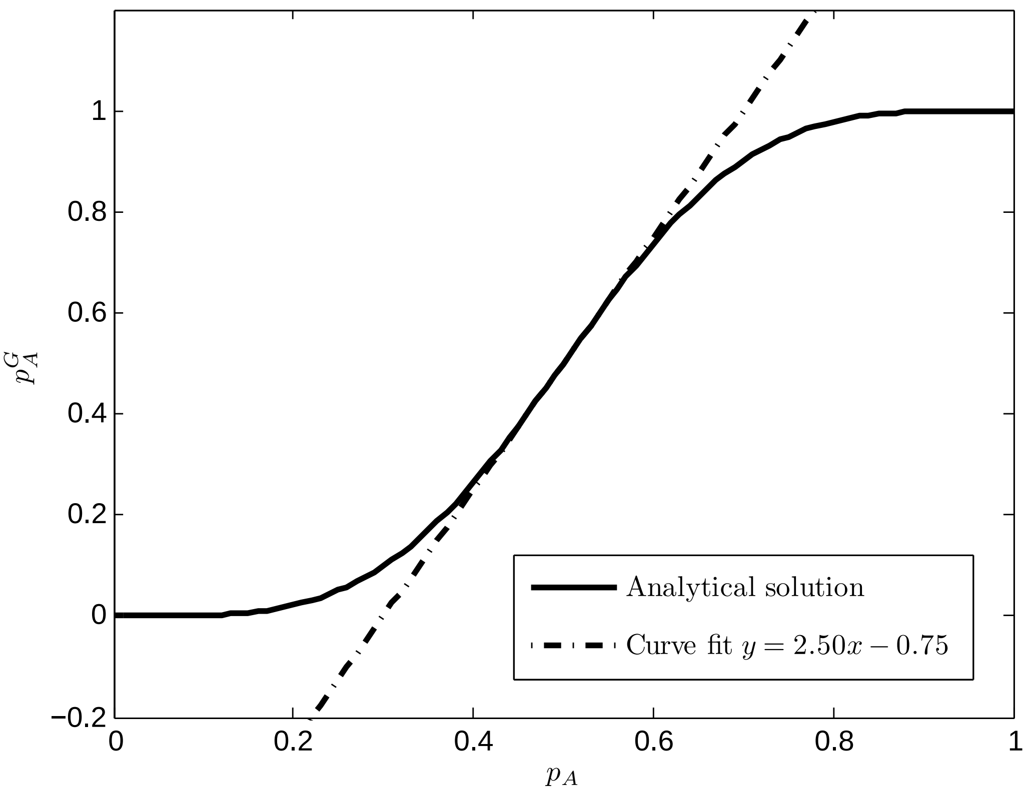

Figure 10: *The functional relation p_A^G (probability of winning a game on serve) vs. p_A (probability of winning a point on serve) as given in formula (5) which links points to games, together with the tangent line approximation taken at value $p_A = 0.5$.*

points on serve, while p_A^G should be interpreted as the sample mean associated with a player's probability of winning a game on serve. The two are related to each other via the equation (5) which we plot in figure 10, together with the tangent line approximation to the curve at the sample mean for the field $p_A = 0.50$. Note that in this region, the graph is highly linear. Because of this, the probability density function governing games won on serve must be linearly related (i.e. proportional) to the pdf governing points won on serve (in the approximate range $0.3 < p_A < 0.7$). Since $p_A^s(x)$ is taken to be Gaussian distributed, so must $p_A^G(x)$. To obtain the pdf $p_A^G(x)$, we run a Monte Carlo simulation based on an ensemble of $10,000$ matches between two players with Gaussian distributed input density functions $p_A^s(x)$, $p_B^s(x)$. Figures 11, 12 and 13 show the results of a Monte Carlo simulation between two players with Gaussian distributed inputs. Figure 11 shows a typical convergence plot (log-

log scale) using inputs $\mu_A^s = 0.63, \sigma_A^s = 0$ for player A, and $\mu_B^s = 0.65, \sigma_B^s = 0$ for player B. As is typical for all parameter values we have run, convergence to the sample mean has power-law form. Figures 12, 13 show the probability density function $p_A^G(x)$ for a player to win a game, using input values for $p_A^s(x)$ to be $\mu_A^s = 0.63316$, with $\sigma_A^2 = 0.094779$ (figure 12) and $\mu_A = 0.8$, with $\sigma_A^2 = 0.094779$ (figure 13). The mean values, in going from points to games, shifts as predicted by the curve shown in figure 10. In particular, for $\mu_A = 0.63316 = p_A$, we obtain $\mu = 0.74367 = p_A^G$, while for $\mu_A = 0.8 = p_A$, we obtain $\mu = 0.92793 = p_A^G$. Our main conclusion is that the distributions for $p_A^G(x)$ remain (approximately) normally distributed throughout a wide range of values.

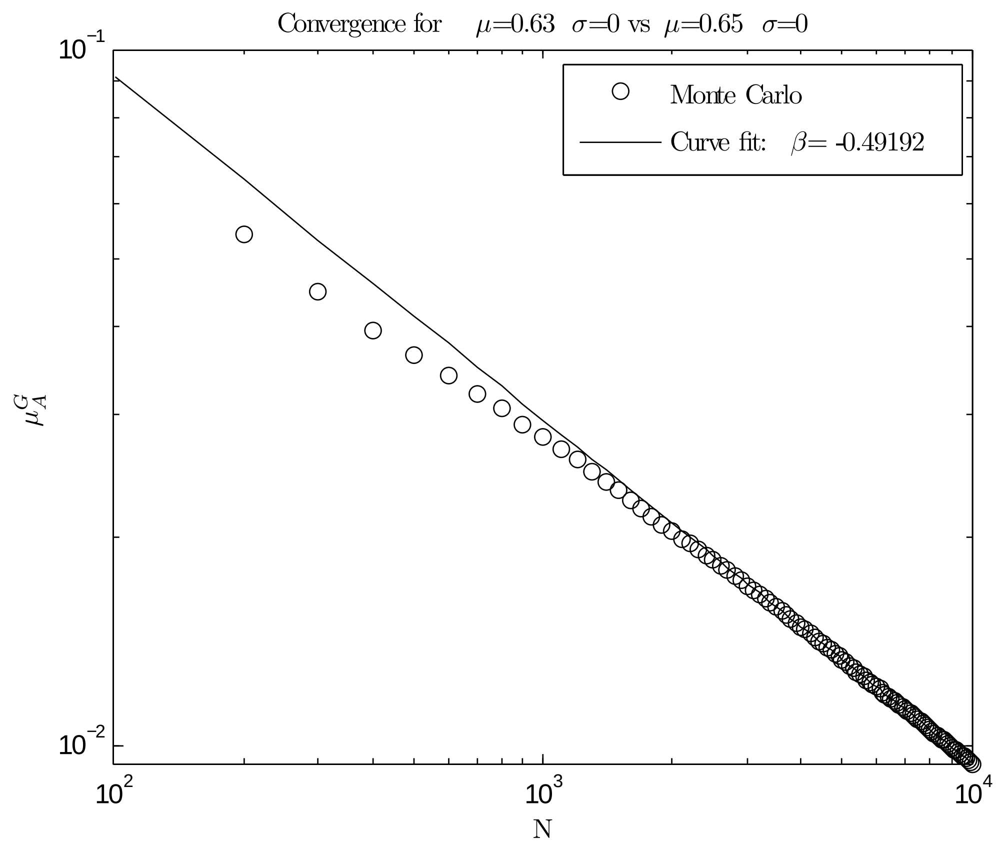

Figure 11: *Convergence plot (log-log) showing μ_A^G (sample mean associated with player A's probability of winning a game on serve) vs. N (number of matches simulated). The input parameters for the simulation are $\mu_A = 0.63$, $\sigma_A = 0$, $\mu_B = 0.65$, $\sigma_B = 0.06$. Data shows that the convergence is of power-law form $(\mu_A^G(N) - \mu_A^G(\infty)) \sim N^{-\beta}, \beta \approx 0.49192$.*

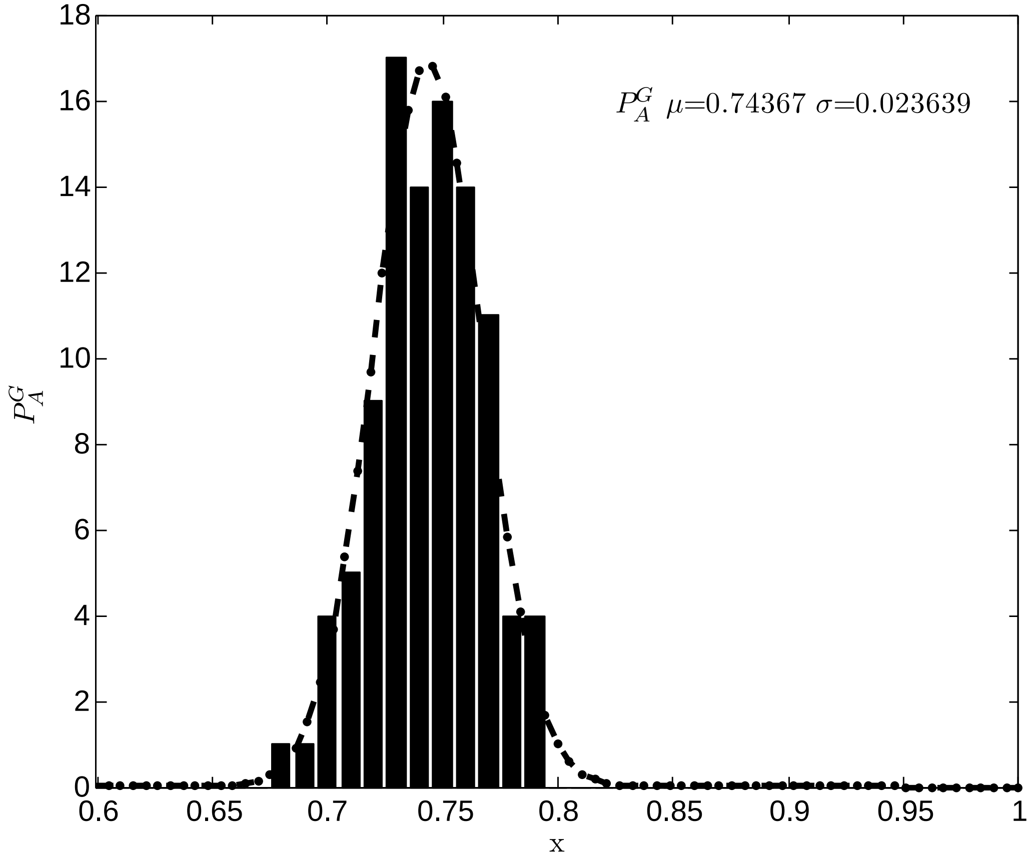

Figure 12: *Probability density function* $p_A^G(x)$ *obtained from Monte Carlo simulations of* $300,000$ *games. Density functions show clear Gaussian form with sample means predicted by the curve shown in figure 10.* $\mu_A = 0.63316$*, with* $\sigma_A^2 = 0.094779$*.*

2.4 Consistency

A player's consistency from match to match is measured by his standard deviations on points won on serve and return of serve. To see the effect of varying this parameter, we show two plots in figures 14 and 15. In figure 14, player A's mean value for points won on serve is taken to be $\mu_A = 0.71$, while his standard deviation varies $0 \leq \sigma_A \leq 0.1$. Player B is the weaker player, with mean value for points won on serve taken as $\mu_B = 0.68$. We take his standard deviation to be $\sigma_B = 0$, hence he is the more consistent player. The figure shows that as the standard deviation for player A increases, his probability of winning the match decreases. Thus, for the better server (i.e. one with higher mean value),

DOI: 10.2202/1559-0410.1169

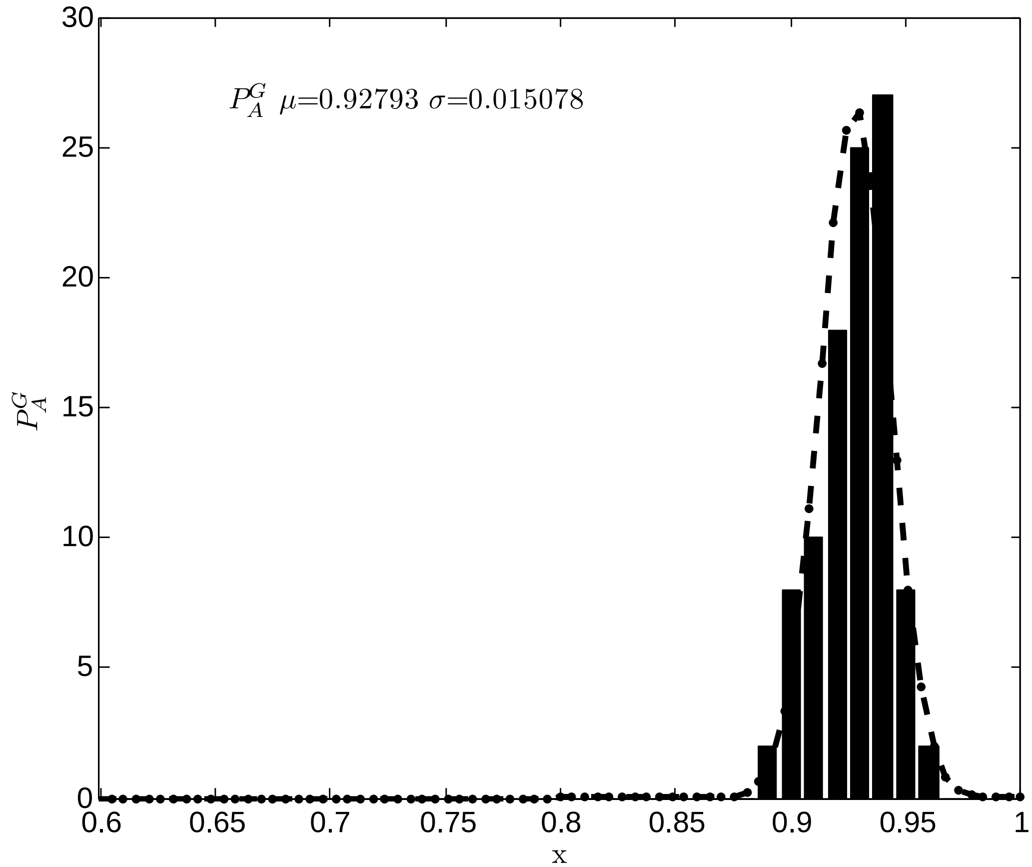

Figure 13: *Probability density function $p_A^G(x)$ obtained from Monte Carlo simulations of* $300,000$ *games. Density functions show clear Gaussian form with sample means predicted by the curve shown in figure 10.* $\mu_A = 0.8$*, with* $\sigma_A^2 = 0.094779$.

lack of consistency (higher standard deviation) hurts his chances of winning the match. In figure 15, we choose parameter values $\mu_A = 0.68$ for player A as we vary his standard deviation: $0 \leq \sigma_A \leq 0.1$. For player B, we take values $\mu_B = 0.71$ with $\sigma_B = 0$. Here, as player A's standard deviation increases, so do his chances of winning the match. Interestingly, for the weaker server (the one with lower mean value for points won on serve), lack of consistency actually *increases* his chance of winning the match. Thus, a player's standard deviation for points won on serve certainly effects his chances of winning or losing a match.

A second important and interesting point regarding a player's standard deviation is shown clearly in figures 12 and 13. The standard deviations associated with the pdf's at the point level *decrease* significantly at the game level. Thus, a player's lack of consistency in winning points on serve as modeled by his standard deviation in $p_A(x)$ becomes less significant in the pdf $p_A^G(x)$. Finally, we point out the connection between the size of the variance of an outcome and the length of the contest. It has been pointed out (see Newton and Keller (2005) and O'Malley (2008)) that upsets occur less often in best out of five set contests than in best out of three set contests. Shorter contests have higher variability and thus more chance for upsets. This is also true in other sports, where upsets are more frequent in individual games than in playoff series. We mention other recent studies of the effects of stochastic variances in paired comparison models, such as the work of Glickman (1999, 2001).

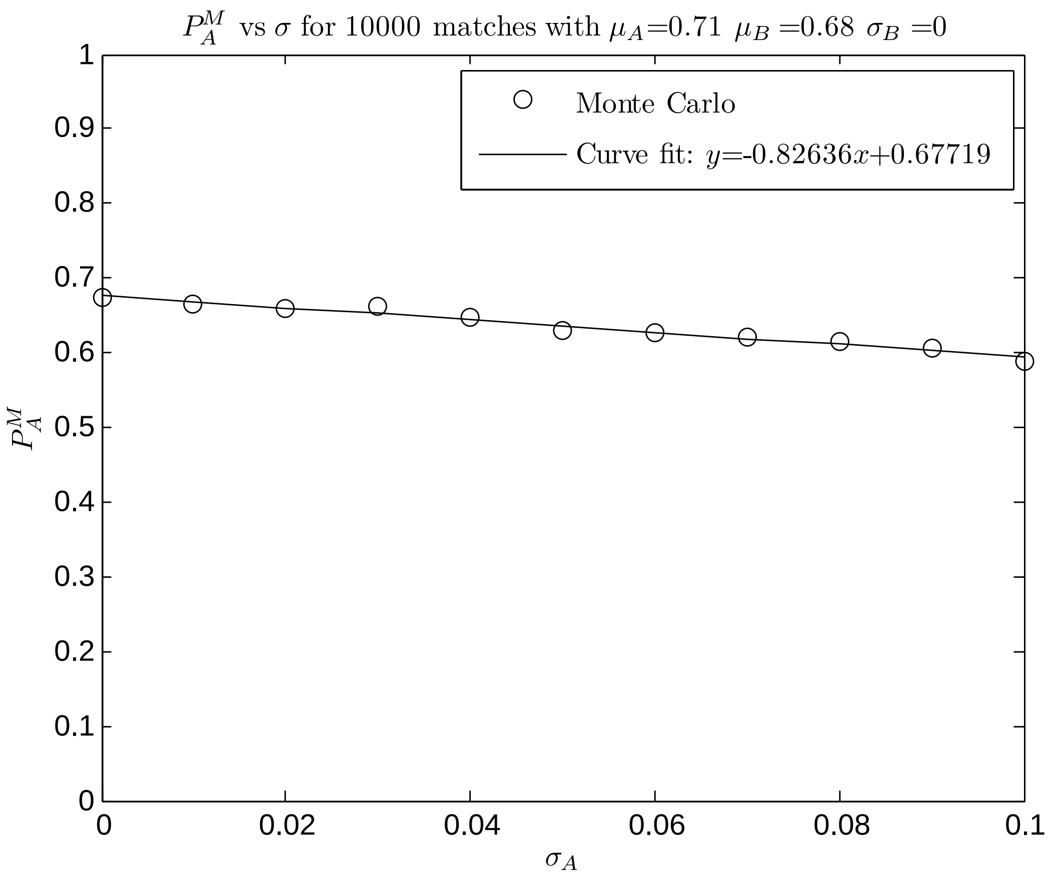

Figure 14: *The effect of varying a player's standard deviation as a measure of his consistency from match to match.* $\mu_A = 0.71$, $0 \leq \sigma_A \leq 0.1$; $\mu_B = 0.68$, $\sigma_B = 0$.

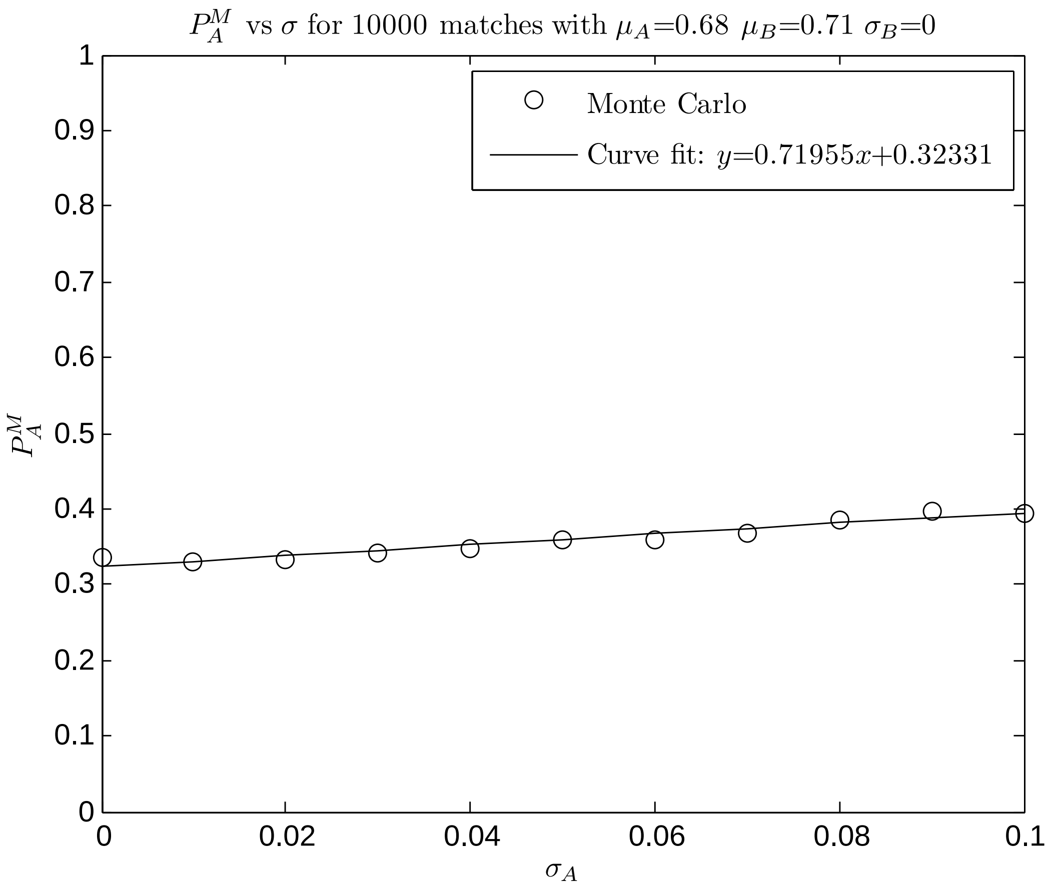

Figure 15: *The effect of varying a player's standard deviation as a measure of his consistency from match to match.* $\mu_A = 0.68$, $0 \leq \sigma_A \leq 0.1$; $\mu_B = 0.71$, $\sigma_B = 0$.

2.5 Implementation of the full model

We now describe how the full model is implemented. In particular, we first describe how each player's Gaussian distributed probability of winning a point on receiving serve is used in the context of the Markov model. In a match between player A and player B, A either wins a point on serve (with probability $p^s_{A|B}$), or he loses a point on serve, in which case his opponent wins the point on return of serve (with probability $p^r_{B|A}$). Therefore, we have:

$$p^s_{A|B} + p^r_{B|A} = 1, \quad (16)$$
$$p^r_{A|B} + p^s_{B|A} = 1. \quad (17)$$

Also, since the 'field' effectively plays itself, we have:

$$p_f^s + p_f^r = 1. \tag{18}$$

From these equations it is clear that there is a balance between a player's probability of winning a point on serve and his opponents probability of winning a point on return of serve. An increase in one comes at the expense of the other. When running a Monte Carlo simulation in a head-to-head matchup between players A and B, we do not use direct head-to-head match data (as typically there is not enough available), but we use each of their results measured against the field. For this, we have:

$$\begin{aligned} p_{A|f}^s + p_{f|A}^r &= 1, & (19) \\ p_{A|f}^r + p_{f|A}^s &= 1, & (20) \end{aligned}$$

and

$$\begin{aligned} p_{B|f}^s + p_{f|B}^r &= 1, & (21) \\ p_{B|f}^r + p_{f|B}^s &= 1. & (22) \end{aligned}$$

Adding (19) with (22) and (20) with (21) and re-arranging gives:

$$\begin{aligned} p_{A|f}^s + p_{B|f}^r &= 2 - (p_{f|A}^r + p_{f|B}^s), & (23) \\ p_{A|f}^r + p_{B|f}^s &= 2 - (p_{f|A}^r + p_{f|B}^s). & (24) \end{aligned}$$

Then, we subtract off (18) from (23), (24) which gives

$$\begin{aligned} p_{A|f}^s + (p_{B|f}^r - p_f^r) &= 1 - (p_{f|A}^r + p_{f|B}^s - p_f^s) \equiv \alpha, & (25) \\ p_{B|f}^s + (p_{A|f}^r - p_f^r) &= 1 - (p_{f|A}^r + p_{f|B}^s - p_f^s) \equiv \alpha. & (26) \end{aligned}$$

The second terms on the left, $(p_{B|f}^r - p_f^r)$ and $(p_{A|f}^r - p_f^r)$ are called 'field-adjusted' variables. They measure the deviation of a player's return of serve ability from that of the field. We will use the 'tilde' notation to denote field adjusted variables. Hence

$$\begin{aligned} \tilde{p}_{A|f}^r &\equiv p_{A|f}^r - p_f^r & (27) \\ \tilde{p}_{B|f}^r &\equiv p_{B|f}^r - p_f^r. & (28) \end{aligned}$$

Then, eqns (25), (26) become

$$\begin{aligned} p_{A|f}^s + \tilde{p}_{B|f}^r &= \alpha, & (29) \\ p_{B|f}^s + \tilde{p}_{A|f}^r &= \alpha. & (30) \end{aligned}$$

In a head-to-head simulation between players A and B, eqn (29) tells us that to include the influence of player B's ability to receive serve, as measured by the 'field adjusted' term $\tilde{p}^r_{B|f}$, we simply adjust the serve parameter for player A, $p^s_{A|f}$, so that the sum, as shown in (29), stays constant. An increase in one comes at the expense of the other. Likewise, to include the influence of player A's ability to receive serve, as measured by the 'field adjusted' term $\tilde{p}^r_{A|f}$, we adjust the serve parameter for player B, $p^s_{B|f}$, so that the sum (30) stays constant. *In short, we account for a player's ability to win a point on return of serve by adjusting his opponent's probability of winning a point on serve, either up or down depending on whether the return of serve ability is better or worse than that of the field.*

As an example, in running a simulation between Roger Federer and Rafael Nadal, Federer's sample mean (measured against the field) for points won on serve is $\mu^s_{RF} = 0.70714$, while his sample mean for points won on receive of serve is $\mu^r_{RF} = 0.41289$. For Rafael Nadal, those values are $\mu^s_{RN} = 0.6877$ and $\mu^r_{RN} = 0.42943$. The field mean for return of serve is $\mu^r_f = 0.36684$. We first calculate the difference between each players return of serve value with that of the field (i.e. the 'field adjusted' value) which we denote with a 'tilde', hence:

$$\tilde{\mu}^r_{RF} \equiv \mu^r_{RF} - \mu^r_f = 0.41289 - 0.36684 = .04605 \tag{31}$$
$$\tilde{\mu}^r_{RN} \equiv \mu^r_{RN} - \mu^r_f = 0.42943 - 0.36684 = .06259. \tag{32}$$

Then, we use 'player adjusted' values for points won on serve by taking:

$$\tilde{\mu}^s_{RF} = \mu^s_{RF} - \tilde{\mu}^r_{RN} = 0.70714 - .06259 = 0.64455 \tag{33}$$
$$\tilde{\mu}^s_{RN} = \mu^s_{RN} - \tilde{\mu}^r_{RF} = 0.6877 - .04605 = .64165. \tag{34}$$

It is these 'player adjusted' values, $\tilde{\mu}^s_{RF}$, $\tilde{\mu}^s_{RN}$ that we actually use as inputs to the truncated Gaussian distributions governing each player's probability distribution of winning a point on serve. Thus, each player's mean probability of winning a point on serve is adjusted depending on how strong his opponents return of serve ability is. A player will win fewer points on serve playing against an opponent with a stronger return of serve than one with a weaker return of serve. In this way, return of serve ability, while playing an important role in our model, manifests itself only by adjusting each players probability of winning a point on his own serve.

Based on these player adjusted values as inputs to the truncated Gaussian distributions, one realization of a statistical simulation between player A and player B proceeds by:

Player	μ^s	σ^s	μ^r	σ^r
J. Blake (13)	0.67364	0.078341	0.39102	0.073779
R. Federer (1)	0.70714	0.072314	0.41289	0.070689
R. Nadal (2)	0.6877	0.083207	0.42943	0.10075
A. Roddick (6)	0.73089	0.084164	0.34411	0.080516

Table 2: *Player data compiled for the full 2007 ATP season. Year end ranking is shown in parenthesis. Note that Roddick has the highest percentage of points won on serve, while Nadal has the highest percentage of points won on receive of serve, yet neither ended the season with the top ranking.*

(i) Choosing a value for p_A and p_B for use in the Markov chain formulas described in §2(b). This is done by drawing a random number with probability density function given by the truncated Gaussian distribution appropriate to each player, with parameters suitably adjusted for the head-to-head encounter;

(ii) Calculating each player's probability of winning a game, set, and match by solving the Markov chain formulas to obtain p_A^M, and p_B^M.

Then, to obtain a statistical ensemble, we:

(iii) Repeat these two steps $30,000$ times (choosing a new random value for each player, for each simulation) to obtain a statistical ensemble from which we obtain the probability density functions shown, for example, in figures 16 - 21. Sample means and standard deviations are then calculated for each of the ensembles. Results from head-to-head simulations between players and full tournament simulations are described in the next section.

3 Monte Carlo simulations

In this section we describe the results of our Monte Carlo simulations of head-to-head matches between Roger Federer, Rafael Nadal, Andy Roddick, and James Blake, using as input, their player data from the 2007 season, as shown in figures 6 - 9 and summarized in Table 2. We also describe the outcome of full tournament simulations of each of the four Grand Slam events using two different methods for obtaining ensembles.

3.1 Head-to-head matches

As shown in Table 2, Federer does not have the highest mean value for points won on serve, or the highest mean value for points won on receive of serve,

DOI: 10.2202/1559-0410.1169

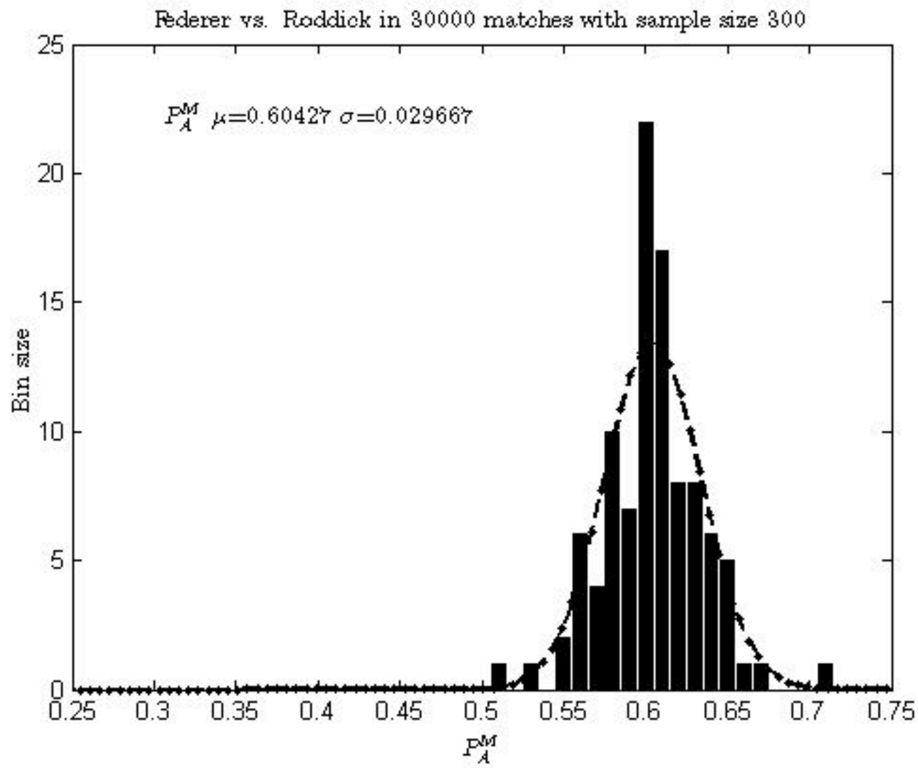

Figure 16: *Histogram and Gaussian distribution of head-to-head match-ups based on 30,000 simulated matches using the Monte Carlo model. Bin size is 300 matches. The sample mean associated with Federer's probability of winning vs. Roddick is* $\mu_{RF|AR} = 0.60733$, *with variance* $\sigma_{RF|AR} = 0.027796$.

but he finished the 2007 season with the best record and the top ranking. We should point out that his standard deviation both on serve and receiving serve was the lowest of the four players. These parameters were used in Monte Carlo simulations of $30,000$ matches between each pair of these four players. The results are histogramed in figures 16 - 21 and the sample means and standard deviations are used to define the probability density functions for each player's probability of winning a match. The output values are shown in Table 3. The results yield important information about how close the players are to each other. For example, in head-to-head matches between Federer and Nadal, Federer is predicted to win a slim majority of 50.74% of their matches.

Since our head-to-head simulation method relies entirely on data for each of the players taken against the 'field', one might wonder whether using actual

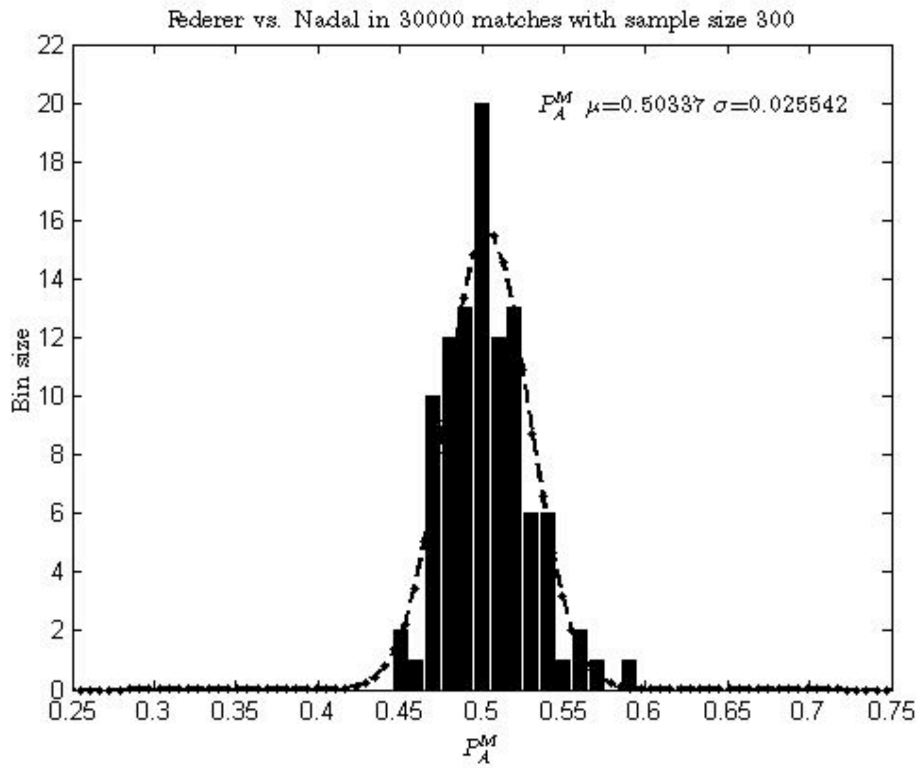

Figure 17: *Histogram and Gaussian distribution of head-to-head match-ups based on 30,000 simulated matches using the Monte Carlo model. Bin size is 300 matches. The sample mean associated with Federer's probability of winning vs. Nadal is* $\mu_{RF|RN} = 0.5074$*, with variance* $\sigma_{RF|RN} = 0.02764$*.*

head-to-head data might be useful as input to a model. We point out that in any given year, the number of actual head-to-head matches between any two players is very small, making it very difficult to use for statistical purposes. For example, Federer defeated Nadal in three out of their five matches in 2007, consistent with his slight statistical edge shown in our simulations but not in enough matches to gather meaningful statistical conclusions. Federer also defeated Blake in their single 2007 match and he defeated Roddick in all three of their head-to-head matches. Nadal defeated Roddick in their single 2007 match, while Nadal never played Blake, and Roddick and Blake never played. All of these outcomes are consistent with our statistical findings, however the sparsity of head-to-head meetings between any two players on the tour in a given year makes it difficult to use this as input data to a model. Taking

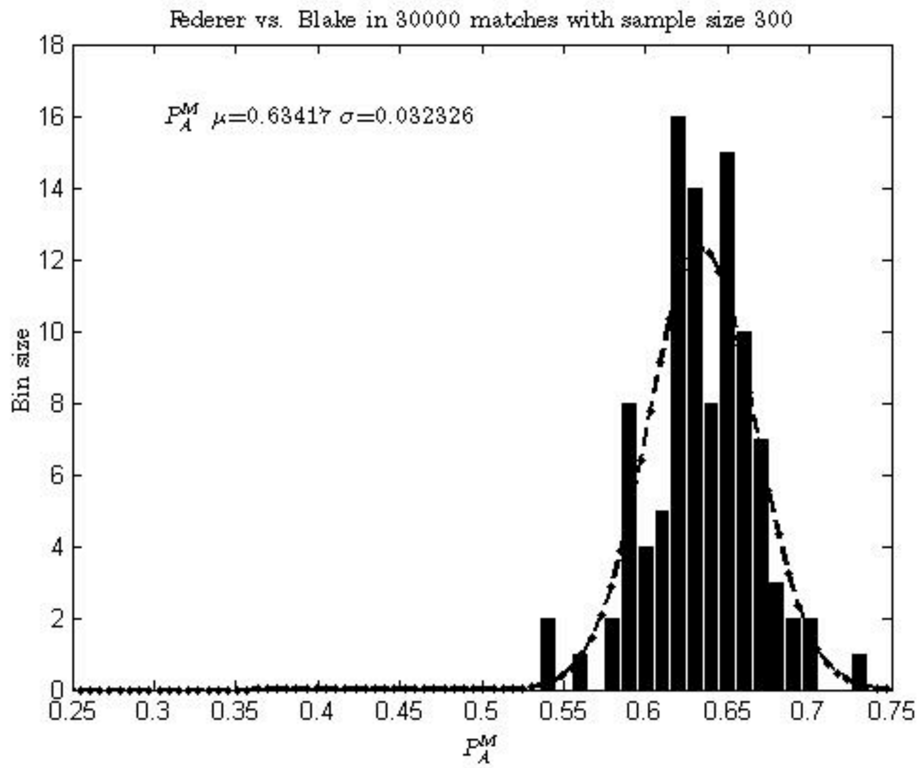

Figure 18: *Histogram and Gaussian distribution of head-to-head match-ups based on 30,000 simulated matches using the Monte Carlo model. Bin size is 300 matches. The sample mean associated with Federer's probability of winning vs. Blake is* $\mu_{RF|JB} = 0.6365$*, with variance* $\sigma_{RF|JB} = 0.03043$*.*

data from head-to-head encounters over several years or a full career would typically increase the amount of data, but would introduce other troublesome questions. For example, it is doubtful that data from matches between Federer and Nadal in 2005, or even 2006 would be useful in predicting outcomes in 2007, as Nadal was in the process of rapidly improving (and altering) his game at that time. Even between two players whose career timelines match, such as Agassi and Sampras, it is hard to argue that data taken from the early stages of their career head-to-head encounters would help in predicting who would win their final encounter at the US Open Finals in 2002. However, it must also be recognized that our use of individual player data taken against the field in the evaluation of head-to-head encounters possibly underestimates the significance of one player having a statistically unusually high (or low) success

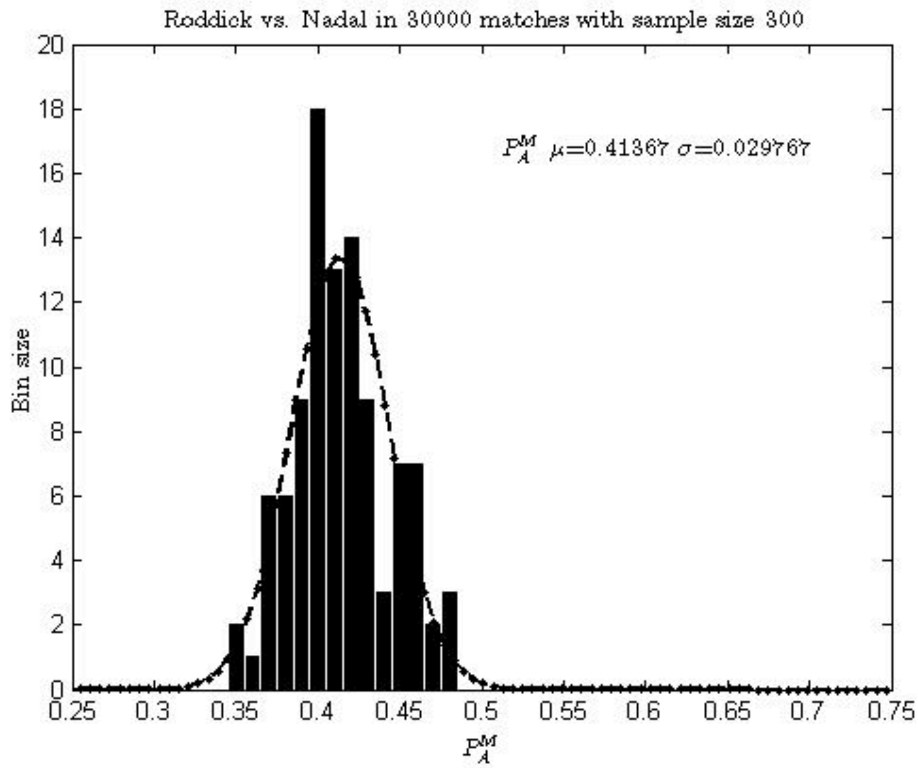

Figure 19: *Histogram and Gaussian distribution of head-to-head match-ups based on 30,000 simulated matches using the Monte Carlo model. Bin size is 300 matches. The sample mean associated with Roddick's probability of winning vs. Nadal is* $\mu_{AR|RN} = 0.4117$*, with variance* $\sigma_{AR|RN} = 0.027895$.

rate against another individual player (a 'bogey' opponent).

3.2 Tournament simulations

Using the full 2007 tournament data for each of the 128 players in the four Grand Slam events, we carried out Monte Carlo simulations of each tournament. The ensembles were gathered in two ways. First, we ran 1000 fictitious tournaments initializing each realization by using the actual first-round match-ups from the event. We call these 'fixed', or 'actual' draws. Then, for comparison, we ran 1000 simulations of each event using random draws chosen in the first round of each realization. The comparison of the two gives valuable insights into the effect of the actual tournament (i.e. player seedings) draw on

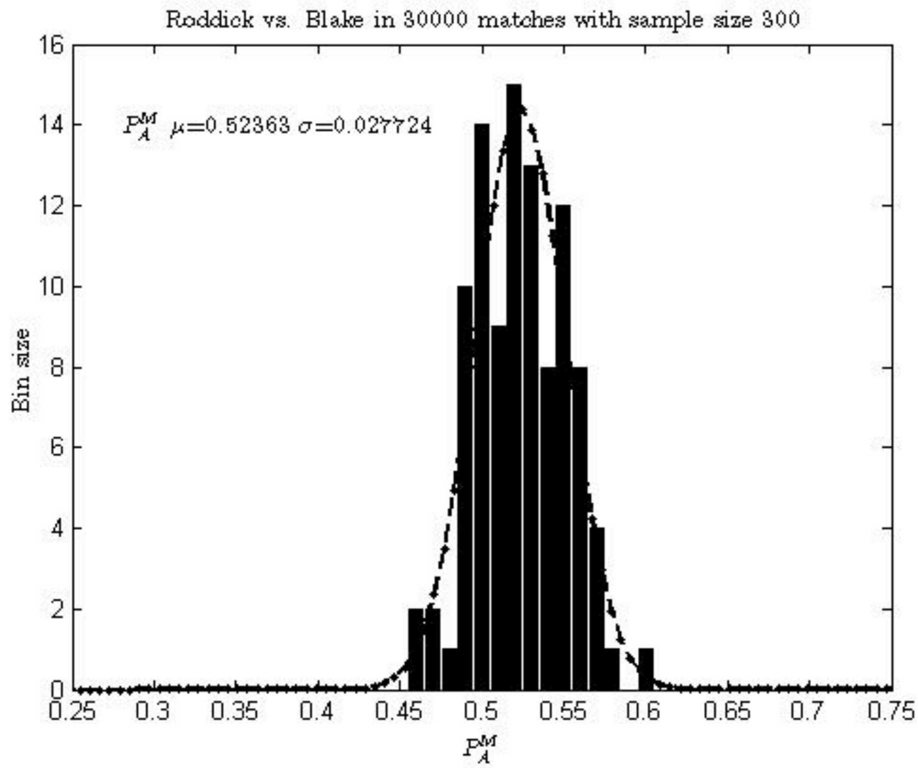

Figure 20: *Histogram and Gaussian distribution of head-to-head match-ups based on 30,000 simulated matches using the Monte Carlo model. Bin size is 300 matches. The sample mean associated with Roddick's probability of winning vs. Blake is* $\mu_{AR|JB} = 0.5206$*, with variance* $\sigma_{AR|JB} = 0.027607$.

outcomes.

The results from the simulations were histogramed for each tournament, showing the number of rounds won by each player in the ensemble of 1000 simulated tournaments. Table 4 shows the number of tournament wins for each of the four players out of 1000 tournament simulations using fixed draws. Table 5 shows the same, but using random draws. The full histograms showing round-by-round statistics for the four players for Wimbledon and the US Open are shown in figures 22 - 33. A number of points are worth making. Generally speaking, the stronger players do better in the actual draw than in random draws. This is evidenced by the fact that the bins from the histograms for Federer and Nadal for the random draws decrease in height as the rounds increase (except for the finals), while those from the actual draws do not.

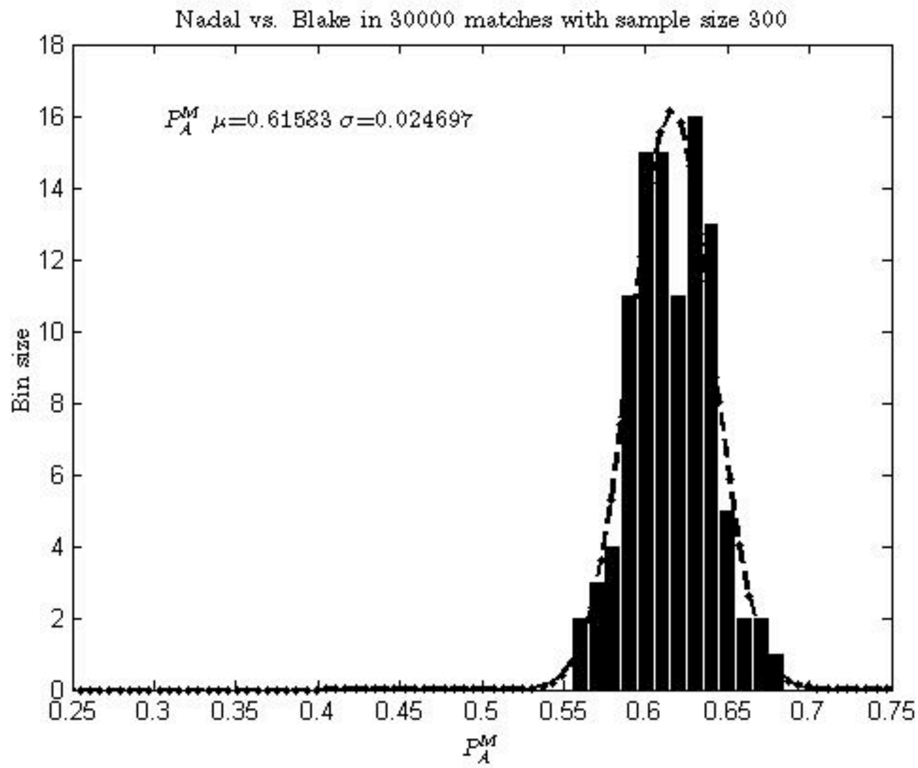

Figure 21: *Histogram and Gaussian distribution of head-to-head match-ups based on 30,000 simulated matches using the Monte Carlo model. Bin size is 300 matches. The sample mean associated with Nadal's probability of winning vs. Blake is* $\mu_{RN|JB} = 0.61307$*, with variance* $\sigma_{RN|JB} = 0.029254$*.*

These two (stronger) players have a better chance of surviving deep into the tournament with the actual draw, as designed by the seeding committee. These two players also win more of the tournaments (summarized in Tables 4 and 5) in the actual draw then in the random draws. This is not true of weaker players who fair better in a random draw then in the actual draw where they are forced to play top players in the early rounds. When comparing the number of times each of the four players actually won each of the four Grand Slam events in our simulation, Federer was the most successful, winning 115, 120, 122, and 94 of each of the events, which in probabilistic terms translates into winning 11.5% of the Australian Open simulations, 12% of the French Open simulations, 12.2% of the Wimbledon simulations, and 9.4% of the US Open simulations. Although these numbers were higher than for any other player

	Blake	Federer	Nadal	Roddick
Blake	NA	$\mu = 0.3635$ $\sigma = 0.0304$	$\mu = 0.3869$ $\sigma = 0.02925$	$\mu = 0.4794$ $\sigma = 0.02761$
Federer	$\mu = 0.6365$ $\sigma = 0.03043$	NA	$\mu = 0.5074$ $\sigma = 0.0276$	$\mu = 0.6073$ $\sigma = 0.0278$
Nadal	$\mu = 0.61307$ $\sigma = 0.02925$	$\mu = 0.4926$ $\sigma = 0.0276$	NA	$\mu = 0.5883$ $\sigma = 0.0279$
Roddick	$\mu = 0.5206$ $\sigma = 0.02761$	$\mu = 0.3927$ $\sigma = 0.0278$	$\mu = 0.4117$ $\sigma = 0.0279$	NA

Table 3: *Resulting means and standard deviations from the Monte Carlo simulations of head-to-head match ups. Based on this input, the matrix based ranking system ordered the four players: 1. Federer; 2. Nadal; 3. Roddick; 4. Blake.*

Player	Australian	French	Wimbledon	US Open
J. Blake	27	22	28	20
R. Federer	115	120	122	94
R. Nadal	79	93	75	105
A. Roddick	27	24	39	38

Table 4: *Number of tournament wins in* 1000 *tournament simulations for each of the Grand Slam events using the actual tournament draw in the first round.*

(except for Nadal's outcomes in the US Open simulations which showed he won 10.5% of those events), they are perhaps surprisingly low given the fact that Federer won 3 out of 4 of these actual events in the 2007 calendar year (Australian Open, Wimbledon, and US Open) and made it to the finals of the French Open. It is worthwhile pointing out that the statistics from Tables 4 and 5 support the notion that Federer has a larger advantage over Nadal in Wimbledon than in the French Open, based on their respective styles of play and the two different surfaces these tournaments use.

Player	Australian	French	Wimbledon	US Open
J. Blake	25	23	23	20
R. Federer	108	95	98	96
R. Nadal	70	90	67	80
A. Roddick	30	28	26	36

Table 5: *Number of tournament wins in* 1000 *tournament simulations for each of the Grand Slam events using random draws in the first round.*

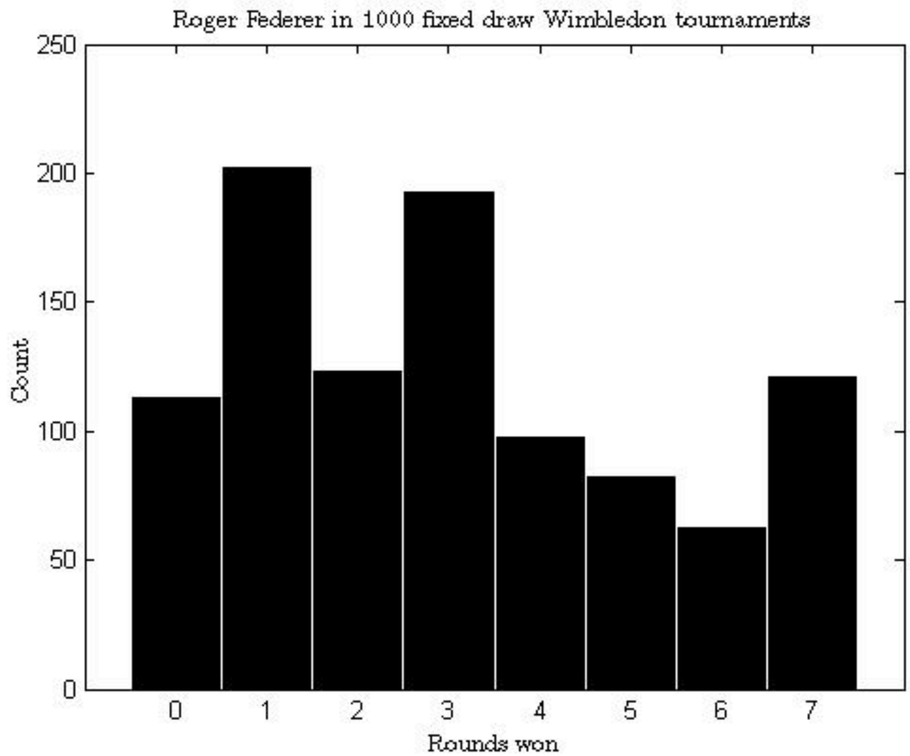

Figure 22: *Histogram of 1000 full tournament simulations for the 2007 Wimbledon tournament using the actual draw. For each tournament realization, the number of rounds won is recorded and binned for each of the players. Results are shown for Roger Federer.*

3.3 Probabilistic ranking schemes

The simulated match-ups can be used in a simple matrix based ranking system (Keener (1992)) which we now describe. Consider the 4×4 'preference' matrix constructed for the four players previously discussed:

$$A = [a_{ij}]. \tag{35}$$

Each entry of this matrix contains the sample mean associated with the probability of the (row) player defeating the (column) player, as obtained from a Monte Carlo simulation of $30,000$ head-to-head matches between the two players, implemented according to our model. We assume there is a vector of ranking values, $\vec{r}$, with positive components j indicating the strength of the

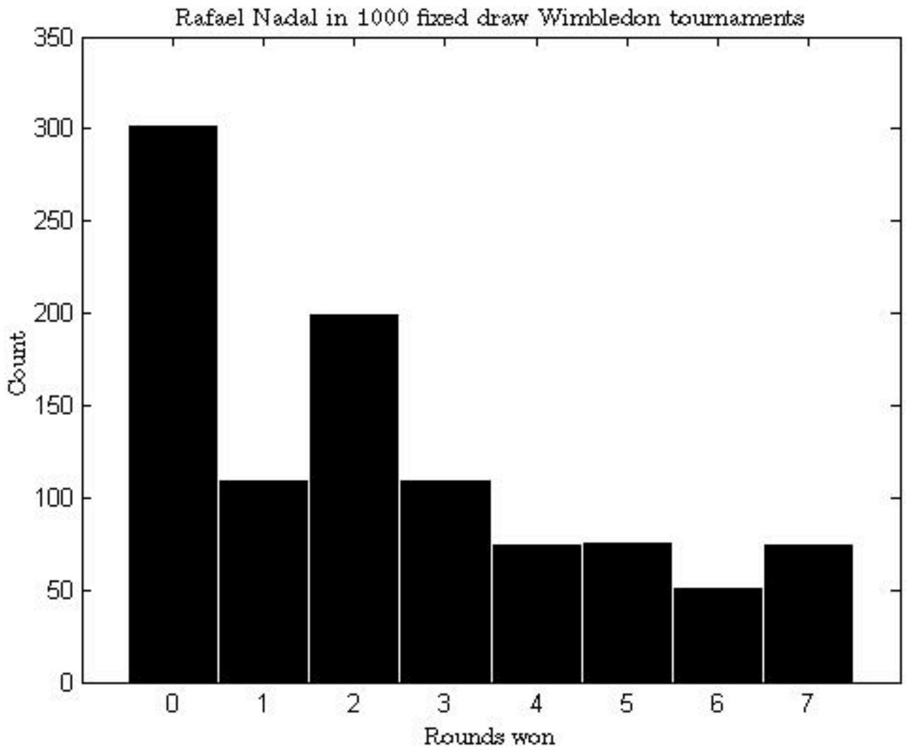

Figure 23: *Histogram of 1000 full tournament simulations for the 2007 Wimbledon tournament using the actual draw. For each tournament realization, the number of rounds won is recorded and binned for each of the players. Results are shown for Raphael Nadal.*

jth player. The normalized 'score' for player i is given by:

$$S_i = \frac{1}{N}\sum_{j=1}^{N} a_{ij} r_j, \tag{36}$$

where a_{ij} is the sample mean associated with the simulated matches between player i and j. Since S_i incorporates information about the strengths of all the players based on their simulated performance against each other, it is a good measure of the comparative strength of the players. Then, as in Keener (1992)[2], we assume that the strength of each player is proportional to their

[2]This method of ranking was first developed by J.B. Keller (personal communication), to rank teams in major league baseball. It is now the basis of Google's pagerank algorithm (see Bryan and Leise (2006)).

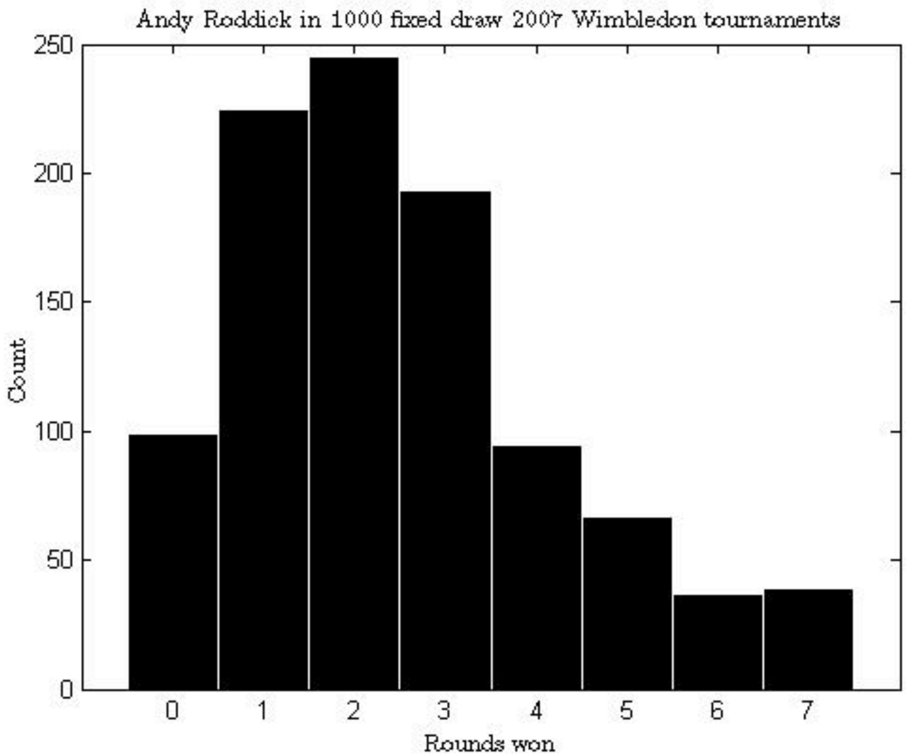

Figure 24: *Histogram of 1000 full tournament simulations for the 2007 Wimbledon tournament using the actual draw. For each tournament realization, the number of rounds won is recorded and binned for each of the players. Results are shown for Andy Roddick.*

score:

$$A\vec{r} = \lambda\vec{r}. \tag{37}$$

Thus the ranking vector is a positive eigenvector of a positive definite matrix A. Each player's score is the result of their interaction with all the other players (i.e. the field), and the assigned score depends both on the outcome of the interactions as well as the strength of the opponents. Of course, since each of the entries a_{ij} are obtained from Monte Carlo simulations with Gaussian distributed inputs obtained from the data, the ranking vector, $\vec{r}$, inherits these desirable features, giving it a natural probabilistic interpretation. Each component of the ranking vector is a random variable with mean value r_{ij}. We note that A has nonnegative entries and is irreducible, hence by the Perron-Frobenius theorem, there exists an eigenvector with nonnegative entries, cor-

DOI: 10.2202/1559-0410.1169

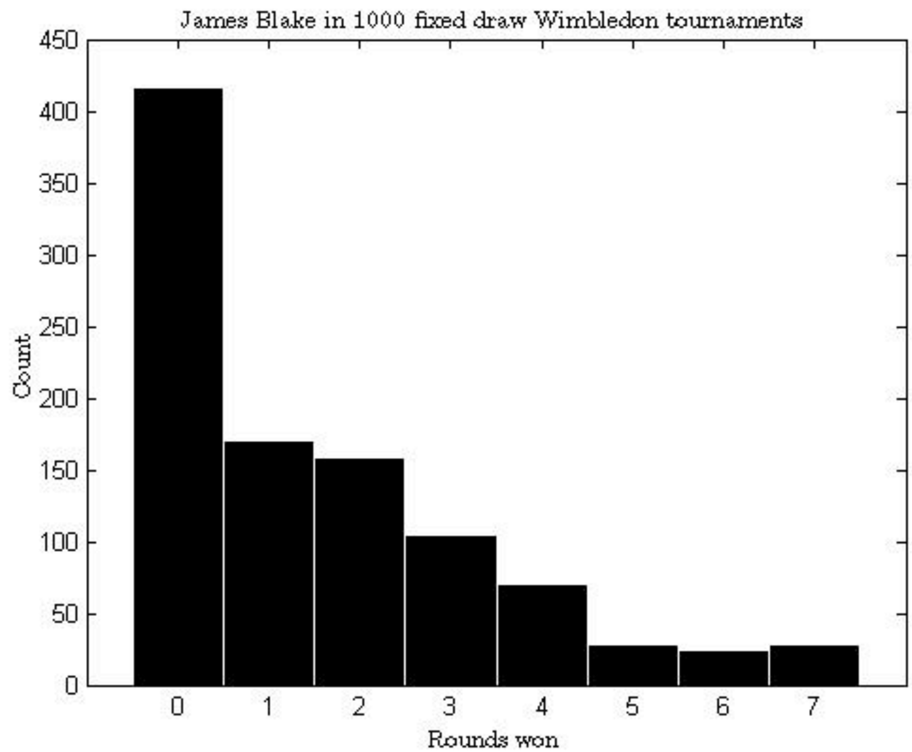

Figure 25: *Histogram of 1000 full tournament simulations for the 2007 Wimbledon tournament using the actual draw. For each tournament realization, the number of rounds won is recorded and binned for each of the players. Results are shown for James Blake.*

responding to λ which is unique and simple, with eigenvalue that is the largest eigenvalue of A (in absolute value).

Based on the data shown in Table 3, we construct the preference matrix with sample means as entries indicating the probability that the row player defeats the column player. The rows and columns are as listed in Table 3. The preference matrix A and associated eigenvector $\vec{\xi}$ is given by:

$$A = \begin{pmatrix} 0 & 0.3635 & 0.3869 & 0.4794 \\ 0.6365 & 0 & 0.5074 & 0.6073 \\ 0.6131 & 0.4926 & 0 & 0.5883 \\ 0.5206 & 0.3927 & 0.4117 & 0 \end{pmatrix}; \quad \vec{\xi} = \begin{pmatrix} 0.4298 \\ 0.5572 \\ 0.5472 \\ 0.4550 \end{pmatrix} \tag{38}$$

Our ranking produces: (4) Blake $= 0.4298$; (1) Federer $= 0.5572$; (2) Nadal $= 0.5472$; (3) Roddick $= 0.4550$. The final rank is in parentheses. Importantly,

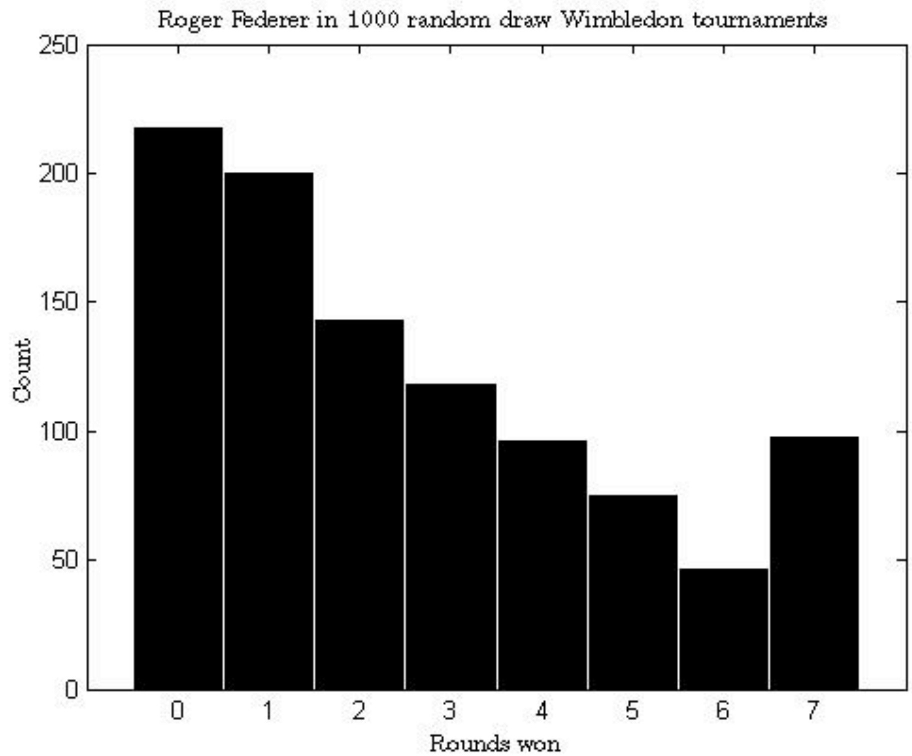

Figure 26: *Histogram of 1000 full tournament simulations for the 2007 Wimbledon tournament using a randomized draw for each tournament realization. The number of rounds won is recorded and binned for each of the players. Results are shown for Roger Federer.*

despite the fact that Federer had neither the highest percentage of points won on serve, or receive of serve (see data in Table 2), he earns the top ranking using this system. The final ordering of the four players also agrees with their ATP year-end rankings also shown in Table 2.

4 Discussion

Novel features of the methods described in this paper are the use of data based Gaussian distributed input variables in the model measuring each player's (i) strength of serve; (ii) strength of return of serve; and (iii) consistency. The data is gathered for each player over their entire portfolio of matches played in the 2007 ATP season. Using this 'field' data allows us to gather enough information on the performance of each of the individual players, despite the

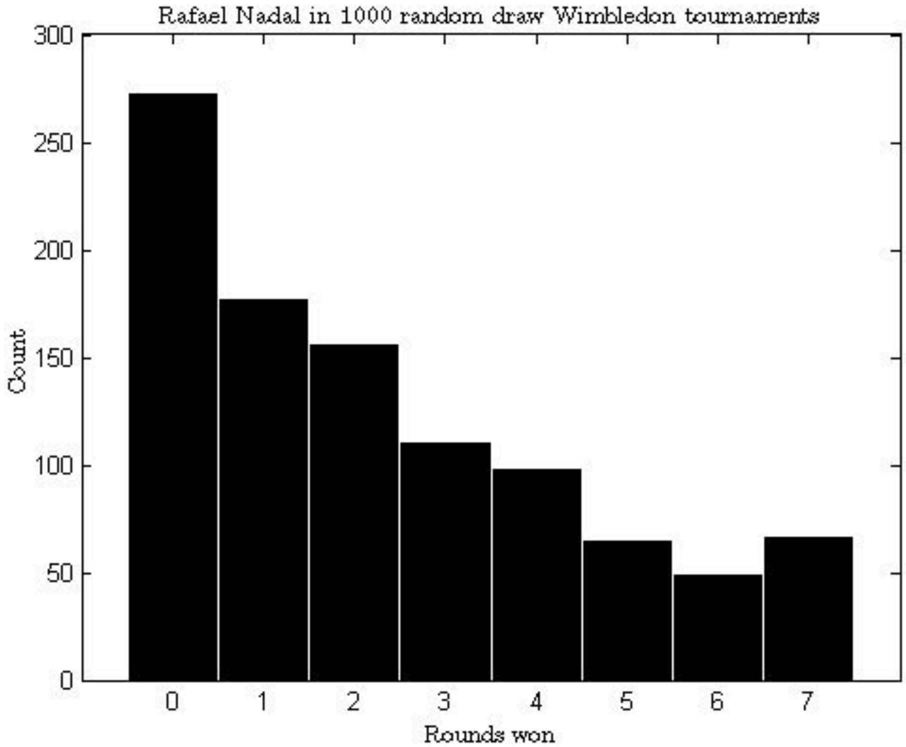

Figure 27: *Histogram of 1000 full tournament simulations for the 2007 Wimbledon tournament using a randomized draw for each tournament realization. The number of rounds won is recorded and binned for each of the players. Results are shown for Raphael Nadal.*

fact that most pairs of players never actually have head-to-head matches (or very few) on a given year. Thus, our predictions on the probability that one player will defeat another are based on how each has performed against the same control group, which in this case is the entire field of players. The calculation of probability density functions as our main outputs of the model, as opposed to single output variables, gives far more detailed and nuanced information regarding a player's ability and probability of winning a match. It also gives us the ability to carry out realistic tournament simulations and gather statistics for each player on a round-by-round basis. There are several ways of using this information in the development of probabilistic ranking schemes, one of which we show using a matrix based method. A fully developed Monte Carlo based ranking system will be described in a separate publication.

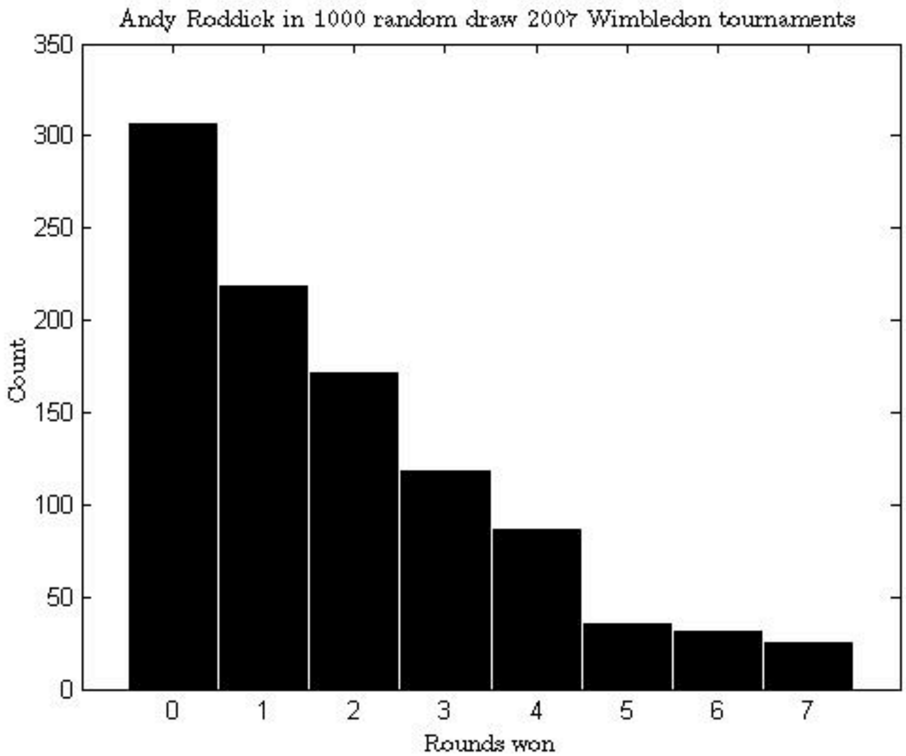

Figure 28: *Histogram of 1000 full tournament simulations for the 2007 Wimbledon tournament using a randomized draw for each tournament realization. The number of rounds won is recorded and binned for each of the players. Results are shown for Andy Roddick.*

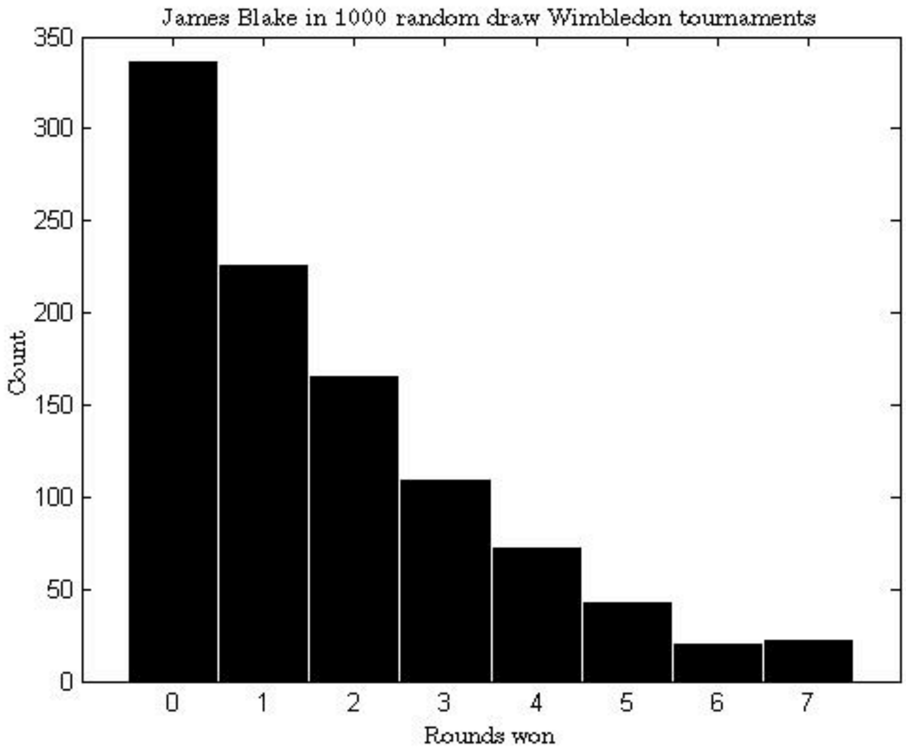

Figure 29: *Histogram of 1000 full tournament simulations for the 2007 Wimbledon tournament using a randomized draw for each tournament realization. The number of rounds won is recorded and binned for each of the players. Results are shown for James Blake.*

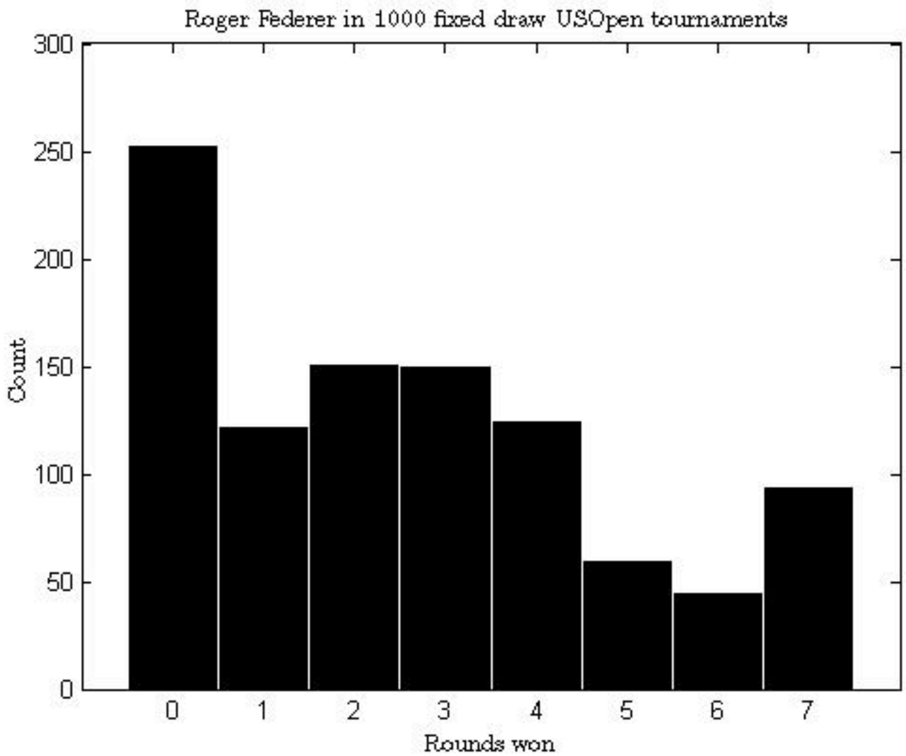

Figure 30: *Histogram of 1000 full tournament simulations for the 2007 US Open using the actual draw. For each tournament realization, the number of rounds won is recorded and binned for each of the players. Results are shown for Roger Federer.*

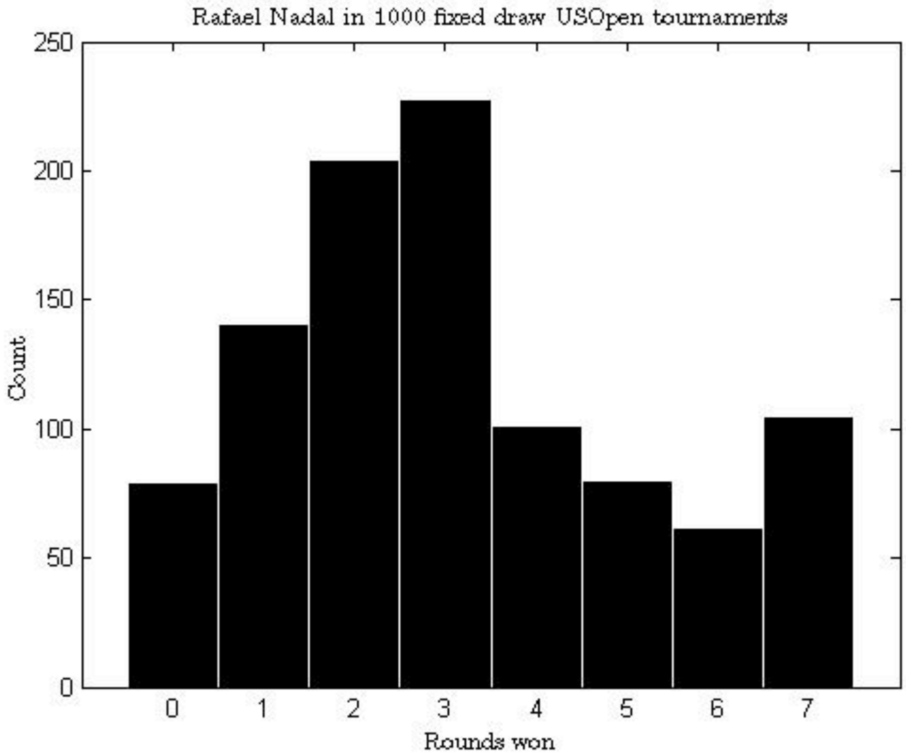

Figure 31: *Histogram of 1000 full tournament simulations for the 2007 US Open using the actual draw. For each tournament realization, the number of rounds won is recorded and binned for each of the players. Results are shown for Raphael Nadal.*

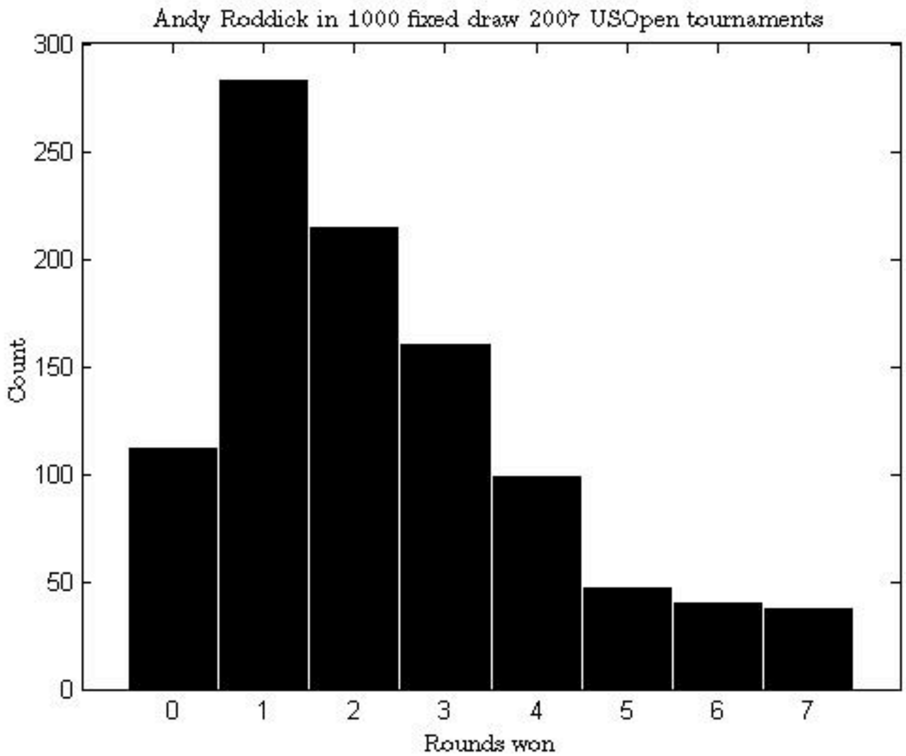

Figure 32: *Histogram of 1000 full tournament simulations for the 2007 US Open using the actual draw. For each tournament realization, the number of rounds won is recorded and binned for each of the players. Results are shown for Andy Roddick.*

DOI: 10.2202/1559-0410.1169

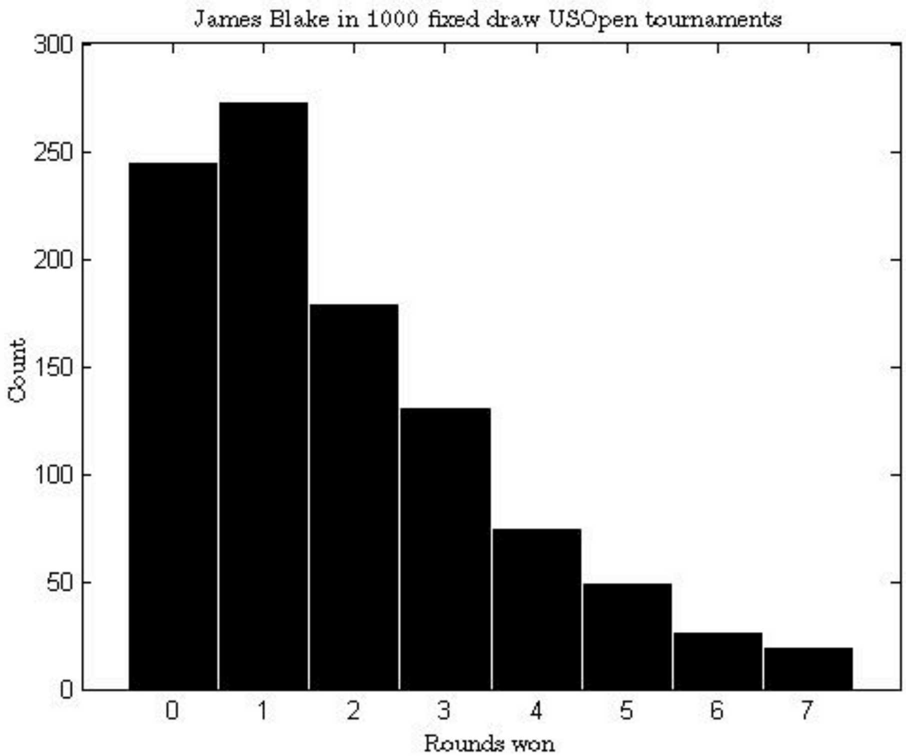

Figure 33: *Histogram of 1000 full tournament simulations for the 2007 US Open using the actual draw. For each tournament realization, the number of rounds won is recorded and binned for each of the players. Results are shown for James Blake.*

References

Asmussen, S., P.W. Glynn (2007) **Stochastic Simulation: Algorithms and Analysis**, Stochastic Modelling and Applied Probability Vol. 57, Springer-Verlag.

Barnett,T., S.R. Clarke (2005) Combining player statistics to predict outcomes of tennis matches, *IMA Journal of Management Mathematics*, **16** 113–120.

Bendat, J.S., A.G. Piersol (1986) **Random Data: Analysis and Measurement Procedures**, 2nd Ed. John Wiley & Sons.

Bryan, K., T. Leise (2006) The $25,000,000,000 Eigenvector: The linear algebra behind Google, SIAM Review, **48**(3) 569-581.

Carter, W.H. , S.L. Crews (1974) An analysis of the game of tennis, *Amer. Statistician* **28** 130–134.

Clarke, S.R., D. Dyte (2000) Using official ratings to simulate major tennis tournaments, *Intl. Trans. in Op. Res.* **7** 585–594.

Cross, R. (2008) Tennis physics, anyone?, *Physics Today* Sept. 84–85.

Glickman, M.E. (1999) Parameter estimation in large dynamic paired comparison experiments, *J. of the Royal Stat. Soc., Ser. C: Appl. Stat.* **48** 377–394.

Glickman, M.E. (2001) Dynamic paired comparison models with stochastic variances, *J. of Appl. Stat.* **28** 673–689.

Keener, J.P. (1993) The Perron-Frobenius theorem and the ranking of football teams, *SIAM Review* **35**(1) 80–93.

Morris, C.N., (1977) The most important points in tennis. In: *Optimal Strategies in Sport*, S.P. Ladany and R.E. Nichol, Eds., pp 131–140. Amsterdam; North-Holland.

Newton, P.K. , J.B. Keller (2005) The probability of winning at tennis I. Theory and data, *Studies in Applied Mathematics* **114** 241–269.

Newton, P.K., K. Aslam (2006) Monte Carlo Tennis, *SIAM Review* **48**(4) 722–742.

Newton, P.K., G.H. Pollard (2004) Service neutral scoring strategies for tennis, *Proceedings of the Seventh Australasian Conference on Mathematics and Computers in Sport*, Massey University Palmerston North, New Zealand, 221–225.

O'Malley, A. James (2008) Probability formulas and statistical analysis in tennis, *J. of Quantitative Analysis in Sports*, Vol. 4(2) Article 15.

Pollard, G.H. (1983) An analysis of classical and tie-breaker tennis, *Austral. J. Stat.* **25** 496–505.

Walker, M., J. Wooders (2001) Minimax play at Wimbledon, *The American Econ. Rev.* **91**(5) Dec. 1521–1538.