March Madness and the Office Pool

Edward H. Kaplan • Stanley J. Garstka
Yale School of Management, New Haven, Connecticut 06520-8200
edward.kaplan@yale.edu • stanley.garstka@yale.edu

March brings March Madness, the annual conclusion to the U.S. men's college basketball season with two single elimination basketball tournaments showcasing the best college teams in the country. Almost as mad is the plethora of office pools across the country where the object is to pick a priori as many game winners as possible in the tournament. More generally, the object in an office pool is to maximize total pool points, where different points are awarded for different correct winning predictions. We consider the structure of single elimination tournaments, and show how to efficiently calculate the mean and the variance of the number of correctly predicted wins (or more generally the total points earned in an office pool) for a given slate of predicted winners. We apply these results to both random and Markov tournaments. We then show how to determine optimal office pool predictions that maximize the expected number of points earned in the pool. Considering various Markov probability models for predicting game winners based on regular season performance, professional sports rankings, and Las Vegas betting odds, we compare our predictions with what actually happened in past NCAA and NIT tournaments. These models perform similarly, achieving overall prediction accuracies of about 58%, but do not surpass the simple strategy of picking the seeds when the goal is to pick as many game winners as possible. For a more sophisticated point structure, however, our models do outperform the strategy of picking the seeds.
(*March Madness; Office Pools; Probability Modeling; Statistical Estimation; Markov Models; Dynamic Programming*)

1. Introduction

The men's college basketball season in the United States culminates in March of each year with the end-of-season tournaments that comprise "March Madness." The National Collegiate Athletic Association (NCAA) places 64 of the best college teams in the country in a single elimination tournament, and the NCAA tournament winner is crowned the national champion. An additional 32 teams are invited to compete in the National Invitational Tournament (or the NIT). In recent years, women's college basketball has gained in popularity, and the NCAA also sponsors a 64-team women's tournament each spring.

Betting on NCAA basketball games is very popular, and the games that comprise March Madness rival the Super Bowl (the championship of the National Football League) in terms of the total amount of money wagered (an estimated $80 million were wagered on the 1998 "Final Four" alone (Barnhouse 1999)). Outside of formal sports betting markets, March Madness has given rise to a plethora of office pools (Atkins 1998). The goal in these pools is typically to predict the winners of as many games as possible. More sophisticated pools incorporate point schemes that award different numbers of points to correct predictions depending on which teams and games are involved. In a pure office pool, entry fees contribute to a prize that is awarded to the player with the highest point total, shared in the case of ties, or perhaps shared among first-, second-, and third-place contestants in some prespecified manner.

In pools with a small number of participants, entrants might consider not only which teams are likely winners in tournament competition, but also

0025-1909/01/4703/0369$5.00
1526-5501 electronic ISSN

how their fellow pool participants are likely to behave. In a manner akin to the avoidance of "popular numbers" in state lotteries, game-theoretic machinations could reasonably lead a player to the deliberate choice of underrated teams in an attempt to distinguish oneself from the competition. A small reduction in the probability of winning the pool might be compensated by the improved conditional winnings associated with a less popular gamble.

Worrying about the play of others in a pool becomes increasingly difficult in pools with a very large number of participants, such as those found on the Internet with tens or even hundreds of thousands of participants. Even in small pools, however, many players are likely to be preoccupied with their ability to correctly forecast tournament outcomes, or to contrast their own assessments of likely team performance against the point scheme employed by the pool to make what seem like point-winning picks. Maximizing the probability of scoring the most points in the pool, or maximizing the expected number of points gained, provide reasonable objectives for such players. The probability that a given player scores the most points in a pool depends upon the actions of other players, which again raises difficult game-theoretic issues, especially in large pools. Maximizing the expected number of points gained, however, depends only upon the tournament structure, the chances that different teams will defeat other teams, and the point scheme employed by the pool.

The authors of this paper have unsuccessfully participated in NCAA pools for more than a decade. Simply put, we decided that we would like to win once in a while, or at least come out close to the top. Surely there must be better and worse strategies for garnering points in an office pool, but our experience suggested that we had only mastered the latter.

This paper began as a concerted effort to reverse this trend, but it has evolved into much more. The rules for office pools combined with the beautiful mathematical structure of single-round elimination tournaments give rise to a rich family of models that are interesting in their own right. Though others have studied the mathematical structure of elimination tournaments (David 1959, Edwards 1996, Horen and Riezman 1985, Knuth 1987), to our knowledge, Breiter and Carlin (1997) (henceforth BC) are the only authors who have addressed office pools thus far. BC first presented a model for estimating the chance that any team would defeat any other team based on Las Vegas point spreads and the ratings of sports expert Jeff Sagarin. For each of the four 16-team regional subtournaments of the NCAA men's basketball tournament, BC used this model to perform 10,000 Monte Carlo simulation runs for each possible slate of game predictions to determine the set of predictions that maximized the expected points gained under office pool rules reflective of actual contests. They compared their optimal-by-simulation and enumeration strategies to the actual results of the 1996 NCAA men's tournament to see how well their strategies would have fared in practice, and showed that their model-derived strategies indeed provided an improvement in the actual (as opposed to expected) score over simpler strategies such as just picking the seeded favorites to advance.

In contrast to the simulation and enumeration approach of BC, we present exact recursions for computing the mean and the variance of the number of points garnered in an office pool for any given slate of predicted game winners, obviating the need for simulation. In addition, our recursion for the expected number of points in an office pool is ideally suited for determining a slate of predicted game winners that maximizes the expected number of points in an office pool via dynamic programming, obviating the need for explicit enumeration. While BC considered the 16-team regional tournaments (which required them to simulate $2^{15} = 32{,}768$ strategies for each tournament, eating 8 hours on a Sparc20 workstation in the process), they did not attempt to optimize over the entire tournament (which would have required enumeration of $2^{63} \approx 9.22 \times 10^{18}$ different strategies). By contrast, in a second or two our dynamic program determines an optimal strategy over the *entire* tournament for a wide class of office pool scoring rules, including those commonly found on the Internet.

In the next section, we review the structure of single elimination tournaments, and develop a reasonably general model for office pools within this structure. In §3, we show how to evaluate office pool prediction strategies by developing recursions for the mean and the variance of the total points awarded corresponding to any particular set of predictions. We demonstrate these ideas via recourse to a random tournament not only because these results are of independent interest, but also because the ensuing predictions serve as a basis for determining whether any given strategy performs better than chance. Determining the variance of the total score in an office pool is not easy, so for continuity we have dispatched the details to the Appendix. In §4, we develop a dynamic program for selecting that prediction strategy that maximizes the expected total score in an office pool. It is actually easier to identify the optimal strategy than it is to calculate the variance of the total score for any single strategy! We return to men's college basketball in §5 where we propose three alternative Markov models for the probability that any team defeats any other team in any encounter. In §6 we apply our models to the 1998 and 1999 NCAA and NIT post-season tournaments. We discover that, although our models perform reasonably well, they do not outperform the simple strategy of picking the seeds when the sole objective is to correctly pick as many game winners as possible. The contribution of our models increases, however, as the scoring system used in the office pool becomes more sophisticated. We offer brief closing remarks in §7.

2. Single Elimination Office Pools

2.1. Single Elimination Tournament Structure

In this section, we describe briefly the structure of the tournaments under consideration (see Edwards 1996 for more detail). We consider k-round single elimination tournaments with 2^k entrants (or teams in the basketball context). In such a tournament, game winners in a given round proceed to play in the next round, while game losers are eliminated from the tournament. This continues until there is only one undefeated team left: the tournament champion! Thus, in the rth round of a k-round single elimination tournament, there are 2^{k-r} games, $r = 1, 2, \ldots, k$. Note that in total, such a tournament contains

$$\sum_{r=1}^{k} 2^{k-r} = 2^k - 1 \tag{2.1}$$

games. This result can be understood intuitively by noting that with the exception of the tournament champion, all $2^k - 1$ other teams in the tournament lose *exactly once*. In the NCAA college basketball tournament, there are $k = 6$ rounds, $2^6 = 64$ teams, and hence 63 games.

Initial pairings, and all possible subsequent games, are dictated by the tournament *bracket* (or *draw*). As shown in Figure 1, the tournament bracket induces a *binary tree* linking game winners from one round to the next. Let V_{rg} denote the identification of the winning team (the *victor*) in game g of round r. As shown in Figure 1, the participants in all first round games are determined by the tournament draw, while in round $r > 1$, game g is played between teams $V_{r-1,\,2g-1}$ and $V_{r-1,\,2g}$, the victors of games $2g - 1$ and $2g$ in round $r - 1$.

The tournament draw dictates which teams are eligible to play in which games of each round of the tournament. Let $\tau(r, g)$ be the set of teams that could conceivably play in game g of round r, $g = 1, 2, \ldots, 2^{k-r}$; $r = 1, 2, \ldots, k$. Because game g in round r is between the winners of games $2g - 1$ and $2g$ of round $r - 1$, only teams playing in these latter games are eligible to play in game g of round r. Consequently, the set of all teams that could conceivably play in game g of round r is given by

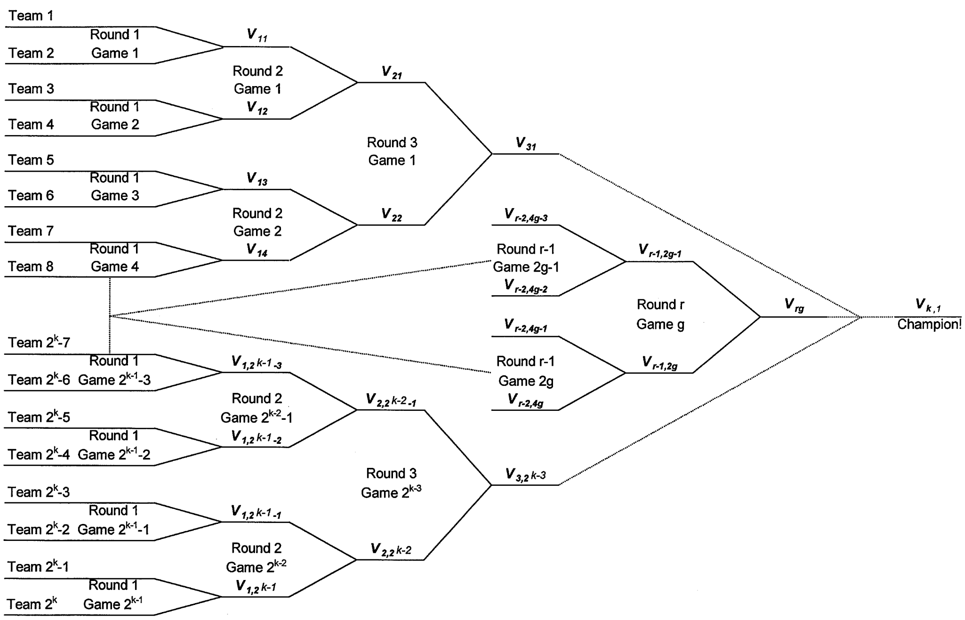
$$\tau(r, g) = \tau(r-1, 2g-1) \cup \tau(r-1, 2g) \quad \text{for } g = 1, 2, \ldots, 2^{k-r}; \; r = 1, 2, \ldots, k, \tag{2.2}$$

since any of the teams that could play in games $2g - 1$ or $2g$ of round $r - 1$ could also play in game g of round r. Now let i denote the index of a team in the tournament draw, $i = 1, 2, \ldots, 2^k$, and equate the team index i to $\tau(0, i)$, the set of teams that play in fictional "game" i of fictional "round" 0. Iterating the recursion in Equation 2.2 from this initial condition yields the result

$$\tau(r, g) = \{i \mid 2^r(g-1) + 1 \le i \le 2^r g\} \quad \text{for } g = 1, 2, \ldots, 2^{k-r}; \; r = 1, 2, \ldots, k. \tag{2.3}$$

Equation 2.3 thus identifies the set of all possible teams that could conceivably play in game g of round r.

Figure 1 The Tournament Bracket in a k-round Single Elimination Tournament



It is also convenient to identify both the exact game any team would play having advanced to a given round r, and the set of possible opponents any team would face in that game. Let $\gamma(i, r)$ be the index for the game team i would play in round r. From 2.3 we see that if team i can play in game g of round r, then

$$\frac{i}{2^r} \leq g \leq \frac{i-1}{2^r} + 1. \tag{2.4}$$

The integrality of g implies that

$$\gamma(i, r) = \left\lceil \frac{i}{2^r} \right\rceil \quad \text{for } i = 1, 2, \ldots, 2^k; \quad r = 1, 2, \ldots, k \tag{2.5}$$

where the ceiling function $\lceil x \rceil$ denotes the smallest integer $\geq x$. We then obtain immediately that $\phi(i, r)$, the set of opponents team i could possibly face in round r, is given by

$$\phi(i,r) = \begin{cases} \tau(r-1, 2\gamma(i,r)) & i \in \tau(r-1, 2\gamma(i,r)-1) \\ \tau(r-1, 2\gamma(i,r)-1) & i \in \tau(r-1, 2\gamma(i,r)) \end{cases}. \tag{2.6}$$

2.2. Office Pools

In an *office pool*, participants make predictions regarding the winners of all games in a tournament. Points are awarded for correct winning predictions, and the pool participant who garners the largest total number of points wins the pool and any associated prizes. With respect to the NCAA men's basketball tournament, there is seemingly no end to the number of office pools across the United States and around the globe, as any cursory Internet search will reveal.

An interesting feature of office pools is that one only has to predict *winners* correctly. For example, in a 2-round tournament, suppose one picks Team 1 to defeat 2 in the first game of Round 1, 3 to defeat 4 in the second game, and 1 to defeat 3 in Round 2 (the championship). If 1 defeats 2, 4 defeats 3, and then 1 defeats 4 in the final, one would receive credit for making two correct winning predictions.

Recall that V_{rg} is the random variable reporting the identification of the victor of game g in round r, and let v_{rg} be the prediction of some pool participant for the victor of game g in round r; $g = 1, 2, \ldots, 2^{k-r}$; $r = 1, 2, \ldots, k$. We only consider *consistent* prediction strategies of the form $v_{rg} \in (v_{r-1,2g-1}, v_{r-1,2g})$, which says that a team can only be predicted to win game g in round r if it is also predicted to win either game $2g-1$ or $2g$ in round $r-1$. A correct prediction occurs when $V_{rg} = v_{rg}$. Define the random variables $W_{rg}(v_{rg})$ as

$$W_{rg}(v_{rg}) = \begin{cases} 1 & V_{rg} = v_{rg} \text{ for } g = 1, 2, \ldots, 2^{k-r}; \quad r = 1, 2, \ldots, k. \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

Clearly, $W_{rg}(v_{rg})$ is the indicator random variable that reveals whether the prediction v_{rg} correctly identifies the winner in game g of round r.

To determine scores in office pools, let $\pi_{rg}(v_{rg})$ denote the number of points that are awarded if team v_{rg} is correctly predicted to win game g of round r. The total score corresponding to a complete tournament *prediction strategy* $\mathbf{v} = \{v_{rg}\}$ for $g = 1, 2, \ldots, 2^{k-r}$; $r = 1, 2, \ldots, k$, denoted by $T(\mathbf{v})$, is then given by

$$T(\mathbf{v}) = \sum_{r=1}^{k} \sum_{g=1}^{2^{k-r}} \pi_{rg}(v_{rg}) W_{rg}(v_{rg}). \tag{2.8}$$

Much of the ensuing analysis in this paper will focus on evaluating the mean and variance of $T(\mathbf{v})$ for a given prediction strategy $\mathbf{v}$, and determining the *optimal* prediction strategy $\mathbf{v}^*$ that serves to maximize the expected total score $E(T(\mathbf{v}))$.

While different pools have different rules for how points are awarded, our scoring model allows for the most common schemes encountered in practice. In the simplest case, single points are awarded for correct winning predictions, in which case $\pi_{rg}(v_{rg}) = 1$ $\forall g, r$ and the total score for a given entrant $T(\mathbf{v})$ corresponds to the total number of games in which the winner was correctly forecasted. With respect to the NCAA basketball tournament, many pools use scoring schemes where the points awarded increase with the rounds of the tournament (so $\pi_{rg}(\bullet)$ would be increasing in r). More elaborate contests encourage pool participants to support weaker teams by offering a point system that greatly rewards the correct prediction of victories by teams judged to be weak relative to the tournament favorites a priori. Our scoring model can incorporate any interaction desired between the predicted winner of a given game and the location of the game itself in the tournament (as identified by the indices g and r), allowing us the flexibility to consider several different scoring schemes.

3. Evaluating Office Pool Prediction Strategies

3.1. Fundamental Recursion of Office Pools

We will now exploit an important property of the tournament bracket. As is clear from Figure 1, any game g in round r completes (or *crowns*) a subtournament involving all of the teams $i \in \tau(r, g)$. Define $T_{rg}(\mathbf{v})$ as the partial score for a prediction strategy $\mathbf{v}$ counting only games within the subtournament crowned by game g in round r. Since game g in round r is played between the victors of games $2g-1$ and $2g$ in round $r-1$, we have

$$T_{rg}(\mathbf{v}) = T_{r-1,2g-1}(\mathbf{v}) + T_{r-1,2g}(\mathbf{v}) + \pi_{rg}(v_{rg}) W_{rg}(v_{rg}), \tag{3.1}$$

for $r = 1, 2, \ldots, k$; $g = 1, 2, \ldots, 2^{k-r}$. Note that $T_{k1}(\mathbf{v}) = T(\mathbf{v})$. Iterating equation 3.1 yields

$$T_{rg}(\mathbf{v}) = \sum_{j=1}^{r} \sum_{h=2^{r-j}(g-1)+1}^{2^{r-j}g} \pi_{jh}(v_{jh}) W_{jh}(v_{jh}), \tag{3.2}$$

for $r = 1, 2, \ldots, k$; $g = 1, 2, \ldots, 2^{k-r}$. We will make use of this result subsequently.

3.2. Evaluating the Mean Total Score

Consider any consistent prediction strategy $\mathbf{v}$ in an office pool. Let $\mu_{rg}(\mathbf{v})$ denote the expected total number of points garnered via $\mathbf{v}$ in all games of the subtournament crowned by game g in round r, and let $\omega_{rg}(v_{rg}) = \Pr\{W_{rg}(v_{rg}) = 1\}$. Taking expectations on both sides of Equation 3.1 yields the recursion

$$\mu_{rg}(\mathbf{v}) = \mu_{r-1,2g-1}(\mathbf{v}) + \mu_{r-1,2g}(\mathbf{v}) + \pi_{rg}(v_{rg}) \omega_{rg}(v_{rg}), \tag{3.3}$$

for $r = 1, 2, \ldots, k$; $g = 1, 2, \ldots, 2^{k-r}$ where $\mu_{0g} = 0$ for $g = 1, 2, \ldots, 2^k$.

3.2.1. Example: A Random Tournament. Consider a *random* tournament where a single point is awarded for each correctly chosen game winner ($\pi_{rg}(v_{rg}) = 1$ $\forall r, g$). The word "random" could have two different meanings here, both of which lead to equivalent results. One interpretation is that in any game between any two teams, each team stands to win with probability 1/2 independently of the results of other games. The other interpretation is that for a tournament where teams have arbitrary probabilities of winning in whatever games they may play, the pool participant selects a prediction strategy from the set of all possible 2^{2^k-1} prediction strategies (since there are two possible outcomes for each of the $2^k - 1$ games in the tournament) at random (so any particular prediction strategy is chosen with a priori probability $1/2^{2^k-1}$).

Under either interpretation, the probability that the winner of a particular game in the rth round is called correctly is given by $1/2^r$. For the first interpretation, note that any team picked to win in round r must have been picked in the $r-1$ previous rounds, and since there is a 1/2 probability of that team winning in any game it plays, the unconditional probability of winning in round r must equal $1/2^r$. For the second interpretation, note that any game in the rth round crowns a single elimination subtournament of exactly $2^r - 1$ games. There are thus 2^{2^r-1} different prediction strategies corresponding to this subtournament, and, of these, exactly 2^{2^r-1-r} select a particular team to win r times in a row (since after removing the r games for which the team in question has been selected to win, there are $2^r - 1 - r$ remaining games in the subtournament, each with two possible outcomes). The probability of correctly picking the winner of a game in the rth round under this interpretation is thus given by $2^{2^r-1-r}/2^{2^r-1} = 1/2^r$ as well.

What is the expected number of correctly predicted winners in a random single elimination tournament? Let μ_{k-1} denote the mean number of correctly predicted wins in a random tournament with $k-1$ rounds, and then consider a random tournament with k rounds formed by playing off the winners of two disjoint random $(k-1)$-round tournaments. Under these assumptions, the expected number of correctly called games will be the same in each of the two $(k-1)$-round tournaments (and there is no dependence on whatever prediction strategy is used). Specializing Equation 3.3 to obtain the expected number of correct predictions for a random tournament with k rounds, we see that

$$\mu_k = 2\mu_{k-1} + \frac{1}{2^k} \quad \text{for } k = 1, 2, 3, \ldots, \tag{3.4}$$

where we define $\mu_0 = 0$. The solution to this is given by

$$\mu_k = \sum_{r=1}^{k} 2^{k-r} \frac{1}{2^r} = \frac{2^k}{3}\left(1 - \frac{1}{4^k}\right). \tag{3.5}$$

Equation 3.5 can be understood as follows: In round r, there are 2^{k-r} games. The chance of correctly identifying the winner of any game in round r is equal to $1/2^r$, and thus the expected number of correct winning predictions in round r equals $2^{k-r}/2^r$. Summing over all rounds yields the expected number of correct predictions for the entire tournament. Since there are $2^k - 1$ games in a k-round single elimination tournament, we see that the fraction of games that are correctly called in a random tournament rapidly approaches 1/3 as the number of rounds in the tournament grows.

3.2.2. Example: A Markov Tournament. In a *Markov tournament*, we assume that the probability that team i defeats team j in any game involving these two teams equals p_{ij}, independently of the outcomes of prior games. The Markov assumption is popular among those who have modeled single elimination tournaments (for examples, see David 1959, Edwards 1996, Horen and Riezman 1985, and Knuth 1987). Use of Equation 3.3 is possible on assessing the probabilities $\omega_{rg}(v_{rg}) = \Pr\{W_{rg}(v_{rg}) = 1\}$. The Markov assumptions enable such an assessment (David 1959, Edwards 1996).

Suppose that the team picked to win game g of round r (v_{rg}) is also the team that was picked to win game $2g-1$ of round $r-1$ (so $v_{rg} = v_{r-1,2g-1}$). Then for v_{rg} to be a successful prediction in game g of round r, this team must both be a successful prediction in game $2g-1$ of round $r-1$ (that is, $W_{r-1,2g-1}(v_{rg}) = 1$) and defeat whatever opponent is faced in game g of round r. If team ℓ is a possible opponent of v_{rg} in round r (so $\ell \in \phi(v_{rg}, r)$), then conditional upon v_{rg} winning its game in round $r-1$, team ℓ will face team v_{rg} in round r with probability $\omega_{r-1,2g}(\ell)$, but lose with probability $p_{v_{rg},\ell}$. Considering all possible opponents, we see that if $v_{rg} = v_{r-1,2g-1}$, then

$$\omega_{rg}(v_{rg}) = \omega_{r-1,2g-1}(v_{rg}) \sum_{\ell \in \phi(v_{rg}, r)} \omega_{r-1,2g}(\ell) p_{v_{rg},\ell}. \tag{3.6}$$

A similar result holds if v_{rg} instead was chosen as the winner of game $2g$ in round $r-1$. Summarizing, we see that

$$\omega_{rg}(v_{rg}) = \begin{cases} \omega_{r-1,2g-1}(v_{rg}) \sum_{\ell \in \phi(v_{rg},r)} \omega_{r-1,2g}(\ell) p_{v_{rg},\ell}. \\ \quad \text{if } v_{rg} = v_{r-1,2g-1} \\ \omega_{r-1,2g}(v_{rg}) \sum_{\ell \in \phi(v_{rg},r)} \omega_{r-1,2g-1}(\ell) p_{v_{rg},\ell}. \\ \quad \text{if } v_{rg} = v_{r-1,2g} \end{cases}, \quad (3.7)$$

for $r = 1, 2, \ldots, k$ and $g = 1, 2, \ldots, 2^{k-r}$. Substituting the results from Equation 3.7 into Equation 3.3 enables rapid evaluation of the expected total score for any given prediction strategy $\mathbf{v}$.

3.3. Evaluating the Variance of the Total Score

Equation 3.1 can also be used to find the variance of the total score. Recall that $T_{rg}(\mathbf{v})$ is the total score obtained from following prediction strategy $\mathbf{v}$ through the subtournament crowned by game g in round r, and let $\sigma^2_{rg}(\mathbf{v}) = Var(T_{rg}(\mathbf{v}))$. From Equation 3.1 we directly obtain the recursion

$$\begin{aligned} \sigma^2_{rg}(\mathbf{v}) = {} & \sigma^2_{r-1,2g-1}(\mathbf{v}) + \sigma^2_{r-1,2g}(\mathbf{v}) \\ & + \pi^2_{rg}(v_{rg}) \omega_{rg}(v_{rg})(1 - \omega_{rg}(v_{rg})) \\ & + 2\{\mathrm{Cov}(T_{r-1,2g-1}(\mathbf{v}), T_{r-1,2g}(\mathbf{v})) \\ & \quad + \mathrm{Cov}(T_{r-1,2g-1}(\mathbf{v}), \pi_{rg}(v_{rg}) W_{rg}(v_{rg})) \\ & \quad + \mathrm{Cov}(T_{r-1,2g}(\mathbf{v}), \pi_{rg}(v_{rg}) W_{rg}(v_{rg}))\}, \quad (3.8) \end{aligned}$$

for $r = 1, 2, \ldots, k$ and $g = 1, 2, \ldots, 2^{k-r}$. For specific probability models, this recursion can be specialized as illustrated in the Appendix. Typically the challenge rests with the evaluation of the last two terms in this equation.

4. Optimal Prediction Strategies for Office Pools

Now consider the problem of selecting a prediction strategy $\mathbf{v}$ to maximize the expected total score $E(T(\mathbf{v}))$. This can be accomplished via a dynamic program that exploits the mean value recursion of Equation 3.3. Define $\mu^*_{rg}(i)$ as the expected total score in the subtournament crowned by game g in round r when team i is predicted to win games $\gamma(i, j)$ for $j = 1, 2, \ldots, r$ and all other predictions within this subtournament are optimal *for this subtournament*; $\mu^*_{rg}(i)$ is thus the value function for this problem. Also, define μ^*_{rg} as the optimal expected total score for the subtournament crowned by game g in round r, and v^*_{rg} as the optimal prediction for the winner of this subtournament. From 3.3 we obtain the recursion:

$$\mu^*_{rg}(i) = \begin{cases} \mu^*_{r-1,2g-1}(i) + \mu^*_{r-1,2g} + \pi_{rg}(i)\omega_{rg}(i) \\ \quad \text{if } i \in \tau(r-1, 2g-1) \\ \mu^*_{r-1,2g}(i) + \mu^*_{r-1,2g-1} + \pi_{rg}(i)\omega_{rg}(i) \\ \quad \text{if } i \in \tau(r-1, 2g) \end{cases}, \quad (4.1)$$

for $r = 1, 2, \ldots, k$, and $g = 1, 2, \ldots, 2^{k-r}$, and $\mu^*_{0i}(i) = \mu^*_{0i} = 0$ for $i = 1, 2, \ldots, 2^k$. The optimal subtournament expected total scores μ^*_{rg} are then given by

$$\mu^*_{rg} = \max_{i \in \tau(r,g)} \mu^*_{rg}(i), \quad (4.2)$$

while the optimal subtournament winning predictions v^*_{rg} are set equal to

$$v^*_{rg} = \arg\max_{i \in \tau(r,g)} \mu^*_{rg}(i) \quad (4.3)$$

again for $r = 1, 2, \ldots, k$ and $g = 1, 2, \ldots, 2^{k-r}$.

To obtain the overall optimal prediction strategy requires "unpacking" the optimal subtournament predictions in a consistent manner to ensure that predicted winners in a given round are also predicted winners in all previous rounds. The optimal prediction for the winner of the entire tournament is equal to v^*_{k1}, while the optimal total score in the office pool over the entire tournament is equal to μ^*_{k1}. The following unpacking algorithm returns the optimal prediction strategy on a game by game basis, along with the corresponding optimal expected total scores.

Algorithm UNPACK
FOR $r = k$ to 2
  FOR $g = 1$ to 2^{k-r}
  IF $v^*_{rg} \in \tau(r-1, 2g-1)$ THEN
  $v^*_{r-1,2g-1} \leftarrow v^*_{rg}$
  $\mu^*_{r-1,2g-1} \leftarrow \mu^*_{r-1,2g-1}(v^*_{rg})$
  ELSE
  $v^*_{r-1,2g} \leftarrow v^*_{rg}$
  $\mu^*_{r-1,2g} \leftarrow \mu^*_{r-1,2g}(v^*_{rg})$
  NEXT g
NEXT r

5. Markov Models for College Basketball Tournaments

To apply the models developed above requires the development of probability models for predicting team performance. Several authors have considered the use of Markov models for predicting game winners in college basketball, but almost all of these have been statistical models calibrated within tournaments based on team seedings (Edwards 1998, Schwertman et al. 1991, 1996). These seedings are meant to reflect the NCAA tournament committee's view of the strengths of the teams in the tournament: the lower the seed, the better the team (and number one seeds are typically favorites to win the tournament). However, the office pool problem seemingly precludes the use of calibrated models based on team seedings from previous tournaments, for although the seeding structure itself is unchanged, the actual teams seeded in the tournament are completely different from year to year. On the other hand, seeding information is the simplest indicator of team quality available at the start of a tournament, and indeed one simple strategy for office pool predictions is simply to "go with the seeds." In the sections that follow, we will consider three Markov models that can be constructed with relative ease before the start of a tournament that do not rely on seeding information. In assessing the performance of these models, however, we will compare the accuracy of model-dependent optimal prediction strategies to the simple strategy of going with the seeds later on.

5.1. A Regular Season Model

The first model we consider is one based on regular season records (publicly available from sources such as *ESPN* and *CBS Sportsline* among others). Essentially, we view the tournament as an extension of the regular season, and argue that tournament performance should reflect ability in the season just ended. The organization of NCAA college basketball is complicated, with different teams belonging to different conferences with differing levels of play. Any one college team will only play a small fraction of all possible teams, and, in a given season, teams that have not faced each other often meet for the first time in tournament play. How can one predict the outcome of a game between two teams that have not played each other?

Our model is very simple. Let *NCAA* denote the set of NCAA Division 1 basketball teams. For team $i \in NCAA$, we assume the existence of a *strength coefficient* $s_i \geq 0$. We then postulate that in *any* game between team i and team j, i will defeat j with probability

$$p_{ij} = \frac{s_i}{s_i + s_j}\,. \tag{5.1}$$

The parameters of this model, known as a Bradley-Terry model (Bradley and Terry 1952), can only be identified to a multiplicative constant, so we force $\sum_{i \in NCAA} s_i = 1$.

To estimate this model, let n_{ij} denote the observed number of times team i defeated team j during the regular season (including conference tournaments). We select the strength coefficients by maximizing the log-likelihood function

$$\log \mathcal{L} = \sum_{i,\, j \in NCAA} n_{ij} \log \frac{s_i}{s_i + s_j}, \tag{5.2}$$

subject to $\sum_{i \in NCAA} s_i = 1$ and $s_i \geq 0\ \forall i \in NCAA$. In our actual applications described below, we did not consider all NCAA teams individually. Rather, in estimating the strength coefficients we included all 64 teams in the NCAA tournament, all 32 teams from the National Invitational Tournament (the NIT), and lumped all remaining nontournament teams into one "megateam" for which we estimated a single strength coefficient. This was achieved by treating any regular season game between one of the 96 tournament teams and a nontournament team as a game between the individual tournament team in question and the megateam.

At first blush, this model might appear overly simplistic, as it seemingly ignores the complicated relationships between teams. However, it is important to note that the interactions between teams are implicitly accounted for in this model via the estimation process itself. For example, suppose that there are only 3 teams, a, b, and c, and suppose that games are only played between a and b, and b and c; a and c never play. Further suppose that the fraction of games in which a defeats b was observed to equal f_{ab} while the fraction of games in which c defeats b equals f_{cb}. The strength coefficients would be uniquely determined in this instance as the solutions to

$$\frac{s_a}{s_a + s_b} = f_{ab}, \tag{5.3}$$

$$\frac{s_c}{s_b + s_c} = f_{cb}, \tag{5.4}$$

$$s_a + s_b + s_c = 1. \tag{5.5}$$

Solution yields a strength coefficient s_a of

$$s_a = \frac{f_{ab}(1 - f_{cb})}{1 - f_{ab} f_{cb}}. \tag{5.6}$$

What is important to note is that even though team a never plays team c, the strength coefficient s_a *depends* on the observed performance of team c. In applying this model to the NCAA data, the same principle applies. The strength coefficient estimated for any given team depends crucially on the performance of other teams, even if there were no games observed between the teams in question.

5.2. An Expert Rating Model

There are various "expert" ratings of team performance provided over the course of the college basketball season, including Sagarin ratings (published in *USA Today*), Massey ratings (⟨www.mratings.com⟩), the NCAA's Ratings Percentage Index (⟨collegerpi.com⟩), and so forth. Perhaps these experts actually know something, in which case we can borrow that knowledge. We focus our attention on the Sagarin ratings, which are meant to be interpreted as expected scoring rates on a team by team basis. Let λ_i denote the Sagarin rating for team i. Our model *assumes* that basketball teams score points in accord with uncorrelated Poisson processes, and consequently the *point spread* X_{ij} between teams i and j has a mean given by $\lambda_i - \lambda_j$, and variance given by $\lambda_i + \lambda_j$. Since the typical values for Sagarin ratings are large (the average for 1999 NCAA tournament teams was 83.5), we further approximate the distribution of the point spread X_{ij} by a normal distribution with mean and variance as stated earlier (for the normal distribution nicely approximates the Poisson for variables with sufficiently large expected values, which makes the point spread approximately the difference in two normally distributed random variables). The probability p_{ij} that team i defeats team j in any game is then simply the probability that the point spread is positive (there are no ties in the tournament), which is given by

$$p_{ij} = \Pr\{X_{ij} > 0\} = \Phi\left(\frac{\lambda_i - \lambda_j}{\sqrt{\lambda_i + \lambda_j}}\right), \tag{5.7}$$

where $\Phi(\bullet)$ is the cumulative distribution function for the standard normal random variable.

We have not attempted to verify whether the actual points scored in basketball games follow the Poisson distribution, as we are less interested in the statistical truths of point spread distributions than in improved performance in office pools. However, empirical support for the use of the normal distribution (albeit with a constant standard deviation) as a model for point spreads in football is presented in Stern (1991).

Is it reasonable to assume that the points scored by two competing teams in a given game are uncorrelated? One could argue that the more points scored by one team, the fewer scored by the other, which would lead to a negative correlation between team scores. Alternatively, one could argue that games take on the characteristics of offensive or defensive battles, which would lead to a positive correlation between team scores. For example, a reviewer suggested that by slowing the offensive tempo of the game via use of the shot clock, a team could both reduce the number of points it scores as well as those scored by its opponent (by denying them time with the ball), inducing a positive correlation. Citing Stern's (1991) football study, Carlin (1996) has argued that the actual point spreads in NCAA tournament games roughly follow a normal distribution, but with a constant standard deviation. Such a model was also employed by Breiter and Carlin (1997).

There is an interesting implication of the Carlin model: If X_i and X_j are the number of points scored by teams i and j in a game against each other, then a correlated Poisson scoring model with a constant point spread standard deviation θ forces the relationship

$$\mathrm{Var}(X_i - X_j) = \lambda_i + \lambda_j - 2\,\mathrm{Cov}(X_i, X_j) = \theta^2. \tag{5.8}$$

If $\theta^2 < \lambda_i + \lambda_j$, Carlin's model implies a positive correlation between X_i and X_j while if $\theta^2 > \lambda_i + \lambda_j$ a negative correlation results. One implication of this model is that the points scored by each team in games between high scoring teams on average (high-λ teams) will be positively correlated, while the points scored by each team in games between low scoring teams on average (low-λ teams) will be negatively correlated.

It is certainly easier to assume that team scoring is uncorrelated as this avoids estimating θ. We also note that the range of values that can be assumed by $\sqrt{\lambda_i+\lambda_j}$ based on actual Sagarin ratings is not large. Considering the 1999 NCAA Tournament as an example, the maximum value possible of $\sqrt{\lambda_i+\lambda_j}$ was equal to 14.2, while the minimum possible value was equal to 11.3; most games would involve values of $\sqrt{\lambda_i+\lambda_j}$ near 13. Carlin (1996) used the constant value of $\theta = 10 < 13$ in his models, which implies a belief that the scores in most games are positively correlated. Again, our interest is less in the nature of the correlation between game scores than in whether the use of a simple model enables better office pool performance. We therefore retain our simplifying assumption that the points scored by competing teams in the same game are uncorrelated random variables.

5.3. A Model Based on Las Vegas Odds

An alternative source for information on team performance is the market. Las Vegas sports books offer a variety of different bets on NCAA tournament games. Two popular bets are based on point spreads and point totals (that is, the sum of the points scored by both teams in a game). Let λ_i and λ_j again represent the average scoring rates per game of teams i and j, and let x_{ij} and y_{ij} be the point spreads and point totals posted just before the start of the tournament (available from *The Las Vegas Sun*, the *Wager Information Network* at ⟨winmor.com⟩, or other sources). If the market is a good judge of the truth in the sense that the posted point spreads and totals are correct on average, then

$$x_{ij} = \lambda_i - \lambda_j, \tag{5.9}$$

and

$$y_{ij} = \lambda_i + \lambda_j. \tag{5.10}$$

These equations solve to yield

$$\lambda_i = \frac{x_{ij}+y_{ij}}{2}, \tag{5.11}$$

and

$$\lambda_j = \frac{y_{ij}-x_{ij}}{2}. \tag{5.12}$$

We have calculated the scoring rates using Equation 5.11–5.12, and again employed Equation 5.7 to estimate the probability that i defeats j in any game.

6. March Madness!

6.1. Maximizing the Number of Correct Predictions

To illustrate our models, we focus on the 1998 and 1999 NCAA and NIT men's basketball tournaments. For each of these four tournaments, we have (i) implemented our three Markov models (referred to as the regular season, Sagarin, and Las Vegas odds models respectively) as described in §5, (ii) optimized the tournament predictions using the methods of §4 after setting $\pi_{rg}(\bullet) = 1\ \forall r, g$, (iii) computed the variance of the total number of correct predictions using the methods described in the Appendix, (iv) counted the *actual* number of correct predictions corresponding to our optimal strategies, and (v) compared these results to predictions based on the tournament seedings, and to what would be expected from random predictions (as discerned from our random tournament models). Virtually all of the data required for our models were collected via the internet from sources such as *ESPN, CBS Sportsline, USA Today*, the *Las Vegas Sun,* and the *Wager Information Network* (⟨winmor.com⟩).

The results are summarized in Tables 1 and 2. Table 1 reports the overall agreement between the three models, picks based on the tournament seedings, and the actual results obtained from all 188 games considered (63 games in each of the 1998 and 1999 NCAA tournaments, and 31 games in each of the 1998 and 1999 NIT tournaments). Each cell in Table 1

Table 1 **Overall Prediction Agreement Over 188 Games (%) (1998 and 1999 NCAA and NIT Tournaments)**

	Regular Season	Sagarin	Las Vegas Odds	Actual Results
Pick the Seeds	146 (78)	146 (78)	156 (83)	106 (56)
Regular Season		152 (81)	133 (71)	111 (59)
Sagarin			141 (75)	108 (57)
Las Vegas Odds				110 (59)

Table 2 Prediction Performance by Tournament

	NCAA 1999 Actual/Expected/Standard Deviation of Wins	NCAA 1998 Actual/Expected/Standard Deviation of Wins	NIT 1999 Actual/Expected/Standard Deviation of Wins	NIT 1998 Actual/Expected/Standard Deviation of Wins
Chance	–/21.3/4.5	–/21.3/4.5	–/10.7/3.2	–/10.7/3.2
Pick the Seeds	36/–/–	39/–/–	17/–/–	13/–/–
Regular Season	39/42.6/4.3	37/43.5/4.1	17/12.9/3.4	18/15.0/3.4
Sagarin	41/41.4/4.2	39/38.5/4.5	15/13.1/3.4	13/13.0/3.4
Las Vegas Odds	38/44.6/3.8	35/45.1/4.3	22/17.5/3.7	15/19.2/3.7

states the number (and percentage) of all 188 games in which the row and column predictions/realizations for the winning team were identical. Overall, the predictive accuracies of the models do not differ greatly from each other or from the simple strategy of picking the seeds. The optimized regular season model correctly identified 111 of the 188 games considered, followed by 110 correct predictions for the Las Vegas odds model, 108 for the Sagarin model, and 106 based on simply picking the seeds. It is interesting to note that in comparing the predictions to each other, the predictions derived from Las Vegas odds agreed with picking the seeds on 156 out of 188 games, while the regular season and Sagarin models agreed on 152 out of 188 games. One possible explanation for such high concordance in predictions is that those placing bets in Las Vegas are relying heavily on the tournament seedings, while the expert Sagarin ratings and the regular season model have extracted essentially the same information from regular season records.

Table 2 reports the performance of each model for each tournament along with the mean and standard deviation of the number of correct predictions associated with each model. Also reported are the tournament by tournament performance of picking by the seeds, and the mean and standard deviation one would expect from random predictions. Perhaps the most striking feature of Table 2 is that for all four tournaments considered and for all three of our models from §5, the actual number of correct predictions falls within two standard deviations of the expected number 11 out of 12 times, and on six occasions the number of correct predictions is within one standard deviation of what is expected. The Sagarin model is particularly noteworthy in its consistency: The expected number of game winners correctly predicted by this model equals 41.4, 38.5, 13.1, and 13.0 across the four tournaments. The actual number of correct predictions for the Sagarin model were equal to 41, 39, 15, and 13, respectively, a remarkable result.

Table 2 shows that is very difficult to discount any of the models we have considered on empirical grounds, for the actual performance of these models is well within the variability expected of each. However, it is also very difficult to discriminate among the models considered based on the data. What is clear is that all of our models perform much better than chance would suggest; all of the models are successful more than $\frac{1}{3}$ of the time. Then again, predictions based solely on tournament seeding are just as competitive as our models. Indeed, for the 1998 NCAA and 1999 NIT tournaments, none of our models performed better than picking by the seeds (though in both cases there were models that performed equally well).

The results from Table 2 suggest that, if the goal is to choose as many correct game winners as possible, our models do not outperform simply picking the seeds. This will also be true in any office pool where the point structure is such that maximizing the total number of points is equivalent to maximizing the number of correctly chosen winners. However, in office pools where the point structures are more complicated, our models can make a difference, as discussed below.

6.2. More Complicated Point Structures: The Packard Pool

An office pool managed by Professor Erik Packard of the mathematics department at Mesa State University offers an interesting test of our models ("Contest 2" at ⟨mesastate.edu/~epackard/hoop/contest.html⟩). Scoring in this pool works as follows: If a number one seed is correctly chosen to win its game in round r, then the pool rewards such predictions with $\rho(r) = 1890, 3960, 7392, 12320, 18480,$ and 27720 points for $r = 1$ through 6, respectively. A correct winning prediction in round r for a team seeded in position n is rewarded with $(1+n)\rho(r)/2$ points for $n = 1, 2, \ldots, 16$ and $r = 1, 2, \ldots, 6$. Thus, the point system used in this contest offers more points to teams perceived as weaker a priori (that is, teams with higher seedings), but also rewards correct predictions that reach deeper into the tournament. This scoring system fits within the general scoring framework we have considered, so it is possible to optimize our models from §5 for use with the "Packard Pool" (see the Internet site referenced above for a justification of the specific point scheme employed). We can compare these results to what would be expected from a simple strategy of picking the seeds. In addition, since the website for this pool reports participants' performance for the 1998 and 1999 NCAA tournaments, we can see how we would have fared had we actually entered the pool.

The results appear in Table 3. Apparently, we would have fared quite well. The winning entry in the 1999 pool garnered a total of 333,572 points. Our optimized Sagarin model scores 405,421 points, which would have won the pool. The models based on Las Vegas odds and the regular season do not perform as well. However, all of our optimizations defeat the strategy of picking the seeds, which would have scored only 239,256 points in the 1999 pool. For the 1998 tournament, the results are even more striking. The winner of the office pool garnered 218,545 points. This is slightly better than our Las Vegas odds model, which would have scored 213,285 points. However, either the regular season or Sagarin models would have won the pool with 356,969 and 274,572 points, respectively. Apparently, simply picking the seeds would also have won the Packard Pool in 1998, achieving 233,702 points! More interesting to us, however, is the fact that our optimizations on the whole do outperform picking the seeds for a complicated point structure such as that offered by the Packard Pool.

Table 3 Performance in the Packard Pool

	NCAA 1999 Actual/Expected/Standard Deviation of Total Score	NCAA 1998 Actual/Expected/Standard Deviation of Total Score
Pick the Seeds	239,256/–/–	233,702/–/–
Regular Season	274,956/302,964/47,293	356,969/304,181/54,213
Sagarin	405,421/299,062/55,738	274,572/273,333/63,974
Las Vegas Odds	299,942/339,887/62,735	213,285/514,516/138,543

7. Conclusions

We have presented a framework for evaluating alternative prediction strategies for office pools. We have developed new models for evaluating the mean and the variance of the total score in an office pool for reasonably general point structures and probability models, and applied these models to random and Markov tournaments. We have also shown how to optimally select prediction strategies to maximize the expected total score in an office pool. We have applied these models to mens' college basketball tournaments, and have been able to correctly predict the winners in about 58% of the games considered. While the models do not offer great improvement over the simple strategy of picking the seeds when the objective is to maximize the number of correctly predicted winners in a tournament, the models did perform well when a more complicated scoring scheme was considered. This seems sensible to us, in that the value of optimization should increase as the trade-offs induced by a scoring scheme become increasingly complex.

Acknowledgments
The authors thank Ben Cumming for research assistance; Rob Langrell of the *Las Vegas Sun*, Terry Masline of the Wager Information Network and JK Sports Service of Los Angeles for providing us with betting odds; and Bobby Choquette of the Imperial Palace, Las Vegas, for explaining the ins and outs of sports betting. This research was supported by funds from the Yale School of Management faculty research program.

Appendix: Evaluating the Variance of the Total Score for Random and Markov Tournaments

Example A.1. A Random Tournament. Consider again the random tournament with $\pi_{rg}(v_{rg}) = 1\ \forall r, g$, and let σ_k^2 denote the variance of the number of correctly-called wins in a tournament with k rounds. As with the mean number of correct predictions, the variance of the number of correctly-called wins in subtournaments of equal size will be equal across subtournaments, and there is no dependence on prediction strategy (so $\sigma^2_{r-1,2g-1}(\mathbf{v}) = \sigma^2_{r-1,2g}(\mathbf{v}) = \sigma^2_{r-1}$ in Equation 3.8). The variance of $W_{rg}(v_{rg})$ simply equals $2^{-r}(1-2^{-r})$ in the random tournament. Also, note that the outcomes of all games in the subtournament crowned by game

$2g-1$ in round $r-1$ are independent of the outcomes of all games in the subtournament crowned by game $2g$ in round $r-1$ (see Figure 1), rendering $\text{Cov}(T_{r-1,2g-1}(\mathbf{v}), T_{r-1,2g}(\mathbf{v})) = 0$ for the random tournament.

Now, the team v_{rg} predicted to win game g in round r must emerge from either of the subtournaments crowned by games $2g-1$ ($v_{rg} \in \tau(r-1, 2g-1)$) or $2g$ ($v_{rg} \in \tau(r-1, 2g)$) in round $r-1$. Suppose $v_{rg} \in \tau(r-1, 2g-1)$. Then in the random tournament, $W_{rg}(v_{rg})$ is independent of $T_{r-1,2g}(\mathbf{v})$ because the probability that v_{rg} wins game g in round r does not depend upon *any* of the outcomes in the subtournament played by teams from the set $\tau(r-1, 2g)$. Consequently, $\text{Cov}(T_{r-1,2g}(\mathbf{v}), W_{rg}(v_{rg})) = 0$ in this case. Alternatively, if $v_{rg} \in \tau(r-1, 2g)$, then $\text{Cov}(T_{r-1,2g-1}(\mathbf{v}), W_{rg}(v_{rg})) = 0$ by symmetry.

Without loss of generality, then, suppose that $v_{rg} \in \tau(r-1, 2g-1)$ and hence $\text{Cov}(T_{r-1,2g}(\mathbf{v}), W_{rg}(v_{rg})) = 0$. The remaining term to consider from Equation 3.8 is

$$\text{Cov}(T_{r-1,2g-1}(\mathbf{v}), W_{rg}(v_{rg})) = E(T_{r-1,2g-1}(\mathbf{v})W_{rg}(v_{rg})) - \frac{\mu_{r-1}}{2^r}, \quad \text{(A.1)}$$

where we have used the notation μ_{r-1} (see Equation 3.5) to denote the expected number of correct predictions in a random tournament with $r-1$ rounds, and recalled that the probability any round r prediction is correct is given by 2^{-r} in the random tournament. To evaluate $E(T_{r-1,2g-1}(\mathbf{v})W_{rg}(v_{rg}))$, note that whichever team is selected to win game g in round r, that same team must have been selected to win $r-1$ previous games within the subtournament crowned by game $2g-1$ in round $r-1$. Hence, for any game h in round $j < r$ where $v_{jh} = v_{rg}$ in this subtournament, $W_{jh}(v_{jh})$ and $W_{rg}(v_{rg})$ are *dependent*, and

$$\Pr\{W_{jh}(v_{jh}) = 1, W_{rg}(v_{rg}) = 1\} = \Pr\{W_{rg}(v_{rg}) = 1\} = \frac{1}{2^r} \quad (\text{for } v_{jh} = v_{rg}). \quad \text{(A.2)}$$

The outcomes of all other games in this same subtournament are independent of $W_{rg}(v_{rg})$, and thus for $v_{jh} \neq v_{rg}$ we have

$$\Pr\{W_{jh}(v_{jh}) = 1, W_{rg}(v_{rg}) = 1\} = \frac{1}{2^{j+r}} \quad (\text{for } v_{jh} \neq v_{rg}). \quad \text{(A.3)}$$

Substituting Equations A.2–A.3 into Equation 3.2 and taking expectations yields the result

$$E(T_{r-1,2g-1}(\mathbf{v})W_{rg}(v_{rg})) = \frac{r-1}{2^r} + \sum_{j=1}^{r-1} \frac{2^{r-j-1}-1}{2^{j+r}}. \quad \text{(A.4)}$$

Collecting all of the results of this section and returning to Equation 3.8 we obtain the recursion

$$\sigma_r^2 = 2\sigma_{r-1}^2 + \frac{1}{2^r}\left(1 - \frac{1}{2^r}\right) + 2\left\{\frac{r-1}{2^r} + \sum_{j=1}^{r-1} \frac{2^{r-j-1}-1}{2^{j+r}} - \frac{2^{r-1}}{3}\left(1 - \frac{1}{4^{r-1}}\right)\frac{1}{2^r}\right\} \quad \text{(A.5)}$$

for $r = 1, 2, 3, \ldots$ and $\sigma_0^2 = 0$. This solves to yield

$$\sigma_k^2 = \frac{20}{63}2^k - \frac{6k-1}{9}\frac{1}{2^k} - \frac{3}{7}\frac{1}{4^k} \quad \text{for } k = 1, 2, 3, \ldots. \quad \text{(A.6)}$$

Note that as the number of rounds k grows, $\sigma_k^2 \to (20/21)\mu_k = \mu_k/1.05$. The overall variability in the total number of correctly predicted wins is thus slightly less than Poisson for the random tournament.

Example A.2. A Markov Tournament. Establishing a recursion for the variance of the total score $\sigma_{rg}^2(\mathbf{v})$ for a Markov tournament is more difficult. First, note that as in the random tournament, the random variables $T_{r-1,2g-1}(\mathbf{v})$ and $T_{r-1,2g}(\mathbf{v})$ are independent, for the outcome of any game between teams in the set $\tau(r-1, 2g-1)$ is independent of the outcome of any game between teams in the set $\tau(r-1, 2g)$ via the Markov assumption. All of the difficulty emerges in evaluating the last two terms of Equation 3.8.

Without loss of generality, suppose that $v_{rg} = v_{r-1,2g-1}$ for all rounds in the tournament, for one can always relabel the tournament bracket to ensure that this condition holds without changing the assignments of which teams can play which other teams. We have already shown how to compute the probabilities $\omega_{rg}(v_{rg})$ and hence the expected subtournament scores $\mu_{rg}(\mathbf{v})$ for a Markov tournament in Equations 3.7 and 3.3. Define the quantities $b_{rg}^+(\mathbf{v})$ and $b_{rg}^-(\mathbf{v})$ as

$$b_{rg}^+(\mathbf{v}) = E[T_{r-1,2g-1}(\mathbf{v})\pi_{rg}(v_{rg})W_{rg}(v_{rg})], \quad \text{(A.7)}$$

and

$$b_{rg}^-(\mathbf{v}) = E[T_{r-1,2g}(\mathbf{v})\pi_{rg}(v_{rg})W_{rg}(v_{rg})]. \quad \text{(A.8)}$$

With this notation, we have

$$\text{Cov}(T_{r-1,2g-1}(\mathbf{v}), \pi_{rg}(v_{rg})W_{rg}(v_{rg})) = b_{rg}^+(\mathbf{v}) - \mu_{r-1,2g-1}(\mathbf{v})\pi_{rg}(v_{rg})\omega_{rg}(v_{rg}), \quad \text{(A.9)}$$

and

$$\text{Cov}(T_{r-1,2g}(\mathbf{v}), \pi_{rg}(v_{rg})W_{rg}(v_{rg})) = b_{rg}^-(\mathbf{v}) - \mu_{r-1,2g}(\mathbf{v})\pi_{rg}(v_{rg})\omega_{rg}(v_{rg}). \quad \text{(A.10)}$$

Because all the computations of this section are contingent on a given prediction strategy $\mathbf{v}$, we will drop explicit dependence upon v_{rg} and $\mathbf{v}$ in the notation to conserve space except where required to avoid confusion. Consider first the computation of b_{rg}^+ ($= b_{rg}^+(\mathbf{v})$). From Equation 3.2 we have

$$\begin{aligned} b_{rg}^+ &= E\left[\pi_{rg}W_{rg}\sum_{j=1}^{r-1}\sum_{h=2^{r-j}(g-1)+1}^{2^{r-j-1}(2g-1)} \pi_{jh}W_{jh}\right] \\ &= \sum_{j=1}^{r-1}\sum_{h=2^{r-j}(g-1)+1}^{2^{r-j-1}(2g-1)} \pi_{rg}\pi_{jh}\Pr\{W_{rg}=1, W_{jh}=1\} \\ &= \sum_{j=1}^{r-2}\sum_{h=2^{r-j}(g-1)+1}^{2^{r-j-1}(2g-1)} \pi_{rg}\pi_{jh}\Pr\{W_{rg}=1, W_{jh}=1\} \\ &\quad + \pi_{rg}\pi_{r-1,2g-1}\Pr\{W_{rg}=1, W_{r-1,2g-1}=1\}. \end{aligned} \quad \text{(A.11)}$$

Now, note that $\Pr\{W_{rg}=1, W_{r-1,2g-1}=1\} = \Pr\{W_{rg}=1\} = \omega_{rg}$, because $v_{rg} = v_{r-1,2g-1}$, and for a correct prediction to occur in game g of round r, a correct prediction must also occur in game $2g-1$ of round $r-1$. Second, note that

$$\begin{aligned} &\Pr\{W_{rg}=1, W_{jh}=1\} \\ &= \Pr\{W_{rg}=1, W_{r-1,2g-1}=1, W_{jh}=1\} \\ &= \Pr\{W_{rg}=1 \mid W_{r-1,2g-1}=1, W_{jh}=1\}\Pr\{W_{r-1,2g-1}=1, W_{jh}=1\} \\ &= \frac{\omega_{rg}}{\omega_{r-1,2g-1}}\Pr\{W_{r-1,2g-1}=1, W_{jh}=1\}. \end{aligned} \quad \text{(A.12)}$$

The last result follows since W_{rg} is conditionally independent of W_{jh} given that $W_{r-1,2g-1}=1$. Inserting these results into equation A.11 we obtain the recursion

$$\begin{aligned} b_{rg}^{+} &= \frac{\pi_{rg}}{\pi_{r-1,2g-1}} \frac{\omega_{rg}}{\omega_{r-1,2g-1}} \sum_{j=1}^{r-2} \sum_{h=2^{r-j}(g-1)+1}^{2^{r-j-1}(2g-1)} \pi_{r-1,2g-1}\pi_{jh} \\ &\quad \times \Pr\{W_{r-1,2g-1}=1, W_{jh}=1\} + \pi_{rg}\pi_{r-1,2g-1}\omega_{rg} \\ &= \frac{\pi_{rg}}{\pi_{r-1,2g-1}} \frac{\omega_{rg}}{\omega_{r-1,2g-1}} (b_{r-1,2g-1}^{+} + b_{r-1,2g-1}^{-}) + \pi_{rg}\pi_{r-1,2g-1}\omega_{rg}. \quad \text{(A.13)} \end{aligned}$$

It remains to evaluate b_{rg}^{-}. While team $v_{rg} \in \tau(r-1, 2g-1)$, all of the teams picked to win game h of round $j < r$ in Equation A.8 are in the set $\tau(r-1, 2g)$. In this case we note that

$$\begin{aligned} &\Pr\{W_{rg}=1, W_{jh}=1\} \\ &\quad = \Pr\{W_{rg}=1 | W_{r-1,2g-1}=1, W_{jh}=1\} \Pr\{W_{r-1,2g-1}=1, W_{jh}=1\} \\ &\quad = \sum_{\ell \in \phi(v_{rg}, r)} p_{v_{rg}, \ell} \Pr\{V_{r-1,2g}=\ell | W_{jh}=1\} \omega_{r-1,2g-1}\omega_{jh}. \quad \text{(A.14)} \end{aligned}$$

For both v_{rg} and v_{jh} to be correct predictions, (i) team v_{rg} must win game $2g-1$ in round $r-1$, (ii) team v_{jh} must win game h in round j, and (iii) in game g of round r, team v_{rg} must defeat the victor of game $2g$ in round $r-1$ (team $V_{r-1,2g}$). However, the probability that any particular team $\ell \in \tau(r-1, 2g)$ wins game $2g$ must be assessed conditionally on the event $W_{jh}=1$. In stating this result, we note that the random variables $W_{r-1,2g-1}$ and W_{jh} are independent (because they refer to games in disjoint subtournaments). The conditional probabilities $\Pr\{V_{r-1,2g}=\ell | W_{jh}=1\}$ themselves can be evaluated via Equation 3.7 after setting $p_{v_{jh}, m}=1$ for all $m \in \tau(j, h) \backslash v_{jh}$ (and $p_{m, v_{jh}}=0$ for all $m \in \tau(j, h) \backslash v_{jh}$). Substituting into Equation A.8 we have

$$\begin{aligned} b_{rg}^{-} &= \sum_{j=1}^{r-1} \sum_{h=2^{r-j-1}(2g-1)+1}^{2^{r-j}g} \pi_{rg}\pi_{jh} \Pr\{W_{rg}=1, W_{jh}=1\} \\ &= \sum_{j=1}^{r-1} \sum_{h=2^{r-j-1}(2g-1)+1}^{2^{r-j}g} \pi_{rg}\pi_{jh} \sum_{\ell \in \phi(v_{rg}, r)} p_{v_{rg}, \ell} \\ &\quad \times \Pr\{V_{r-1,2g}=\ell | W_{jh}=1\} \omega_{r-1,2g-1}\omega_{jh}. \quad \text{(A.15)} \end{aligned}$$

Substituting Equations A.13 and A.15 into Equations A.9 and A.10, and further substituting into Equation 3.8 yields a recursion for calculating the variance of the total score $\sigma_{rg}^2(\mathbf{v}) = \mathrm{Var}(T(\mathbf{v}))$ for any given prediction strategy $\mathbf{v}$. Note that in using this recursion, the initialization $b_{rg}^{+} = b_{rg}^{-} = 0$ for $r = 0, 1$ (and associated g values) must be used. Also note that the probabilities ω_{rg} and conditional probabilities $\Pr\{V_{r-1,2g}=\ell | W_{jh}=1\}$ are built recursively, easing the implementation of this computational scheme.

References

Atkins, L. 1998. Sports betting is the other "March Madness." *The Detroit News*, Op-Ed Page, March 28.

Barnhouse, W. 1999. NCAA tournament a scandal-in-waiting: Bet on it. *Fort Worth Star-Telegram*, Sports Section, January 17.

Bradley, R. A., M. E. Terry. 1952. Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* **39** 324–345.

Breiter, D. J., B. P. Carlin. 1997. How to play office pools if you must. *Chance* **10** 5–11.

Carlin, B. P. 1996. Improved NCAA basketball tournament modeling via point spread and team strength information. *Amer. Statist.* **50** 39–43.

David, H. A. 1959. Tournaments and paired comparisons. *Biometrika* **46** 139–149.

Edwards, C. T. 1996. Double-elimination tournaments: Counting and calculating. *Amer. Statist.* **50** 27–33.

——. 1998. Non-parametric procedure for knockout tournaments. *J. App. Statist.* **25** 375–385.

Horen, H., R. Riezman. 1985. Comparing draws for single elimination tournaments. *Oper. Res.* **33** 249–262.

Knuth, D. E. 1987. A random knockout tournament. *SIAM Rev.* **29** 127–129.

Schwertman, N. C., T. A. McCready, L. Howard. Probability models for the NCAA regional basketball tournaments. *Amer. Statist.* **45** 35–38.

——., K. L. Schenk, B. C. Holbrook. More probability models for the NCAA regional basketball tournaments. *Amer. Statist.* **50** 34–38.

Stern, H. 1991. On the probability of winning a football game. *Amer. Statist.* **45** 179–183.

Accepted by Hau L. Lee; received December 8, 1999. This paper was with the authors 1 month for 1 revision.