A Characterization of the Degree of Weak and Strong Links in Doubles Sports

Paramjit S. Gill and Tim B. Swartz *

Abstract

This paper proposes a model that characterizes the degree to which a doubles sport (i.e. two team members) is a weak or a strong link game. The model is applied to the sport of pickleball where interest is focused on the doubles version of the sport. As a byproduct of the analysis, individual player rankings are obtained.

Keywords: Bayesian analysis, Markov chain Monte Carlo, Pickleball, Regularized regression, WinBUGS.

*Paramjit Gill is Associate Professor, Irving K. Barber School of Arts and Sciences, University of British Columbia Okanagan, 3333 University Way, Kelowna BC, Canada V1V1V7. Tim Swartz is Professor, Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby BC, Canada V5A1S6. Swartz has been partially supported by grants from the Natural Sciences and Engineering Research Council of Canada. The authors thank the Associate Editor and two reviewers for their comments on an earlier version of the manuscript. The authors also thank Dave McLure, Administrator and Data Curator of the Men's League at the Pickleball Kelowna Club for help with data management.

1 Introduction

In team sports, there is some interest in knowing the degree to which a sport is a weak or a strong link game. For example, it is well accepted that basketball is a strong link game. In the NBA, it is difficult for teams to be successful without "superstars" (strong links). To win an NBA championship without a superstar, one would need to go back to the 2003/2004 season where the Detroit Pistons were arguably a team without a superstar.

The topic of weak and strong link sports has received minimal attention in the literature. Anderson and Sally (2013) have devoted two entertaining chapters (chapters 8 and 9) to the topic where it is argued that soccer is a weak link game. The idea is that a weak soccer player (especially a defender) can be exploited and goals can be scored against the player's team. Since goal scoring is rare in soccer (less than three goals per game on average in top European leagues), the mistakes caused by a weak link player are often detrimental to the team. By replacing the weakest player with a better player, it is possible for team performance to improve considerably. Anderson and Sally (2013) provide graphical displays of team performance versus player quality.

Novet (2017) provided a weak/strong link analysis to the sport of hockey based on data from the National Hockey League (NHL). Novet (2017) concluded that hockey, like basketball is a strong link game. Similar to Anderson and Sally (2013), Novet (2017) provided two regressions; total team points versus the strength of the team's weak link player (worst player) and total team points versus the strength of the team's strong link player (best player). Novet also regressed total team points against measures of talent distribution such as the Gini coefficient.

From the point of view of team composition, it is important to know whether a particular sport is a weak or a strong link game. For example, would it be better for a team to spend its money to replace their three weakest players with average players or would the team be better served by spending its money on a single superstar? In a non-sports setting, Gladwell (2016) extends the weak/strong link concepts in Anderson and Sally (2013) to educational philanthropy. Gladwell makes the case that financial contributions to small (weak link) programs make a greater difference to society than contributions to large (strong link) programs.

In this paper, we attempt to formalize the notion of weak and strong links in doubles sports. In particular, we develop a regression model where a single parameter provides the degree to which the sport is a weak or a strong link sport. Our measure lies on the scale from 0 to 1 where 0 represents a completely weak link sport (total reliance on the lesser quality teammate), 1 represents a completely strong link sport (total reliance on the higher quality teammate), and 0.5 denotes that the overall quality of a team is the mean of the individual qualities of the two teammates. The proposed model contains non-standard covariates and synchronicity terms that are treated using regularized regression.

In Section 2, we propose a continuous measure of the weak/strong link characterization with an emphasis on parameter interpretability in the case of two players per team (i.e. doubles). In Section 3, we introduce a regression model that contains the weak/strong link measure as a parameter. As a byproduct, the model contains parameters that provide player rankings. The model is set in a Bayesian framework where special paired player relationships are introduced and are handled via regularization. Computational issues associated with the model are discussed. The model is well-suited for analysis using the WinBUGS software package (Spiegelhalter et al. 2003). In section 4, we apply the methodology to doubles pickleball, a racquet sport with growing participation rates. We conclude with a short discussion in Section 5.

2 Weak and Strong Link Formulation

In sports applications, there is no shortage of papers that consider models of the form

$$\mathrm{E}(d) = q_i - q_j \tag{1}$$

where $\mathrm{E}(d)$ is the expected point differential between teams i and j, q_i describes the strength of team i and q_j describes the strength of team j. There are often variations to the model (1) such as the inclusion of a term for the home team advantage and sometimes a link function is imposed on $\mathrm{E}(d)$ as is done in generalized linear models. Another variation involves modelling the points scored by one team rather than the point differential. The q term may also consist of components involving individual players. For example, models of the above type have been considered in basketball (Fearnhead and Taylor 2011), hockey (Macdonald

2011), soccer (Karlis and Ntzoufras 2000) and cricket (de Silva, Pond and Swartz 2001).

In (1), q represents team strength, and we would like to express q in terms of its component players to explore the weak/strong link aspect of the sport. In doubles sports (i.e. teams with two players), we denote θ_1 and θ_2 as the individual qualities of the two players. Accordingly, $\min(\theta_1, \theta_2)$ is the strength of the lesser quality teammate and $\max(\theta_1, \theta_2)$ is the strength of the higher quality teammate. With a weight function $w \in (0, 1)$, the quality (team strength) of the two-player team is given by

$$q = w \max(\theta_1, \theta_2) + (1 - w) \min(\theta_1, \theta_2) \tag{2}$$

where the parameter w describes the weak/strong link relationship. When $w = 0$, we have a completely weak link game; team strength $q = \min(\theta_1, \theta_2)$ is directly reliant on the lesser quality teammate. When $w = 1$, we have a completely strong link game; team strength $q = \max(\theta_1, \theta_2)$ is directly reliant on the higher quality teammate. When $w = 0.5$, the overall strength of the team is $q = (\theta_1 + \theta_2)/2$, the mean of the individual qualities of the two teammates. Therefore, w provides a simple interpretation as to the degree to which the sport is a weak or strong link game.

Introducing some additional notation, let θ_1 and θ_2 be the individual qualities of the teammates on team A, and let θ_3 and θ_4 be the individual qualities of the teammates on team B. Further, let γ_{ij} denote a special synchronicity term involving players i and j on the same team such that their individual effects are not additive. For example, i and j may play better or worse together than the "sum of their parts". When γ_{ij} is positive (negative), i and j play particularly well (poorly) as teammates. Putting these ideas together, we propose the following weak/strong link model for doubles sports

$$\mathrm{E}(d) = w(\max(\theta_1, \theta_2) - \max(\theta_3, \theta_4)) + (1 - w)(\min(\theta_1, \theta_2) - \min(\theta_3, \theta_4)) + \gamma_{12} - \gamma_{34} \tag{3}$$

where d is the score differential by which team A defeats team B.

Alternatively, we can express model (3) as

$$\mathrm{E}(d) = \begin{cases} w(\theta_2 - \theta_4) + (1-w)(\theta_1 - \theta_3) + \gamma_{12} - \gamma_{34} & \theta_1 < \theta_2,\ \theta_3 < \theta_4 \\ w(\theta_2 - \theta_3) + (1-w)(\theta_1 - \theta_4) + \gamma_{12} - \gamma_{34} & \theta_1 < \theta_2,\ \theta_3 \geq \theta_4 \\ w(\theta_1 - \theta_4) + (1-w)(\theta_2 - \theta_3) + \gamma_{12} - \gamma_{34} & \theta_1 \geq \theta_2,\ \theta_3 < \theta_4 \\ w(\theta_1 - \theta_3) + (1-w)(\theta_2 - \theta_4) + \gamma_{12} - \gamma_{34} & \theta_1 \geq \theta_2,\ \theta_3 \geq \theta_4 \end{cases} \tag{4}$$

What makes model (3)/(4) unusual from a linear models perspective is that the right hand terms do not follow the typical pattern βx where β is an unknown parameter and x is a known covariate. Rather, in model (3)/(4), both w and the θ_i are unknown parameters. Also, (4) has a piecewise representation which is nonstandard in traditional linear models. To address these model features and a sparsity issue involving the γ_{ij} terms, we set the problem in a Bayesian framework in Section 3.

3 Modelling and Computation

Referring to model (3)/(4), we first require a distribution for the error term associated with the score differential d. Tentatively, we assign a Normal$(0, \sigma^2)$ distribution to the error term associated with the score differential d. With this error term, we may specify the hyperparameter σ taking into account the range of score differentials. Another possibility (and the one which we adopt in Section 4) is the vague prior specification $\sigma^2 \sim$ Inverse-Gamma$(1, 1)$. We also note that the model (3)/(4) does not contain an intercept term. This is appropriate when there is no systematic advantage given to one of the two teams such as a home court advantage or a serving advantage. Although we assume that player form does not change over time, it is possible to weight the error term σ so that recent matches receive more weight.

Apriori, we have no knowledge whether the sport in question is a weak or a strong link sport. We therefore assign the prior specification $w \sim$ Uniform$(0, 1)$.

Given m players, the strength parameters $\theta_1, \ldots, \theta_m$ suffer from non-identifiability. For example, it is immediate that the likelihood is invariant if θ_i is substituted with $\theta_i + k$ for any constant k and for all players $i = 1, \ldots, m$. The non-identifiability is managed by assigning independent priors $\theta_i \sim$ Normal$(0, \sigma_\theta^2)$. This parameterization has several advantages. First, the zero-mean specification provides a convenient interpretation as players are classified according to $\theta < 0$ (below average) and $\theta > 0$ (above average). Second, provided σ_θ is not too large, the prior specification keeps the estimation of θ_i rooted about zero, and does not allow θ_i to grow large as may happen with the non-identifiability issue. We have used $\sigma_\theta^2 \sim$ Inverse-Gamma$(1, 1)$. Alternatively, non-identifiabilities of this type are sometimes handled by imposing a constraint such as $\sum_{i=1}^{m} \theta_i = 0$ (Ntzoufras 2009). Thirdly, the common mean for the θ_i's provides a shrinkage advantage where players who have not competed in many matches are not characterized as excessively weak nor strong.

The model may be further expanded to include a gender-specific variance component alternative to σ_θ^2 involving male and female players in mixed doubles. This may be helpful to rank male and female players separately and to compare the heterogeneity of abilities.

A byproduct of the estimation of the θ_i terms is a ranking of players in terms of their ability. This differs from the traditional ranking method used in doubles pickleball which is described in Section 4.

In model (3)/(4), the γ_{ij} are the adjustment parameters (interaction terms) that describe the special relationship between players i and j. It may not be the case that the team strength involving players i and j is adequately described by (2). We expect these sorts of special relationships to be rare, and given that there are $m(m-1)/2$ parameters γ_{ij}, $i < j$, it would be advantageous if many of the $\gamma_{ij} = 0$. Fortunately, regularized regression is a popular technique that forces many of the γ terms equal to zero, and therefore reduces the dimensionality of the parametrization. Only outstandingly large or small relationships characterized by the γ's are assigned non-zero values. Although regularization is prominent in a classical context (Hastie, Tibshirani and Friedman 2001), regularization also has a Bayesian analogue. One version of Bayesian regularization is L_1 regularization which is carried out using the Double-Exponential(λ) prior with probability density function $\pi(\gamma) \propto \exp\{-\lambda \sum_{i \neq j} |\gamma_{ij}|\}$ where $\lambda > 0$ is a tuning parameter. The larger the value of λ, the greater the "penalty" and the fewer γ_{ij} are assigned non-zero values. Regularization in a Bayesian framework via Markov chain Monte Carlo (MCMC) permits straightforward inference as the generated parameters may be used to infer relevant posterior distributions.

One complication concerning the γ_{ij} parameters is that the estimation is nonsensical if i and j have not played together. Moreover, if i and j have only played together a single time then model (3)/(4) suggests that γ_{ij} will be overfit. Accordingly, we let n_{ij} denote the number of times that i and j have played together. We therefore modify the prior distribution

$$\gamma_{ij} \sim \begin{cases} \text{Double-Exponential}(\lambda) & n_{ij} \geq 2 \\ I_0 & n_{ij} = 0, 1 \end{cases} \tag{5}$$

where I_0 denotes the distribution with point mass at zero.

Our Bayesian model is conveniently implemented in the WinBUGS programming environment (Spiegelhalter et al. 2003). One of the advantageous of WinBUGS is that the user does not have to write detailed MCMC code. Instead, the user only has to provide the model specification and MCMC calculations are done in the background. Our WinBUGS program consists of fewer than 40 lines of code and is available from the authors upon request. Lykou and Ntzoufras (2011) provide a tutorial on WinBUGS programming and devote a section to a particular implementation of regularization.

3.1 Model Selection

A popular method for model selection uses the predictive performance criterion proposed by Laud and Ibrahim (1995). Given a finite number of candidate models, the criterion is based on the predictive performance of a model in terms of its ability to predict a replicate of the data.

Let $\mathbf{Y}_{\text{pred}}$ denote a replicate of the observed data $\mathbf{Y}_{\text{obs}}$. That is, $\mathbf{Y}_{\text{pred}}$ is generated from a predictive distribution. The particular predictive distribution is the one whose covariates match up to the covariates of $\mathbf{Y}_{\text{obs}}$. In this way, $\mathbf{Y}_{\text{pred}}$ is a replicate of $\mathbf{Y}_{\text{obs}}$. The predictive density of $\mathbf{Y}_{\text{pred}}$ under model M is

$$f^{(M)}(\mathbf{Y}_{\text{pred}}|\mathbf{Y}_{\text{obs}}) = \int f(\mathbf{Y}_{\text{pred}}|\eta^{(M)})\ \pi(\eta^{(M)}|\mathbf{Y}_{\text{obs}})\, d\eta^{(M)} \tag{6}$$

where $\eta^{(M)}$ denotes all the parameters under model M, $\pi(\eta^{(M)}|\ \mathbf{Y}_{\text{obs}})$ is the posterior density and $f(\mathbf{Y}_{\text{pred}}|\ \eta^{(M)})$ is the density of a predicted (or future) value. The model selection criterion called the expected predictive deviance (EPD), chooses the model M with the smallest value of

$$\mathrm{E}^{(M)}[d(\mathbf{Y}_{\text{pred}}, \mathbf{Y}_{\text{obs}})|\mathbf{Y}_{\text{obs}}] \tag{7}$$

where $d(\mathbf{Y}_{\text{pred}}, \mathbf{Y}_{\text{obs}})$ is a discrepancy function and the expectation is with respect to the predictive distribution given by (6). One discrepancy measure that we use is the absolute prediction error: $d(\mathbf{Y}_{\text{pred}}, \mathbf{Y}_{\text{obs}}) = ||\mathbf{Y}_{\text{pred}} - \mathbf{Y}_{\text{obs}}||$. It is straightforward to estimate (7) as part of MCMC sampling. In each loop of an MCMC run, $\mathbf{Y}_{\text{pred}}$ is generated and $d(\mathbf{Y}_{\text{pred}}, \mathbf{Y}_{\text{obs}})$ is calculated. The sample mean of the generated $d(\mathbf{Y}_{\text{pred}}, \mathbf{Y}_{\text{obs}})$ values is then used to estimate the EPD (7).

In the examples that follow, we also compute the model selection diagnostic DIC known as the Deviance Information Criterion (Carlin and Louis 2008). DIC is a counterpart to the Akaike Information Criterion (AIC) and is an immediate byproduct of sampling from the posterior. In WinBUGS, DIC is an option on the *Inference* menu. As with AIC, smaller values of DIC suggest preferred models where differences greater than 3-5 are typically regarded as meaningful.

4 Example: Doubles Pickleball

The sport of pickleball was invented in the summer of 1965 by Joel Pritchard, Bill Bell and Barney McCallum of Bainbridge Island, Washington, USA (http://ifpickleball.org/history-of-pickleball-and-the-ifp/). The sport has evolved from humble beginnings into a popular sport played throughout the US and Canada. The game is growing internationally as well, with Mexico and many European and Asian countries introducing pickleball clubs. It is a sport that is related to tennis, badminton and ping-pong. Pickleball is played both indoors and outdoors on a court which is the same size as a doubles badminton court (measuring 20 $\times$ 44 feet). The same court is used for both singles and doubles play. The net height is 36 inches at the sidelines and 34 inches in the middle. The court is divided into right and left service areas with a 7-foot non-volley zone in front of the net (referred to as the "kitchen"). Courts can be constructed specifically for pickleball or they can be converted using existing tennis or badminton courts. Figure 1 shows the pickleball court layout.

In a pickleball game, the team first reaching the score of 11 points is declared the winner.

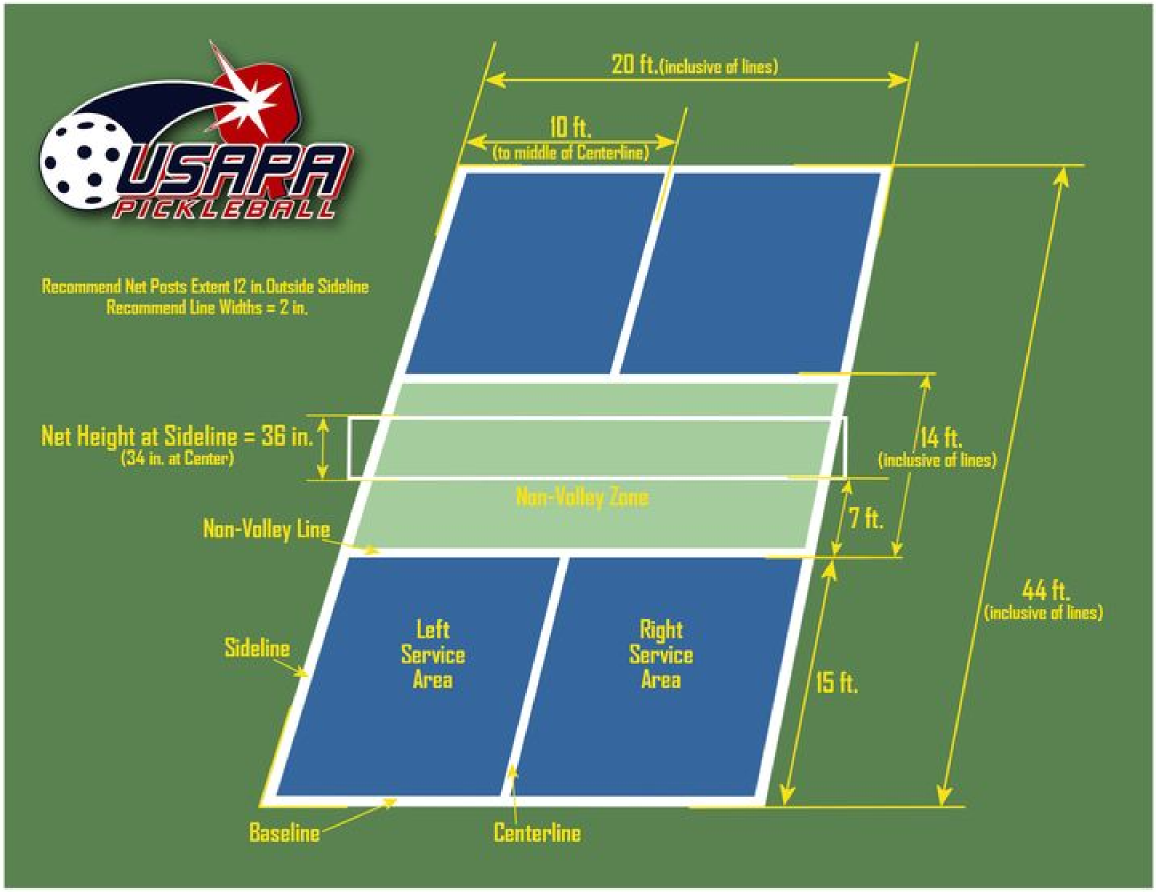

Figure 1: Pickleball court layout (www.usapa.org).

A point is scored when the serving team wins the rally. The alternation between serving and receiving teams is described at www.usapa.org. The scoring system has implications for modeling the score differential d where we had previously assigned $d \sim \text{Normal}(0, \sigma^2)$. To partially account for the scoring system, we truncate the $\text{Normal}(0, \sigma^2)$ distribution described in Section 3 to the interval $(-11.1, 11.1)$. We note that the truncated normal still lacks some realism since score differentials are discrete, and in a no-draw contest such as pickleball, the outcome $d = 0$ is impossible.

The Pickleball Kelowna Club (in Kelowna, British Columbia) has more than 500 members of varying skill levels. The club organizes various doubles leagues from May to October.

Detailed match data are kept for two of the leagues called the "Monday Mixed Ladder" and the "Tuesday Men's League".

For the two leagues, players register weekly through an online system called TrackitHub (www.trackithub.com). Player groups are then formed by the software so as to have groups of similar skill levels. Therefore, players are compared within homogeneous strata and not across all of the club players.

In a typical pickleball doubles league, each player is awarded a percent performance index. This index is the percent of maximum possible points earned by a player in all the games played that he played during a session. For example, if a player participated in four doubles matches and the game scores were 11-9, 4-11, 11-3 and 8-11 in favour of his teams, then his total score for that day is $(11+4+11+8) = 34$ and his percent performance index is $(34/44)100\% = 77.3\%$. The player's performance index over all sessions is a measure of the player's ability. It is evident that a major failing of the performance index is that it does not directly take into account the abilities of his partner and his opponents. Recall that the parametrization (3) characterizes the strength of all four players in a doubles game. Another issue with the performance index is that it has high variability when a player has participated in a small number of matches. On the other hand, in the Bayesian model, the estimates of θ_i shrink towards the mean value of zero for players with a small number of games.

4.1 Tuesday Men's League

In the Tuesday Men's League, the players who sign up ($\sim$60), are mostly divided into groups of 8 players each. Within each group, there are $\binom{8}{2} = 28$ possible pairs of players where each pair plays exactly one game within a session. The player rotation is arranged so that each player plays twice against each of the other 7 players. Table 1 provides one such configuration (a special resolvable balanced incomplete block design with 8 treatments and block size 4) where the integers from 1 to 8 indicate players' labels.

Since the total number of players signing up on a given day may not be a multiple of 8, some groups consist of five or six players. Typically, five-player groups play games to 15 points and each player plays four games in four of the five rounds. Rather than throwing away data from five-player groups, we rescaled the scores out of 11 points. For example, if

round 1	round 2	round 3	round 4	round 5	round 6	round 7
26 v 18	45 v 16	38 v 14	67 v 13	58 v 36	12 v 35	23 v 46
34 v 57	28 v 37	56 v 27	48 v 25	17 v 24	47 v 68	15 v 78

Table 1: An example of a session involving 8 players in doubles pickleball. In a session, there are 7 rounds each consisting of two games.

a game score is 15-10, we recorded this as 11-7.3.

For the Tuesday Men's League, we have data from 1278 matches recorded for the 2017 season, with $m = 135$ players. Table 2 shows posterior summaries for the parameters corresponding to three Bayesian models. The three models that we considered involve decreasing penalization (i.e. decreasing the tuning parameter λ) in the regularization of the synchronicity terms γ_{ij}. The resulting estimates were based on 5,000 simulations in WinBUGS with a burn-in of 5,000 iterations. Under these inputs, standard convergence diagnostics appeared satisfactory. The simulation procedure required approximately two minutes of computation on an ordinary laptop computer. We observe that the posterior mean of the weak/strong link parameter $w \approx 0.87$ for all three models. This indicates that doubles pickleball is a strong link game, an observation that coincides with our intuition. In doubles pickleball, the stronger player may assert himself to execute more of the shots and take control of the game. We also observe that the parameter estimates σ and σ_θ are similar under all three models suggesting that regularization does not have a great effect in this dataset. However, we do see that the effect of regularization (although small) behaves in the expected way. For example, for smaller values of the tuning parameter λ, "unusual" data are better modeled such that σ decreases (i.e. there is less error variability).

In Table 3, we provide diagnostics of fit for the three proposed models. We observe that the simple unregularized model (i.e. no γ_{ij} terms) does not fit quite as well as the regularized models (cf. EPD and DIC). As expected, as λ increases (i.e. a greater penalty on the synchronicity terms γ_{ij}), a smaller number of "special" effects λ_{ij} become active. However, in comparing the three models, our preferred model is the simple Model 1 (without regularization). A reason for this is based on the prior distribution (5) used in regulatization. In the Tuesday Men's League dataset, there are only 1502 instances amongst the $m(m - 1)/2 = 9045$ pairs of players where $n_{ij} \geq 2$. It is difficult for us to believe that special

Model	Parameter	Post Mean	Post Std Dev
Model 1	w	0.87	0.10
(unregularized;	σ	5.08	0.09
i.e. $\lambda \to \infty$)	σ_θ	0.88	0.17
Model 2	w	0.86	0.11
(regularized;	σ	5.07	0.09
$\lambda = 100.0$)	σ_θ	0.91	0.13
Model 3	w	0.87	0.10
(regularized;	σ	4.98	0.09
$\lambda = 2.0$)	σ_θ	0.90	0.16

Table 2: Posterior summaries of the model parameters for the Tuesday Men's League.

synchronicity exists in such a high percentage (1359/1502)100% = 90.5% of the pairings. We have defined the existence of special synchronicity for players i and j when $|\gamma_{ij}| > 0.05$. Moreover, an "optimized" value of the tuning parameter λ can be obtained by introducing a prior distribution. We have done this (i.e. $\lambda \sim$ Inverse-Gamma(1,1)) where the resultant posterior mean $\hat{\lambda} = 1.7$ is close to the tuning value $\lambda = 2$ utilized in Model 2. We prefer the simplicity of Model 1 over the small gains in fit obtained by Model 2 and Model 3. However, we believe that the consideration of regularization is a useful exercise, as we have demonstrated that the computations are feasible, and it may be the case that regularization proves useful in other pickleball leagues or other doubles sports.

Model	EPD	DIC	No. of $\gamma_{ij} \neq 0$
Model 1	5.9	8047.0	0
Model 2	5.9	8045.0	10
Model 3	5.7	8031.0	1359

Table 3: Model fit diagnostics for the Tuesday Men's League.

We also compare the ranking procedures provided by our estimated θ_i terms and the traditional performance index. In Figure 2, we provide a scatterplot of the player skill parameters θ_i based on Model 1 versus the traditional pickleball peformance index, $i = 1, \ldots, 135$. Although there is good agreement between the two measures (correlation =

0.82), there are cases where disagreements exist. For example, there is a player (far right in Figure 2) who is a little better than average ($\theta_i = 0.56$) but is the best player in the league according to the player performance index. In this case, the player only played 7 games, winning five and losing two. We believe that such a player should not be rated extremely high since there is little information. We view this as a feature of the methodology as ability estimates have a shrinkage property when the data are sparse.

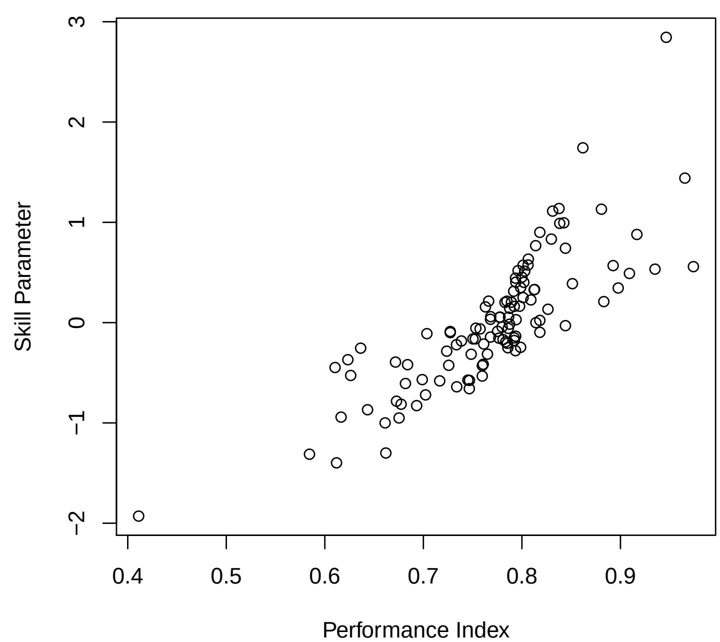

Figure 2: Scatterplot of the player skill parameters θ_i based on Model 1 versus the traditional pickleball peformance index.

4.2 Monday Mixed League

Both men and women players compete in the Monday Mixed League. Players sign up every week and are divided into groups of five players. Each player plays four out of five games where they sit out one of the games. That way, each player pairs up with every other player exactly once and plays exactly two games against every other player. Whenever the total number of players is not an exact multiple of five, some of the sessions consist of four players who each play three games and a game is played to 15 points. For such games, we rescale the team scores out of 11.

For the Monday Mixed League, we have data from 846 matches recorded during the 2017 summer season with 166 active players (84 men and 82 women) who have each played at least four matches.

We fit two Bayesian models to these data. Model 1a corresponds to the best fitting Model 1 in the Tuesday Men's League. This is the model without regularization. Model 1b is an analogue of Model 1a where we have introduced gender-specific variance components $\sigma^2_{\theta_M}$ and $\sigma^2_{\theta_F}$ as described in Section 3. In Table 4, we provide model summaries for Model 1a and Model 1b. Although the fit diagnostics are similar for both models, we prefer the more complex Model 1b. We prefer Model 1b since there is some evidence of gender-specific variance components. Specifically, the men appear to have skills that are more variable than the women. Based on the first author's personal experience, not only do some highly skilled men from the Tuesday Men's League compete in the Monday Mixed League, but there are also beginner men who avoid the Tuesday Men's League and only compete in the Monday Mixed League. The parameters σ and w in Models 1a and 1b are consistent with the estimates from Model 1 in the Tuesday Men's League (Table 2). However, we observe that weak/strong link parameter $w = 0.80$ in Model 1b is slightly smaller than $w = 0.87$ found in Model 1 in the Tuesday Men's League (Table 2). We conjecture that the slight decrease in w for the Monday Mixed League may be that men are playing in a more "polite" fashion where strong players do not dominate the ball in the same way that they do in the Tuesday Men's League.

5 Discussion

We have provided a model which allows us to characterize the degree to which doubles pickleball is a weak or a strong link game. From the analysis of data based on a large number of games from two leagues, we learn that doubles pickleball is a strong link game. Of course, with other leagues of different strengths, the style of play may differ and result

Model	Parameter	Post Mean	Post Std Dev	EPD	DIC
Model 1a	w	0.72	0.13		
	σ	5.16	0.11	6.0	5408.0
	σ_θ	1.47	0.18		
Model 1b	w	0.80	0.16		
	σ	5.18	0.12	6.0	5410.0
	σ_{θ_M}	1.60	0.29		
	σ_{θ_F}	1.15	0.32		

Table 4: Posterior summaries of the model parameters for the Monday Mixed League.

in different estimates of the weak/strong link parameter w.

As a byproduct of the analysis, player rankings are obtained which may be superior to the traditional performance indices. A limitation of our approach exists when there is not sufficient mixing of players during matches. Without mixing, a weak player A belonging to a strong group may have a lower ranking than a strong player B belonging to a weak group even though A may otherwise be a better player than B.

The applicability of the approach to other doubles sports appears limited to amateur events where team composition is randomized by league organizers. For example, although professional tennis has doubles competitions, partners do not tend to mix greatly. Pairs of top-end tennis players tend to stay together for long periods of time. Without sufficient mixing, it is difficult to separate the component skill levels θ_1 and θ_2 of the teammates.

In the case of teams consisting of more than two players, it is even more difficult to think of tournament situations where team members are randomized over games. If there were such a scenario, then it may be possible to generalize the model in this paper to account for the strength of all players. Instead of team strength characterized by equation (2), it may be possible to obtain some formulation that is a function of the order statistics $\theta_{(i)}$.

6 References

Anderson, C. and Sally, D. (2013). *The Numbers Game: Why Everything You Know about Soccer is Wrong*, Penguin Books: New York.

Carlin, B.P. and Louis, T.A. (2008). *Bayesian Methods for Data Analysis*, Chapman and Hall/CRC:

New York.

de Silva, B.M., Pond, G.R. and Swartz, T.B. (2001). Estimation of the magnitude of victory in one-day cricket. *The Australian and New Zealand Journal of Statistics*, 43, 259-268.

Fearnhead, P. and Taylor, B.M. (2011). On estimating the ability of NBA players. *Journal of Quantitative Analysis in Sports*, 7(3), Article 11.

Gladwell, M. (2016). My little hundred million. *Revisionist History*, Podcast - Season 1, Episode 6 accessed on July 13/17 at http://revisionisthistory.com/episodes/06-my-little-hundred-million

Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Springer: New York.

Karlis, D. and Ntzoufras, I. (2000). On modelling soccer data. *Student*, 3, 229-244.

Laud, P., and Ibrahim, J. (1995). Predictive model selection, *Journal of the Royal Statistical Society B*, 57, 247-262.

Lykou, A. and Ntzoufras, I. (2011). WinBUGS: a tutorial. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3, 385-396.

Macdonald, B. (2011). A regression-based adjusted plus-minus statistic for NHL players. *Journal of Quantitative Analysis in Sports*, 7(3), Article 4.

Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*, Volume 698 of Wiley Series in Computational Statistics, Wiley: New York.

Novet, A. (2017). Strong and weak links: talent distribution within teams. *Hockey Graphs*, accessed online July 13/17 at https://hockey-graphs.com/2017/03/14/strong-and-weak-links-talent-distribution-within-teams/#more-16199

Spiegelhalter, D.J., Thomas, A., Best, N. and Lund, D.J. (2003). *WinBUGS (Version 1.4.3) User Manual.* MRC Biostatistics Unit, Cambridge: UK.