Paired comparisons analysis: an axiomatic approach to ranking methods*

Julio González-Díaz†
Department of Statistics and Operations Research
University of Santiago de Compostela

Ruud Hendrickx
CentER and Department of Organization and Strategy
Tilburg University

Edwin Lohmann
CentER and Department of Econometrics and Operations Research
Tilburg University

Published in Social Choice and Welfare (2013)
Published version available at http://www.springer.com

DOI 10.1007/s00355-013-0726-2

Abstract

In this paper we present an axiomatic analysis of several ranking methods for general tournaments. We find that the ranking method obtained by applying maximum likelihood to the (Zermelo-)Bradley-Terry model, the most common method in statistics and psychology, is one of the ranking methods that perform best with respect to the set of properties under consideration. A less known ranking method, generalised row sum, performs well too. We also study, among others, the fair bets ranking method, widely studied in social choice, and the least squares method.

1 Introduction

In a world full of choices and alternatives, rankings are becoming an increasingly important tool to help individuals and institutions make decisions. In this paper we study, from an axiomatic point of view, the classic problem of ranking a series of alternatives when we have information about paired comparisons between them. The set of alternatives and a matrix containing this information are referred to as a *ranking problem* or *general tournament*. Ranking problems appear in a wide variety of situations such as sports, product testing, evaluation of political candidates and policies to be chosen. Because of this, the issue of defining rankings has been studied in various fields and ranking methods based on different

*We thank Miguel Brozos-Vázquez and José Carlos Díaz-Ramos for helpful discussions. We also thank Pavel Chebotarev and an anonymous referee for their comments on earlier versions of the manuscript.

†Corresponding author: julio.gonzalez@usc.es

motivations have been defined. Sport events and in particular chess motivated the seminal work on rankings by Zermelo (1929). Later on, sparked by Arrow's Impossibility Theorem (Arrow 1963), this topic emerged in the context of social choice and voting theory. The theory of rankings has also attracted statisticians and psychologists, who have studied it under the name of paired comparisons analysis. It is worth noting that the literature on rankings and tournaments is somehow split depending on whether the goal is to identify the winner (set of winners) or, as in this paper, to define a more detailed ranking of the alternatives. The recent book by Langville and Meyer (2012) provides an up to date review of different fields where rankings are important, ranging from relatively classic fields of application such as college sports in the United States to human development indexes and more recent applications like rankings of websites (Google PageRank) or movies (Netflix and IMDb).

To be more specific, suppose that we want to rank several alternatives and that we have at our disposal a series of nonnegative matrices R^1, R^2, ..., R^p, where each matrix contains information about some pairwise comparisons between the different alternatives. Each of these matrices may represent the opinion of a given expert/judge, the outcome of a certain poll or voting procedure, the results of a certain round of a competition, the outcomes of the comparisons in different regions/competitions...The entries of each of these matrices are such that $R^\ell_{ij} + R^\ell_{ji} = 1$ if i and j have been compared at stage ℓ and $R^\ell_{ij} + R^\ell_{ji} = 0$ otherwise. The definition of a ranking of the alternatives based on this information can be seen as a problem of adding "pairwise" preferences. To accomplish this, here we suppose that the information of the R^ℓ matrices has already been aggregated into a single nonnegative matrix A (possibly putting different weights on different rounds/judges/experts/regions...). It is probably because of the different interpretations mentioned above of the R^ℓ matrices that the theory of paired comparisons has developed simultaneously in a wide variety of fields and has an important potential for applications.

In this paper a *ranking problem* is represented by a set N consisting of n alternatives and a nonnegative $n \times n$ matrix A with zeros on the diagonal, where A_{ij} is the total score of alternative i against alternative j after their (possibly many) pairwise comparisons. This approach is the usual one in fields such as statistics, psychology and applications to sports. On the other hand, in voting theory, a field that has devoted considerable attention to this topic, a *tournament* is typically defined through weakly complete and asymmetric binary relations, *i.e.*, for each pair of alternatives we know which is the preferred one (and nothing else); there is no measure of intensity of preference. These *binary* tournaments are a particular case of our more general setting, which is able to accommodate the following extra features:[1] i) incomplete tournaments, in which we may not have information about direct confrontations between pairs of alternatives ($A_{ij}+A_{ji}$ may be zero), ii) tournaments in which alternatives may have been compared with each other more than once ($A_{ij} + A_{ji} > 1$),[2]

[1]In matrix form, a binary tournament corresponds to a binary matrix $A \in \{0, 1\}^{n\times n}$ such that for each pair of different alternatives i and j, $A_{ij} + A_{ji} = 1$.

[2]This is specially important, for instance, in testing objects, where each pair of objects may be tested by several experts.

iii) tournaments with ties ($A_{ij} = A_{ji}$) and iv) tournaments in which intensities of preference are present (captured, for instance, by $(A_{ij} - A_{ji})/(A_{ij} + A_{ji})$).

Although ideally we would like to work with ranking problems in which there is information about all possible pairwise comparisons, there are many situations where it is unfeasible to obtain direct information about each pair of alternatives. This may be because there is a high number of alternatives to be ranked or just because it is costly to undergo each pairwise comparison and it is preferable to base the ranking on an incomplete set of comparisons. From a conceptual point of view, whether or not tournaments are restricted to be binary has important implications in defining ranking methods. In a binary tournament all the alternatives have "faced" each other exactly once and simple rules that look at the number of "victories" of each alternative may have good properties; these rules would include, for instance, the well known Borda count and the Copeland methods (Copeland 1951). However, in general tournaments it does not suffice to know how well an alternative has scored. We need to take into account the quality of the "opponents".

Our goal in this paper is to take some of the most relevant ranking methods considered in the different fields and compare them by looking at their performance with respect to a set of (mostly standard) properties. Axiomatic approaches to ranking theory have been taken before in the literature, especially in social choice and voting theory (see, for instance, Chebotarev and Shamis (1998) for a survey). However, most of these contributions mainly deal with binary tournaments. Laslier (1997) presents a thorough analysis of different ranking methods and properties defined for binary tournaments.[3] Another characteristic of the voting theory approach is that it is common to focus on the set of winners (thus, allowing for ties) and the spirit of many of the properties revolves around this possibility. Also within the axiomatic approach, Bouyssou (2004) revisits the main ranking methods in Laslier (1997) and studies their monotonicity properties (responsiveness to the beating relation).[4]

Because of the large amount of ranking methods and properties that have been discussed in the different fields, some selection is needed. Our analysis mainly concentrates on the ranking methods listed below.

- **Scores:** A natural choice for binary tournaments (see Rubinstein (1980) for an axiomatic characterisation).
- **Maximum likelihood:** The most common choice in statistics and psychology (see, for instance, Zermelo (1929) and Bradley and Terry (1952)).

[3]Dutta and Laslier (1999) consider a setting where this last restriction is generalised to allow for intensities, but completeness is still a requirement.

[4]In recent years, the related issue of ranking scientific journals has received a lot of attention in economics (see, for instance, Liebowitz and Palmer (1984) and Palacios-Huerta and Volij (2004)). In this setting, the rankings are defined on the basis of citation matrices, which contain information regarding the number of times each journal has been cited by any other journal. There is a fundamental difference between the two settings. In our setting, a victory of i over j should be seen as something good for i and bad for j. However, when looking at scientific journals, a citation from journal j to journal i should be good for journal i, but not necessarily bad for journal j. Clearly, this cannot be ignored when defining properties of a ranking method and, therefore, it would be inappropriate to include in our axiomatic analysis ranking methods that are based on citation matrices.

- **Fair bets:** A ranking widely studied in social choice and economics (see, for instance, Daniels (1969), Moon and Pullman (1970), Slutzki and Volij (2005) and Slutzki and Volij (2006)).
- **Least squares:** Another common choice in statistics and psychology (see, for instance, Horst (1932), Mosteller (1951), Gulliksen (1956) and Kaiser and Serlin (1978)).
- **Recursive performance and recursive Buchholz:** These ranking methods are the result of a new approach developed in Brozos-Vázquez et al. (2008).
- **Generalised row sum:** A parametric family of ranking methods that on one end generalise the scores and on the other end generalise least squares (see Chebotarev (1994)).[5]

The main contribution of this paper is to study how the above ranking methods perform with respect to a set of properties. This analysis is important not only to get a better understanding of the different ranking methods, but also to learn about the strength and implications of the different properties. Interestingly, maximum likelihood is one of the rankings that perform best with respect to the chosen properties. This is somewhat surprising since, because of its nature, one would expect maximum likelihood to have good statistical properties (and it does, for instance, in terms of asymptotic behaviour), but there is no reason to expect good behaviour with respect to some of the properties we work with. The other method that stands up from our approach is the generalised row sum method.

The rest of the paper is structured as follows. In Section 2 we present the main definitions and ranking methods. In Section 3 we give some background on their origins. Then, in Sections 4-6 we define and discuss several families of properties. Finally, in Section 7 we discuss the results, which are summarized in a table that allows to visually compare the performance of the different ranking methods.

2 Ranking problems and ranking methods

A *ranking problem* is a pair (N, A), where N is a finite set of $n \geq 2$ players and $A \in \mathbb{R}^{n \times n}$ is the *results matrix*. Each A_{ij} represents the aggregate score of player i against j. We assume $A_{ij} \geq 0$ for all $i, j \in N$ and $A_{ii} = 0$ for all $i \in N$.[6] We say that i has *scored* against j if $A_{ij} > 0$ and that i has *beaten* j if $A_{ij} > A_{ji}$. When no confusion arises, we denote a ranking problem (N, A) by A.

We make the standard assumption that the matrix A is irreducible.[7] This means that for every pair of players $i, j \in N$, $i \neq j$, there is a sequence of players $(i = k_0, k_1, \ldots, k_n = j)$ such that, for each $\ell \in \{0, \ldots, n-1\}$, k_ℓ has scored against $k_{\ell+1}$. In Slutzki and Volij (2005), the authors show that, under a natural set of axioms, this restriction can be circumvented

[5] More precisely, they generalise the *aggregate net scores*, which we define in Section 2.

[6] We do not restrict the non-zero entries in A to be natural numbers as in, *e.g.*, Slutzki and Volij (2005). This choice of domain has no impact on the analysis, but allows for simpler definitions of the properties under study.

[7] This ensures that all the rankings methods we discuss in this paper are well defined.

by partitioning the set of players into maximal irreducible components, which they call leagues, analysing each league separately and then recombining the leagues' rankings to get a ranking on the whole set of players.[8]

To each ranking problem (N, A) we associate a (symmetric) *matches matrix* $M(A) = A + A^\top$, where $M_{ij}(A)$ is interpreted as the number of matches between i and j.[9] When no confusion can arise, we denote $M(A)$ by M. For each player $i \in N$, define $m_i = \sum_{j\in N} M_{ij}$ to be the total number of matches played by i, so $m = Me$, where $e \in \mathbb{R}^n$ is the vector $e = (1, \ldots, 1)^\top$. For $i, j \in N$, define $\bar{M}_{ij} = M_{ij}/m_i$ to be the proportion of player i's matches that he plays against j. A ranking problem is called *round-robin* if $M_{ij} = 1$ for all $i, j \in N, i \neq j$; in a round-robin ranking problem each player has played once against any other player.[10]

A *ranking method* φ assigns to each ranking problem (N, A) a weak order $\varphi(A)$ on N (transitive and complete). Given a ranking problem (N, A), a vector $r \in \mathbb{R}^n$ is a *rating vector*, where each r_i is a measure of the performance of player $i \in N$ in the ranking problem.[11] The ranking methods considered in this paper are all induced by rating vectors: for each ranking method φ there is an underlying rating vector r^φ such that the players are ranked according to it, *i.e.*, φ ranks i weakly above j if and only if $r^\varphi_i \geq r^\varphi_j$. A ranking φ is called *flat* on A if $\varphi(A)$ ranks all players equally. Abusing terminology slightly, we also refer to the underlying rating vector $r^\varphi(A)$ as flat. Below, we present the definitions of the ranking methods studied in this paper. We defer their interpretation and motivation until next section.

Scores: The vector of average scores, s, is defined by $s_i = \sum_{j\in N} A_{ij}/m_i$ for all $i \in N$. It follows from the assumption that A is irreducible that $s_i \in (0, 1)$ for all $i \in N$.

Maximum likelihood: The rating of player i is given by $r^{\mathrm{ml}}_i = \log(\pi_i)$, where $\pi \in \mathbb{R}^n$ is the unique and positive solution of the system of non-linear equations given by $\pi^\top e = 1$ and, for each $i \in N$,

$$\pi_i = \frac{m_i s_i}{\sum_{j\in N\setminus\{i\}} \frac{M_{ij}}{\pi_i+\pi_j}}. \tag{2.1}$$

To facilitate the comparison with other ranking methods, when presenting the maximum likelihood ratings in the examples (and only in the examples), we re-normalise them so that they add up to 0 (just by adding the same constant to each component).

[8] It is worth noting that, although the approach developed by Slutzki and Volij is quite appealing normatively, there may be settings where one might desire to define rankings directly on reducible tournaments. In particular, consider the following situation (we thank P. Chebotarev for suggesting this example). There are $n + 1$ players, with player i having beaten player j whenever $i < j < n + 1$. Further, player $n + 1$ has beaten player n. According to the approach in Slutzki and Volij (2005), players 1 and $n+1$ are incomparable (both of them have a perfect score). However, it might be argued that player 1, who has beaten $n - 1$ other players, should be regarded as stronger than player $n + 1$, who has only beaten the weakest player.

[9] Note that, although in our setting the number of matches $M_{ij}(A)$ may not be an integer number, we always have in mind situations in which as many (possibly non-integer) points are divided between two players as matches they play. This extra generality could be useful in settings where partial comparisons are possible or, as we mentioned in the Introduction, different matches have different weights.

[10] Despite being very special, round-robin ranking problems are still more general than binary tournaments, since they allow for intensities (A_{ij} needs not be 0 or 1) and ties (A_{ij} may equal A_{ji}).

[11] Sometimes we use the notation $r(A)$ to indicate that the rating vector is derived from matrix A.

Neustadtl: Let $\hat{A}$ be defined, for each pair $i, j \in N$, by $\hat{A}_{ij} = A_{ij}/m_i$. Then, the Neustadtl rating vector is given by $r^{\mathrm{n}} = \hat{A}s$.

Fair bets: Let $L_A = \mathrm{diag}(A^\top e)$. So, for every $i \in N$, $(L_A)_i$ represents how much other players have scored against player i, *i.e.*, i's total number of "losses". Consider the system of linear equations given by $L_A^{-1}Ax = x$ or, equivalently, $\sum_{j\in N} A_{ij}x_j = \sum_{j\in N} A_{ji}x_i$ for all $i \in N$. The rating vector r^{fb} is defined to be the unique positive solution of the above system such that $(r^{\mathrm{fb}})^\top e = 1$.

Least squares: Let D be defined, for each pair $i, j \in N$ by $D_{ij} = \frac{A_{ij}-Aji}{M_{ij}}$. The ratings are then obtained via *least squared errors estimation*:

$$\min_{x\in\mathbb{R}^n} Q(x) = \min_{x\in\mathbb{R}^n} \sum_{i,j\in N} M_{ij}(D_{ij} - (x_i - x_j))^2.$$

The rating vector r^{ls} is defined to be the unique minimiser of the above problem such that $x^\top e = 0$.

Buchholz: The Buchholz rating vector is given by $r^{\mathrm{b}} = \bar{M}s + s$.

Recursive performance: Define $c \in \mathbb{R}^n$ by $c_i = F_L^{-1}(s_i)$ for all $i \in N$, where F_L is the (standard) logistic distribution $F_L(x) = 1/(1 + \exp(-x))$. Next, define $\hat{c} = c - \frac{m^\top c}{m^\top e}e$. Then the recursive performance rating vector, r^{rp}, is the unique solution of the system of linear equations given by $x^\top e = 0$ and $\bar{M}x + \hat{c} = x$.

Recursive Buchholz: The recursive Buchholz rating vector, r^{rb}, is the unique solution of the system of linear equations given by $x^\top e = 0$ and $\bar{M}x + \hat{s} = x$, where $\hat{s} = s - \frac{e}{2}$.

Generalised row sum: This is a parametric family of ranking methods. First, define $\hat{m} = \max_{i,j\in N} M_{ij}$, $A^* = A - A^\top$ and $C = \mathrm{diag}(m) - M$. Define the vector of *aggregate net scores* $s^* \in \mathbb{R}^n$ so that, for each $i \in N$, $s^*_i = \sum_{j\in N} A^*_{ij}$. Then, given $\varepsilon > 0$, the generalised row sum ratings are defined as the unique solution of the linear system of equations $(I + \varepsilon C)x = (1 + \hat{m}n\varepsilon)s^*$. To facilitate the comparison with other ranking methods we present the above ratings divided by $m(n-1)$ and, further, we take $\varepsilon = \frac{1}{\hat{m}(n-2)}$ and denote the corresponding ratings by r^{grs}.[12]

As we explain in Section 3, the (rather naive) Neustadtl and Buchholz methods find their origins in practice: they are methods commonly used, for instance, in chess tournaments. Their role in this paper is twofold. On one hand, we use them to motivate theoretically more sophisticated ranking methods: fair bets and recursive performance/Buchholz. On the other hand, the axiomatic analysis shows that this sophistication pays off.

The ranking methods defined above are illustrated in the following example.

[12] We postpone a discussion on the role of ε and this particular choice for this parameter to Section 3. Yet, it is worth noting that $\varepsilon = \frac{1}{\hat{m}(n-2)}$ is not well defined for the (trivial) class of two-player ranking problems. Thus, $n > 2$ is implicitly assumed when the generalised row sum method is discussed.

Example 2.1. *Consider the following three-player ranking problem. Player 1 has beaten player 2 four matches to one and player 3 five matches to none. On the other hand, player 2 has beaten player 3 quite consistently over a larger set of matches, twenty five to one. In our setting, this ranking problem would be represented through matrix A below:*

A			scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	r-Buch	GRS	s^*
0	4	5	0.9	2.051	0.352	0.801	0.533	1.335	1.764	0.267	0.39	8
1	0	25	0.839	0.608	0.055	0.192	0.172	1.011	0.491	0.086	0.407	21
0	1	0	0.032	−2.659	0.027	0.006	−0.705	0.881	−2.255	−0.353	−0.798	−29

We see that all the ranking methods defined above except for generalised row sum (and s^*) agree in the chosen ranking, with player 1 being ranked on top.

The above example can already be used to make some observations about the different ranking methods. First, we note that the least squares ratings are twice the recursive Buchholz ratings (the discrepancy comes from the rounding). In Proposition 3.1 below we show that this is indeed a general result: least squares and recursive Buchholz are the same ranking method. Second, we observe that, apparently, the generalised row sum method puts relatively more weight on the fact that player 2 has an excellent aggregate net score s^*_2 than on the average quality of his opponents. In this parametric family this weight is controlled by the choice of ε, which we discuss a bit more in Section 3.

We finish this section with three basic properties for a ranking method φ that all methods presented above satisfy.

Anonymity (ANO): Let $i, j \in N$ and let A' be the ranking problem obtained from A by permuting columns i and j and rows i and j. Then, the rankings $\varphi(A)$ and $\varphi(A')$ are the same but with players i and j interchanged. This standard property just requires that the rankings should be independent on the player's "names".

Homogeneity (HOM): For all $k > 0$, $\varphi(kA) = \varphi(A)$. Note that homogeneity is an ordinal property. It relates to the ordering of the players and not necessarily to the underlying rating vector. The rankings should be invariant to rescaling; in particular, if we get a ranking problem by adding a number of identical subproblems, the rankings for the big problem should coincide with those for the subproblems.

Symmetry (SYM): φ is flat on any symmetric ranking problem ($A = A^\top$). So if everyone has a 50% score against all opponents, not necessarily with the same number of matches, all players end up equally ranked.

3 Background on ranking methods

In this section we elaborate on the origins and motivations of the different ranking methods. In a nutshell, maximum likelihood and fair bets are probably the most outstanding representatives of the statistical and social choice approaches, respectively. We show below that recursive Buchholz is just a reinterpretation of least squares, which is another popular ranking method in statistics. Recursive performance is a relatively new ranking method which borrows from both maximum likelihood and least squares approaches and that has not been studied axiomatically. The generalised row sum approach also has statistical interpretations in terms of ridge estimation. Neustadtl and Buchholz are merely instrumental.[13]

Maximum likelihood is a classic ranking method whose origins can be traced as far back as Zermelo (1929). It has been studied in several fields, but it is specially popular in statistics (see, for instance, Bradley and Terry (1952), Moon and Pullman (1970) and David (1988)). The starting point is to view determining players' ratings as an estimation question in which the result matrix A is used as statistical evidence. Maximum likelihood looks for the ratings vector that maximises the probability of the matrix A being realised when all the matches given in matrix M take place.

Formally, this ranking method assumes that each player $i \in N$ has a rating r_i and that, given two players $i, j \in N$, $i \neq j$, the probability that player i beats player j is given by $F(r_i, r_j)$. Although there are several possible choices for the function F, here we follow a classic approach used in Zermelo (1929) and rediscovered by Bradley and Terry (1952). Under this approach, F is based on the (standard) logistic distribution $F_L(x) = 1/(1 + \exp(-x))$, so that $F(r_i, r_j) = F_L(r_i - r_j) = \exp(r_i)/(\exp(r_i) + \exp(r_j))$. Given this assumption on the beating probabilities, matrix A gives rise to a likelihood function (as function of r). The first order conditions for maximising this likelihood function can be rewritten as Eq. (2.1).[14]

Horst (1932) and, more recently, Mosteller (1951) also viewed ranking the players as an estimation question with A as statistical evidence. Their approach was extended beyond round-robin ranking problems by Gulliksen (1956) and revisited in Kaiser and Serlin (1978). Rather than explicitly modeling probabilities, they identified D_{ij}, the "realised difference" between players i and j in matrix A, as an estimate for $r_i - r_j$, the rating difference between them. The exact definition of D_{ij} is left as a degree of freedom in the model, but one common choice is $D_{ij} = \frac{A_{ij} - A_{ji}}{M_{ij}}$. Ideally, one would like to choose the vector $r \in \mathbb{R}^n$ such that $D_{ij} - (r_i - r_j) = 0$ for all pairs (i, j). This results in a system of $n^2 - n$ equations with n variables. To get around this issue, this approach proposes to estimate the r_i ratings via *least squared errors minimization*:

$$\min_{r \in \mathbb{R}^n} Q(r) = \min_{r \in \mathbb{R}^n} \sum_{i,j \in N} M_{ij}(D_{ij} - (r_i - r_j))^2.$$

From the minimization problem above, it may seem that we actually have a weighted least squares problem. However, this problem can be equivalently stated without the M_{ij} weights by taking the sum over all matches, instead of taking it over all pairs of players.

The Neustadtl ranking method (Neustadtl 1882) is widely used as a tie-breaker in round-robin ranking problems. It computes a weighted average of the individual scores of each player i, where the weight of his score against player j is proportional to the score of

[13]The reader interested in a deeper discussion of these and other ranking methods for this and other settings may refer to David (1988), Laslier (1997), Chebotarev and Shamis (1998), Chebotarev and Shamis (1999) and Brozos-Vázquez et al. (2008).

[14]Refer, for instance, to Ford (1957) or David (1988).

player j.[15] Thus, the idea behind Neustadtl is to reward a win against a player with a high score more than a win against a player with a lower score.

The scientific literature offers several ranking methods that build upon the idea above of rewarding wins without punishing losses. Two early contributions in this field are Wei (1952) and Kendall (1955). The most widely studied method within this stream of literature is the fair bets ranking method (Daniels 1969). Fair bets is a ranking method that was originally defined for round-robin ranking problems and that has been studied in social choice and voting theory under different names and interpretations: from the classic papers by Daniels (1969) and Moon and Pullman (1970), to more recent references such as Slutzki and Volij (2005) and Slutzki and Volij (2006).[16] In Laslier (1992) this ranking method is called the "ping-pong winners" because of the following interpretation in round-robin ranking problems. Suppose several players are waiting to play table tennis. The first two players i and j are randomly chosen and play. Player i wins with probability A_{ij}/M_{ij} and player j with probability A_{ji}/M_{ij}. The winner stays, a new opponent is randomly chosen, and the likelihood of each of them being the winner is derived again from matrix A. If we rank the players according to the amount of time they would play under the above rules, we would get the fair bets rankings.

Finally, with respect to the Neustadtl ranking, fair bets adds depth to the idea of rewarding results against good players. It is not only important to have beaten players who have high scores, but also that they have achieved these high scores beating players with high scores. In the table tennis example, given two players with the same average score, it is better to have beaten the one who has beaten stronger players. This reasoning can be given further levels of depth and the system of equations defining the fair bets ranking method captures them.

Whereas the Neustadtl ranking method rewards victories against strong players, the Buchholz[17] method takes one step back and takes the average strength of *all* your opponents into consideration ($\bar{M}s$), in addition to your own score (s).[18]

The recursive performance ranking method, defined in Brozos-Vázquez et al. (2008), combines ideas present in maximum likelihood and Buchholz. Instead of finding the ratings for which the likelihood of the observed results is maximised, as maximum likelihood does, recursive performance finds the rating that explains the "average" result of each player, giving it a Buchholz flavour. Given a ranking problem (N, A), a rating vector $r \in \mathbb{R}^n$, and a player i, the average opponent of i in the ranking problem is $(\bar{M}r)_i$, *i.e.*, the average rating of the opponents of i (weighted by the number of matches played against each of them).

[15] This (tie-breaking) rule, which originally was only defined for round-robin ranking problems, is commonly referred to as Sonneborn-Berger. In fact, Sonneborn and Berger criticised Neustadtl's method by arguing that the players's scores should be added to the Neustadtl ranking vector.

[16] Similar ideas have been also used in slightly different settings in papers such as Borm et al. (2002), Herings et al. (2005) and Slikker et al. (2010).

[17] Although Bruno Buchholz is widely recognised to have developed the method named after him (see World Chess Federation (2012)) in 1932, we were unfortunately not able to find a document on this by the inventor himself.

[18] Actually, Buchholz is commonly used as a tie-breaker in non round-robin chess tournaments and is computed as the average of the scores of the opponents of each player, *i.e.*, $\bar{M}s$. For players with an equal score, this definition results in the same relative ranking as the definition we present here.

Recursive performance looks for a rating such that for each player $i \in N$, $F(r_i, (\bar{M}r)_i) = \sum_{j\in N} A_{ij}/m_i = s_i$, where F is based on the logistic distribution function F_L. Hence, for each player i, r_i^{rp} takes into account the average strength of i's opponents $\bar{M}r^{\text{rp}}$ and his own score in the ranking problem ($\hat{c}_i$ is increasing in s_i).

Recursive Buchholz is a ranking method that combines the ideas of Buchholz and recursive performance by adding to the Buchholz ranking method the same kind of depth that the fair bets adds to Neustadtl.[19] Not only the average score of your opponents ($\bar{M}s$) should be important, but also whether your opponents have achieved this average score against weak or strong opponents. All else equal, having faced opponents with a high score who have themselves played against strong opponents should be better than having faced opponents with a high score who have played against weak opponents. Again, further depth can be given to this argument and recursive Buchholz captures this idea.

Interestingly, the next proposition shows that recursive Buchholz coincides with the least squared errors approach.[20]

Proposition 3.1. *The recursive Buchholz ranking and the least squares ranking coincide.*

Proof. Recall that the least squares ratings, r^{ls}, were defined as the unique solution of the minimisation problem $\min_{r\in\mathbb{R}^n} Q(r) = \min_{r\in\mathbb{R}^n} \sum_{i,j\in N} M_{ij}(D_{ij} - (r_i - r_j))^2$ such that $\sum_{i\in N} r_i = 0$, where $D_{ij} = \frac{A_{ij}-A_{ji}}{M_{ij}}$.

We show that r^{rb} solves the least squares problem with distances $D'_{ij} = \frac{A_{ij}-A_{ji}}{2M_{ij}}$. It is readily verified from the first order conditions that this implies that $2r^{\text{rb}}$ solves original problem, so r^{rb} provides the same ranking on the players.

Let $i \in N$. From the first order conditions we obtain:

$$\begin{aligned}
0 &= \sum_{j\in N} M_{ij}(D'_{ij} - r_i + r_j) \\
&= \sum_{j\in N} M_{ij}D'_{ij} - \sum_{j\in N} M_{ij}r_i + \sum_{j\in N} M_{ij}r_j \\
&= \sum_{j\in N} M_{ij}\frac{A_{ij}-A_{ji}}{2M_{ij}} - m_i r_i + \sum_{j\in N} M_{ij}r_j \\
&= \sum_{j\in N} A_{ij} - \frac{1}{2}\sum_{j\in N}(A_{ji} + A_{ij}) - m_i r_i + \sum_{j\in N} M_{ij}r_j.
\end{aligned}$$

Division by m_i yields

$$\begin{aligned}
0 &= s_i - \frac{1}{2m_i}\sum_{j\in N} M_{ij} - r_i + \sum_{j\in N} \bar{M}_{ij}r_j \\
&= \hat{s}_i - r_i + \sum_{j\in N} \bar{M}_{ij}r_j.
\end{aligned}$$

Hence, r^{ls} solves $x = \bar{M}x + \hat{s}$ and therefore r^{ls} and r^{rb} deliver the same rankings. □

[19] Although recursive Buchholz was not defined in Brozos-Vázquez et al. (2008), it can be seen as a variation of recursive performance where F_L is taken to be the identity. Thus, the existence and uniqueness of r^{rb} follows from Theorem 2 in that paper.

[20] We thank P. Chebotarev for pointing this out.

In the light of the previous result, hereafter we use the well established name of least squares rankings to refer to the rankings associated with r^{ls} and r^{rb}. Yet, in all the proofs in the paper it is more convenient to work with the vector r^{rb}.[21]

Finally, we have the generalised row sum method. It was first considered in Chebotarev (1989) and its properties were thoroughly analysed in Chebotarev (1994). This is actually a parametric family of ranking methods that range from the aggregate net scores s^* when $\varepsilon = 0$ to least squares when $\varepsilon \to \infty$; in general, this parameter measures how much "weight" the method puts on the opponents of the players with respect to s^*. In particular, based on some reasonableness condition, the upper bound $\varepsilon = \frac{1}{\hat{m}(n-2)}$ is identified, and this is the value we use throughout this paper. In Chebotarev (1994) this parametric family is obtained as the set of ranking methods satisfying certain conditions on how the pairwise results or the players are to be aggregated. Further, some statistical interpretations of the methods in this family are also discussed.

4 Response to victories and losses

In this section we consider two types of properties for a ranking method φ. The first type deals with preserving a ranking when two ranking problems (N, A) and (N, A') are combined. The second type deals with the (a)symmetric role victories and losses play in a ranking method.

Flatness preservation (FP): If $\varphi(A)$ and $\varphi(A')$ are both flat, then so is $\varphi(A + A')$. This property just says that if all players are regarded as equal in two ranking problems, this should not change when we add up the ranking problems.

Order preservation (OP): Let $i, j \in N$. If both $\varphi(A)$ and $\varphi(A')$ rank i strictly above j and $\frac{m_i}{m_j} = \frac{m'_i}{m'_j}$, then $\varphi(A + A')$ ranks i strictly above j as well. If i is better than j in two ranking problems, this should not change when we add them up. The condition on m and m' imposes some balance between the number of matches played in ranking problems A and A'. In Example 4.3 below we show that OP without this condition is not even satisfied by the score method.

Inversion (INV): Let $i, j \in N$. Then $\varphi(A)$ ranks i weakly above j if and only if $\varphi(A^\top)$ ranks j weakly above i. If we reverse all the results in a ranking problem, then the ranking should be reversed as well (Chebotarev and Shamis 1998). The spirit of this natural property, which trivially implies SYM, is to require a symmetric treatment between victories and losses.[22]

[21] Given the similarity in the definitions of recursive Buchholz and recursive performance, one may wonder whether the latter can also be rewritten as a least squares method for some adequately chosen D_{ij}. For recursive Buchholz, one of the key elements in the proof of Proposition 3.1 is that the D_{ij} entries can be seen as a disaggregation of the scores of the players, which are then recovered through the sums $\sum_{j \in N} M_{ij} D_{ij}$. In the recursive performance, the role of the scores vector is played by the vector $\hat{c}$. Yet, this vector aggregates the pairwise results of each player via the nonlinear function F_L, which makes it impossible to disaggregate in a natural way the components $\hat{c}$ to obtain the D_{ij} values.

[22] In Chebotarev (1994) this property is referred to as transposability.

Negative response to losses (NRL): Let $\lambda \in \mathbb{R}^n$, $\lambda > 0$ and define $\Lambda = \text{diag}((\lambda_i)_{i \in N})$. If $\varphi(A)$ is flat, then $\varphi(A\Lambda)$ ranks i weakly above j if and only if $\lambda_i \leq \lambda_j$. This property is introduced in Slutzki and Volij (2005) and is the key ingredient of the characterisation they obtain for the fair bets ranking method. In words of the authors: "Negative responsiveness to losses concerns situations in which all players are equally ranked and the problem is irreducible. If a new problem is obtained by multiplying each player's losses by some positive constant (which may be different for each player), then the players should be ranked in the new problem in a way that is inversely related to these constants".

It is rather straightforward that the score ranking method satisfies FP. The following proposition relates flatness of recursive performance to flatness of scores.

Proposition 4.1. *Let A be a ranking problem. Then $r^{\text{rp}}(A)$ is flat if and only if $s(A)$ is flat.*

Proof. "$\Rightarrow$": Assume that $r^{\text{rp}}(A)$ is flat, so there is $k \in \mathbb{R}$ such that $r^{\text{rp}}(A) = ke$. Recall that $r^{\text{rp}}(A)$ is a solution of $(I - \bar{M})r^{\text{rp}} = \hat{c}$, where $\hat{c}_i$ is strictly increasing in s_i. Then, $(I - \bar{M})r^{\text{rp}} = ke - k\bar{M}e = 0$. Hence, $\hat{c} = 0$ and therefore, $s(A)$ is flat.
"$\Leftarrow$": Assume that $s(A)$ is flat, so $s = \frac{1}{2}e$. Then, $c = 0$ and $\hat{c} = 0$. So, a particular solution of $(I - \bar{M})x = \hat{c}$ is 0 and the solution set is $\text{span}\{e\}$. Hence, $r^{\text{rp}}(A)$ is flat. □

In a similar way as in Proposition 4.1 one can show that both least squares and maximum likelihood are flat if and only if scores are flat. Yet, since the definition of generalised row sum builds upon the aggregate net scores s^* instead of the scores s, it is necessary to provide a separate result.

Proposition 4.2. *Let A be a ranking problem. Then $r^{\text{grs}}(A)$ is flat if and only if $s(A)$ is flat.*

Proof. "$\Rightarrow$": Assume that $s(A)$ is flat. Recall that, for each $i \in N$, $s_i = \sum_{j \in N} A_{ij}/m_i$ and $s_i^* = \sum_{j \in N}(A_{ij} - A_{ji})$. Since we always have that $\sum_{i \in N} s_i = n/2$, s is flat if and only if, for each $i \in N$, $s_i = 1/2$. In such a case, since $m_i = \sum_{j \in N} A_{ij} + \sum_{j \in N} A_{ji}$ we have $\sum_{j \in N} A_{ji} = \sum_{j \in N} A_{ij}$. Thus, if s is flat we have that s^* is the zero vector. Now, the zero vector is a solution of the system $(I + \varepsilon C)x = (1 + \hat{m}n\varepsilon)s^*$ and so $r^{\text{grs}}(A)$ is flat.

"$\Leftarrow$": Assume that $r^{\text{grs}}(A)$ is flat. By the *centering* property of generalised row sums, it is always true that $\sum_{i \in N} r_i^{\text{grs}} = 0$ (see Chebotarev (1994)). Thus, if $r^{\text{grs}}(A)$ is flat it must be the zero vector which, in turn, implies that also s^* is the zero vector. Now, for each $i \in N$, $\sum_{j \in N} A_{ji} = \sum_{j \in N} A_{ij}$ and $s_i = \sum_{j \in N} A_{ij}/m_i = \sum_{j \in N} A_{ij}/(2\sum_{j \in N} A_{ij}) = 1/2$. □

By FP of the score ranking method we obtain the following corollary.

Corollary 4.3. *Maximum likelihood, least squares, recursive performance and generalised row sum satisfy* FP.

Slutzki and Volij (2005) show that fair bets satisfies FP. The following example shows that Neustadtl and Buchholz do not.

Example 4.1. *Consider the ranking problems A and A' described below:*

A				scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0	0	2	2	0.667	0.347	0.222	0.333	0	1	0	0.222
0	0	2	2	0.667	0.347	0.222	0.333	0	1	0	0.222
1	1	0	0	0.333	−0.347	0.222	0.167	0	1	0	−0.222
1	1	0	0	0.333	−0.347	0.222	0.167	0	1	0	−0.222

A'				scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0	1	1	1	0.5	0	0.25	0.25	0	1	0	0
1	0	1	1	0.5	0	0.25	0.25	0	1	0	0
1	1	0	1	0.5	0	0.25	0.25	0	1	0	0
1	1	1	0	0.5	0	0.25	0.25	0	1	0	0

$A+A'$				scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0	1	3	3	0.583	0.203	0.257	0.3	0.1	1.028	0.202	0.133
1	0	3	3	0.583	0.203	0.257	0.3	0.1	1.028	0.202	0.133
2	2	0	1	0.417	−0.203	0.229	0.2	−0.1	0.972	−0.202	−0.133
2	2	1	0	0.417	−0.203	0.229	0.2	−0.1	0.972	−0.202	−0.133

Then $r^{\mathrm{n}}(A)$, $r^{\mathrm{n}}(A')$, $r^{\mathrm{b}}(A)$ and $r^{\mathrm{b}}(A')$ are all flat, but $r^{\mathrm{n}}(A+A')$ and $r^{\mathrm{b}}(A+A')$ are not.

Proposition 4.4. *The score ranking method satisfies* OP.

Proof. Let (N, A), (N, A') and $i, j \in N$ be such that $s_i > s_j$, $s'_i > s'_j$ and $\frac{m_i}{m_j} = \frac{m'_i}{m'_j}$. Note that $\frac{m_i+m'_i}{m_i} = \frac{m_j+m'_j}{m_j}$. Hence, $\frac{m_i}{m_i+m'_i} = \frac{m_j}{m_j+m'_j}$ and, clearly, $\frac{m'_i}{m_i+m'_i} = \frac{m'_j}{m_j+m'_j}$ as well. It is straightforward to check that the score of player i in the combined ranking problem $A + A'$ equals $\frac{m_i}{m_i+m'_i} s_i + \frac{m'_i}{m_i+m'_i} s'_i$. Then,

$$\frac{m_i}{m_i + m'_i} s_i + \frac{m'_i}{m_i + m'_i} s'_i = \frac{m_j}{m_j + m'_j} s_i + \frac{m'_j}{m_j + m'_j} s'_i > \frac{m_j}{m_j + m'_j} s_j + \frac{m'_j}{m_j + m'_j} s'_j,$$

which coincides with the score of player j in the combined ranking problem, so we have established OP. □

The following example shows that the other ranking methods do not satisfy OP.

Example 4.2. *Consider the ranking problems A and A' described below:*

A				scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0	5	15	5	0.625	0.383	0.278	0.344	0.183	1.075	0.378	0.21
5	0	5	15	0.625	0.315	0.253	0.322	0.15	1.05	0.307	0.19
5	5	0	7	0.425	−0.17	0.213	0.189	−0.083	0.975	−0.164	−0.11
5	5	3	0	0.325	−0.528	0.188	0.144	−0.25	0.9	−0.521	−0.29

A'				scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0	10	9	7	0.65	0.839	0.287	0.388	0.292	1.144	0.792	0.295
10	0	9	6	0.625	0.762	0.277	0.363	0.258	1.131	0.72	0.255
1	1	0	1	0.075	−1.956	0.048	0.026	−0.658	0.719	−1.8	−0.735
3	4	19	0	0.65	0.354	0.147	0.223	0.108	1.006	0.288	0.185

$A+A'$				scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0	15	24	12	0.637	0.452	0.265	0.347	0.203	1.088	0.43	0.261
15	0	14	21	0.625	0.475	0.291	0.358	0.217	1.109	0.46	0.264
6	6	0	8	0.25	−0.854	0.143	0.097	−0.383	0.828	−0.827	−0.486
8	9	22	0	0.488	−0.073	0.203	0.198	−0.037	0.975	−0.062	−0.039

Except for the score rankings, all other ranking methods rank player 1 strictly above player 2 in both A and A'. However, they rank player 2 on top of player 1 in ranking problem $A+A'$. Hence, none of these ranking methods satisfies OP. *Moreover, note that all players have played the same number of matches in A, A' and A'', whereas this was not required in the definition of* OP. *Hence, a weakening of* OP *in this direction would also be violated by all these ranking methods.*

The following example shows that a stronger version of OP without requiring the balance between m and m' is not even satisfied by the scores.

Example 4.3. *Consider the ranking problems A and A' described below:*

$$\begin{array}{cc} A & \text{scores} \\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 99 \\ 0.01 & 1 & 0 \end{pmatrix} & \begin{array}{c} 0.9901 \\ 0.9900 \\ 0.0100 \end{array} \end{array} \qquad \begin{array}{cc} A' & \text{scores} \\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0.01 \\ 99 & 1 & 0 \end{pmatrix} & \begin{array}{c} 0.0100 \\ 0.0099 \\ 0.9900 \end{array} \end{array} \qquad \begin{array}{cc} A+A' & \text{scores} \\ \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 99.01 \\ 99.01 & 2 & 0 \end{pmatrix} & \begin{array}{c} 0.0198 \\ 0.9802 \\ 0.5000 \end{array} \end{array}$$

Then, according to the score method, player 1 is ranked strictly above player 2 in both A and A' and yet, when adding them up, player 2 has a higher score.

The score ranking method trivially satisfies INV. For various other ranking methods, INV can be shown by explicitly transforming the rating vector.

Proposition 4.5. *The maximum likelihood ranking method satisfies* INV.

Proof. Recall that maximum likelihood orders the players according to r^{ml} where, for each $i \in N$, $r_i^{\text{ml}} = \log(\pi_i)$ and vector π is such that $\pi^\top e = 1$ and, for each $i \in N$,

$$\pi_i = \frac{m_i s_i}{\sum_{j \in N \setminus \{i\}} \frac{M_{ij}}{\pi_i + \pi_j}}.$$

Let $\bar{x}$ be defined, for each $i \in N$, by $\bar{x}_i = r_i^{\text{ml}} - \frac{1}{n} \sum_{j \in N} r_j^{\text{ml}}$. Then, $\sum_{i \in N} \bar{x}_i = 0$ and, for each $i \in N$, $\pi_i = \alpha \exp(\bar{x}_i)$ with $\alpha = (\prod_{j \in N} \pi_j)^{1/n}$. Hence, for each $i \in N$, we have

$$\exp(\bar{x}_i) = \frac{m_i s_i}{\sum_{j \in N \setminus \{i\}} M_{ij} \frac{1}{\exp(\bar{x}_i) + \exp(\bar{x}_j)}}. \tag{4.1}$$

Now consider the following system of equations in $y \in \mathbb{R}^n$: $\sum_{i \in N} y_i = 0$ and, for each $i \in N$,

$$\exp(y_i) = \frac{m_i (1 - s_i)}{\sum_{j \in N \setminus \{i\}} M_{ij} \frac{1}{\exp(y_i) + \exp(y_j)}}. \tag{4.2}$$

If we show that $y = -\bar{x}$ solves this system, then (because the transformation from π to $\bar{x}$ is monotonic) we are done since player i's score in $A^\top$ is $1 - s_i$ and M is the same in both ranking problems. Filling in $y = -\bar{x}$ in the right hand side of Eq. (4.2) yields

$$\begin{aligned}\frac{m_i(1-s_i)}{\sum_{j\in N\setminus\{i\}} M_{ij}\frac{1}{\exp(-\bar{x}_i)+\exp(-\bar{x}_j)}} &= \frac{m_i(1-s_i)}{\sum_{j\in N\setminus\{i\}} M_{ij}\frac{\exp(\bar{x}_i)\exp(\bar{x}_j)}{\exp(\bar{x}_i)+\exp(\bar{x}_j)}} \\ &= \frac{1}{\exp(\bar{x}_i)}\frac{m_i(1-s_i)}{\sum_{j\in N\setminus\{i\}} M_{ij}\frac{\exp(\bar{x}_j)}{\exp(\bar{x}_i)+\exp(\bar{x}_j)}} \\ &= \frac{1}{\exp(\bar{x}_i)}\frac{m_i(1-s_i)}{\sum_{j\in N\setminus\{i\}} M_{ij}(1-\frac{\exp(\bar{x}_i)}{\exp(\bar{x}_i)+\exp(\bar{x}_j)})} \\ &= \frac{1}{\exp(\bar{x}_i)}\frac{m_i(1-s_i)}{m_i - \exp(\bar{x}_i)\sum_{j\in N\setminus\{i\}}\frac{M_{ij}}{\exp(\bar{x}_i)+\exp(\bar{x}_j)}},\end{aligned}$$

which, by Eq. (4.1), reduces to

$$\frac{1}{\exp(\bar{x}_i)}\frac{m_i(1-s_i)}{m_i - \exp(\bar{x}_i)\frac{m_i s_i}{\exp(\bar{x}_i)}} = \frac{1}{\exp(\bar{x}_i)} = \exp(y_i).$$

So $y = -\bar{x}$ solves the system for $A^\top$ and therefore, maximum likelihood satisfies INV. □

Proposition 4.6. *Recursive performance, least squares, Buchholz and generalised row sum satisfy* INV.

Proof. To show that recursive performance satisfies INV, observe that if r^{rp} solves $\bar{M}x+\hat{c} = x$, then $-r^{\text{rp}}$ solves the corresponding equation for $A^\top$, because $\bar{M} = \bar{M}^\top$ and $\hat{c}(A^\top) = -\hat{c}(A)$ as a result of F^{-1} being symmetric around $\frac{1}{2}$. The argument for least squares and generalised row sum is analogous. For Buchholz, observe that $s(A^\top) = e - s(A)$, from which it readily follows that $r^{\text{b}}(A^\top) = 2e - (\bar{M}s(A) + s(A)) = 2e - r^{\text{b}}(A)$ and so the Buchholz ranking method satisfies INV as well. □

Not all ranking methods satisfy INV, as is shown in the following example.

Example 4.4. *Consider the following ranking problems:*

$$A = \begin{pmatrix} 0 & 0.5 & 0.2 & 1 \\ 0.5 & 0 & 0.3 & 0.8 \\ 0.8 & 0.7 & 0 & 0.9 \\ 0 & 0.2 & 0.1 & 0 \end{pmatrix} \quad \textit{and} \quad A^\top = \begin{pmatrix} 0 & 0.5 & 0.8 & 0 \\ 0.5 & 0 & 0.7 & 0.2 \\ 0.2 & 0.3 & 0 & 0.1 \\ 1 & 0.8 & 0.9 & 0 \end{pmatrix}$$

A: Neus	A: f-bets	$A^\top$: Neus	$A^\top$: f-bets
0.176	0.195	0.131	0.065
0.201	0.210	0.179	0.137
0.306	0.559	0.106	0.054
0.062	0.036	0.329	0.744

Since player 2 is ranked strictly above player 1 in both A and $A^\top$ and for both Neustadtl and fair bets, these ranking methods do not satisfy INV*. Note that A is a round-robin ranking problem. Hence, Neustadtl and fair bets do not satisfy* INV *even if we restrict to round-robin ranking problems.*

Our analysis of NRL builds upon Slutzki and Volij (2005), though some care is needed. On the one hand, they develop their characterisation of the fair bets ranking method for a larger class that allows for reducible ranking problems. On the other hand, they restrict to results matrices with integer entries.

A ranking problem A is called *balanced* if $Ae = A^\top e$, *i.e.*, if each player has the same number of victories and losses. It is strongly balanced if, moreover, there is a constant k such that $Ae = ke$, so the number of victories (and losses) is equal across all players. The next result is an adaptation of Lemmas 3 and 4 in Slutzki and Volij (2005).

Lemma 4.7. *Let φ be a ranking method satisfying* ANO*,* HOM*,* SYM *and* FP*. Then φ is flat on balanced ranking problems.*

Proof. First, suppose that A is strongly balanced with $Ae = ke$. Then by Birkhoff's theorem (Birkhoff 1946), matrix A can be written as k times a convex combination of permutation matrices. By ANO, φ is flat on permutation matrices. By HOM, φ is also flat on the ranking problems that result after the multiplication of the permutation matrices by positive numbers. Finally, by FP and HOM again, φ is flat also on matrix A.

If A is not strongly balanced, then A can be decomposed as the sum of a strongly balanced ranking problem, in which we have just seen that φ is flat, and a symmetric ranking problem (see the proof of Lemma 4 in Slutzki and Volij (2005)). By SYM, φ is flat on the symmetric ranking problem as well, and by FP it is then flat on the original ranking problem A. □

Most of the ranking methods we consider in this paper satisfy ANO, HOM, SYM and FP, and, therefore, all of them coincide (and are flat) for balanced ranking problems. The next result, which is the adaptation of the main result in Slutzki and Volij (2005) to our setting, illustrates the strength of the NRL property.

Proposition 4.8. *The fair bets ranking method is the unique ranking method satisfying* ANO*,* HOM*,* SYM*,* FP *and* NRL*.*

Proof. Fair bets has already been shown to satisfy ANO, HOM, SYM and FP. NRL follows from Slutzki and Volij (2005).

To show the converse, let φ be a ranking method satisfying ANO, HOM, SYM, FP and NRL. Given an irreducible ranking problem A and corresponding fair bets rating vector, r^{fb}, the ranking problem $A' = A\,\text{diag}((r_i^{\text{fb}})_{i\in N})$ is a balanced (and irreducible) ranking problem because, by definition, for all $i \in N$,

$$\sum_{j\in N} A_{ij} r_j^{\text{fb}} = \sum_{j\in N} A_{ji} r_i^{\text{fb}}.$$

Then, $A = A'(\text{diag}((r_i^{\text{fb}})_{i\in N}))^{-1}$. Since φ satisfies ANO, HOM, SYM and FP, by Lemma 4.7, $\varphi(A')$ is flat. Then, by NRL, $\varphi(A)$ ranks i weakly above j if and only if $1/r_i^{\text{fb}} \leq 1/r_j^{\text{fb}}$. Hence, φ is the fair bets ranking method. □

As a result of Proposition 4.8, scores, maximum likelihood, least squares, recursive performance and generalised row sum do not satisfy NRL because, being all of them different from fair bets, they satisfy all other properties in the characterisation. The following example shows that Neustadtl and Buchholz do not satisfy NRL either.

Example 4.5. *Let* $\lambda = (0.99, 2, 1, 1)$ *and* $\Lambda = \text{diag}((\lambda_i)_{i \in N})$. *Let* A *and* $A\Lambda$ *be as follows:*

$$A = \begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 2 & 0 \end{pmatrix}$$

Neus	Buch
0.25	1
0.25	1
0.25	1
0.25	1

and

$$A\Lambda = \begin{pmatrix} 0 & 4 & 1 & 1 \\ 1.98 & 0 & 1 & 1 \\ 0.99 & 2 & 0 & 2 \\ 0.99 & 2 & 2 & 0 \end{pmatrix}$$

Neus	Buch
0.245	1.024
0.192	0.911
0.264	1.046
0.264	1.046

Note that $r^{\text{n}}(A)$ *and* $r^{\text{b}}(A)$ *are both flat. Despite* $\lambda_1 \leq \lambda_3$, *we have* $r^{\text{n}}_3(A\Lambda) > r^{\text{n}}_1(A\Lambda)$ *and* $r^{\text{b}}_3(A\Lambda) > r^{\text{b}}_1(A\Lambda)$.

5 Score consistency

In this section we investigate to what extent a ranking method φ preserves some of the features of the score ranking method, making it appealing for round-robin ranking problems.

Score consistency (SCC): A ranking method φ satisfies this property if it coincides with the score ranking method on the class of round-robin ranking problems.

Homogeneous treatment of victories (HTV): Let $i, j \in N$. If $M_{ik} = M_{jk}$ for all $k \in N \backslash \{i, j\}$, then $\varphi(A)$ ranks i above j if and only if $s_i(A) \geq s_j(A)$. Roughly speaking, if i and j play the same number of matches against the other players, then they should be ranked according to their aggregate scores. Note that HTV trivially implies SCC.

Note that both SCC and HTV relate to the ranking of the players, not necessarily to the underlying rating vectors. It follows from the ranking problem A in Example 4.4 that Neustadtl and fair bets do not satisfy SCC.

The remaining ranking methods all satisfy HTV, and hence SCC.[23]

Proposition 5.1. *The maximum likelihood ranking method satisfies* HTV.

Proof. Let (N, A) and $i, j \in N$ be such that $M_{ik} = M_{jk}$ for all $k \in N \backslash \{i, j\}$. Rewriting the equations used to define r^{ml} we have

$$s_i = \frac{1}{m_i} \sum_{k \in N \backslash \{i\}} M_{ik} \frac{\pi_i}{\pi_i + \pi_k}.$$

Since $\frac{\pi_i}{\pi_i + \pi_k}$ is increasing in π_i, the right hand side of the equation is increasing in π_i. Then, because $M_{ik} = M_{jk}$ for all $k \neq i, j$ and therefore $m_i = m_j$, we have that $s_i \geq s_j$ if and only if $\pi_i \geq \pi_j$. Hence, maximum likelihood satisfies HTV. □

If $|N| = 2$ we have that $Ms + s = (s_1 + s_2, s_1 + s_2)^\top$, so Buchholz is flat in two-player tournaments and therefore satisfies neither HTV nor SCC.

Proposition 5.2. *If* $n > 2$, *then Buchholz satisfies* HTV.

[23] It is worth noting that the coincidence of the rankings proposed by maximum likelihood and the scores was already established in Zermelo (1929).

Proof. Let (N, A) and $i, j \in N$ be such that $M_{ik} = M_{jk}$ for all $k \in N\backslash\{i, j\}$. Given $i, j \in N$, since $M_{ik} = M_{jk}$ for all $k \neq i, j$, we have that $m_i = m_j$ and, hence, $\bar{M}_{ij} = \bar{M}_{ji}$. Then,

$$r_i^{\mathrm{b}} - r_j^{\mathrm{b}} = (\bar{M}s + s)_i - (\bar{M}s + s)_j = (1 - \bar{M}_{ij})(s_i - s_j).$$

Since A is irreducible and $n > 2$, it cannot be the case that $\bar{M}_{ij} = 1$. Then, $(1 - \bar{M}_{ij}) > 0$ and the Buchholz ranking method coincides with the scores. □

Proposition 5.3. *Least squares and recursive performance satisfy* HTV.

Proof. Recall that r^{rb} solves $(I - \bar{M})x = \hat{s}$. So, in particular

$$x_i - \bar{M}_{ij}x_j - \sum_{k\in N\backslash\{i,j\}} \bar{M}_{ik}x_k = \hat{s}_i \quad \text{and} \quad -\bar{M}_{ji}x_i + x_j - \sum_{k\in N\backslash\{i,j\}} \bar{M}_{jk}x_k = \hat{s}_j.$$

Subtracting the two equations and using that $\bar{M}_{ij} = \bar{M}_{ji}$ and $\bar{M}_{ik} = \bar{M}_{jk}$ for all other k yields

$$(1 + \bar{M}_{ij})(x_i - x_j) = \hat{s}_i - \hat{s}_j.$$

Therefore, $x_i - x_j$ and $\hat{s}_i - \hat{s}_j$ have the same sign. Hence, the least squares ranking method satisfies HTV.

The proof for recursive performance is analogous, but with $\hat{c}$ on the right hand side. Since $\hat{c}$ and $\hat{s}$ induce the same ranking, the same argument works. □

Proposition 5.4. *Generalised row sum satisfies* HTV.

Proof. Let (N, A) and $i, j \in N$ be such that $M_{ik} = M_{jk}$ for all $k \in N\backslash\{i, j\}$. Thus, $m_i = m_j$. Recall that r^{grs} solves $(I + \varepsilon C)x = (1 + \hat{m}n\varepsilon)s^*$. Thus, we have

$$\begin{aligned} x_i + \varepsilon m_i x_i - \sum_{k\in N, k\neq j} \varepsilon M_{ik}x_k - \varepsilon M_{ij}x_j &= s_i^*(1 + \varepsilon\hat{m}n) \quad \text{and} \\ x_j + \varepsilon m_j x_j - \sum_{k\in N, k\neq i} \varepsilon M_{jk}x_k - \varepsilon M_{ji}x_i &= s_j^*(1 + \varepsilon\hat{m}n). \end{aligned}$$

Subtracting the above two equations we get $x_i - x_j = (s_i^* - s_j^*)\frac{1+\varepsilon\hat{m}n}{1+\varepsilon(m_i+M_{ij})}$. Thus, we just have to show that $s_i^* - s_j^*$ and $s_i - s_j$ have the same sign. Let $m_0 = m_i = m_j$. Then, $\sum_{k\in N} A_{ki} = m_0 - \sum_{k\in N} A_{ik}$ and we get

$$s_i^* = \sum_{k\in N} A_{ik} - \sum_{k\in N} A_{ki} = 2\sum_{k\in N} A_{ik} - m_0 = 2m_0 s_i - m_0.$$

Similarly, $s_j^* = 2m_0 s_j - m_0$, so $s_i^* - s_j^* = 2m_0(s_i - s_j)$. □

6 Monotonicity

In this section we present several properties that deal with changes in the results matrix. If an existing result is changed or a new one is added, how should the rankings change? Because of the logical relationships existing between some of the properties we study in this section, we separate a bit from the rest of the analysis in this paper and present and discuss each property in turn, instead of defining all of them at the beginning of the section. In total we discuss five properties. We start by analyzing the implications in our general setting of the classic property of *independence of irrelevant matches*, IIM. We continue with two standard properties, PRB and NNRB that state that winning a match should always be beneficial to your ranking. Then we present a new property, which we call BPI, that illustrates the strength of PRB and why it might not always be desirable. Finally, we conclude by analyzing SCM, a natural monotonicity property introduced in Chebotarev and Shamis (1997) and thoroughly discussed in Chebotarev and Shamis (1999).

The IIM property deals with the responsiveness of the relative ranking of players i and j to matches not involving either of them.

Independence of irrelevant matches (IIM): We follow the definition introduced in Rubinstein (1980): take four different players $i, j, k, \ell \in N$. Suppose that A and A' are identical, except for the results between k and ℓ. Then the relative ranking between i and j in both $\varphi(A)$ and $\varphi(A')$ is the same.

Rubinstein (1980) uses ANO, IIM and PRB below to characterise the score ranking method on the class of binary ranking problems. Clearly, in our wider class of ranking problems the scores also satisfy IIM. We show below that none of the other ranking methods does.

Example 6.1. *Consider the ranking problems A and A' described below:*

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0.429	−0.203	0.224	0.2	−0.1	0.959	−0.201	−0.125
0.571	0.203	0.265	0.3	0.1	1.041	0.201	0.125
0.571	0.203	0.265	0.3	0.1	1.041	0.201	0.125
0.429	−0.203	0.224	0.2	−0.1	0.959	−0.201	−0.125

$$A' = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0.5	−0.016	0.237	0.233	−0.008	0.998	−0.015	−0.007
0.571	0.225	0.278	0.308	0.11	1.056	0.224	0.133
0.571	0.184	0.26	0.292	0.09	1.038	0.182	0.117
0.375	−0.393	0.205	0.167	−0.192	0.92	−0.39	−0.243

In ranking problem A, all ranking methods rank players 2 and 3 equally. In ranking problem A', except for the scores, all ranking methods rank player 2 on top of player 3, violating IIM.

Note that whereas IIM is a very natural property in round-robin ranking problems, it is questionable in our more general setting. Indeed, we argue in Section 7 that when players face different opponents, IIM is a property not to be desired.

Positive responsiveness to the beating relation (PRB): Let A be a ranking problem such that $\varphi(A)$ ranks i weakly above j. Let A' be a ranking problem identical to A, except that there is $k \in N \backslash \{i\}$ such that $M'_{ik} = M_{ik}$ and $A'_{ik} > A_{ik}$ (thus, $A'_{ki} < A_{ki}$). Then, $\varphi(A')$ ranks i strictly above j. Note that this should hold in particular for $k = j$.

The score ranking method obviously satisfies PRB. In fact, among the ranking methods under study, generalised row sum is the only other method satisfying PRB. This follows from the fact that generalised row sum satisfies a stronger property introduced in Chebotarev (1994) and called *monotonicity.*

Example 6.2. *Consider the ranking problems A and A' described below:*

A				scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0	1	20	20	0.5	0	0.25	0.25	0	1	0	0
1	0	20	0	0.5	0	0.25	0.25	0	1	0	0
20	20	0	0	0.5	0	0.25	0.25	0	1	0	0
20	0	0	0	0.5	0	0.25	0.25	0	1	0	0

A'				scores	max-lik	Neus	f-bets	least-sq	Buch	r-perf	GRS
0	1	20	39	0.732	0.916	0.14	0.331	0.238	1	0.416	0.369
1	0	20	0	0.5	0.916	0.256	0.331	0.238	1.011	1.398	0.04
20	20	0	0	0.5	0.916	0.308	0.331	0.238	1.116	1.17	0.102
1	0	0	0	0.025	−2.748	0.018	0.008	−0.713	0.757	−2.984	−0.51

According to all methods under consideration, players 1, 2 and 3 are equally ranked in A. In A', player 1 has a better result against player 4 than in A, but only the score ranking method and generalised row sum rank him above players 2 and 3. Hence, all other methods violate PRB.

A crucial aspect in Example 6.2 is that player 1 is what we call a *bridge player*, which in general is defined as follows. Given a ranking problem (N, A), a player $b \in N$ is a *bridge player* if there exist $N^1, N^2 \subseteq N$ with cardinalities $n^1 \geq 2$ and $n^2 \geq 2$ such that $N^1 \cup N^2 = N$, $N^1 \cap N^2 = \{b\}$ and $M_{ij} = 0$ for all $i \in N^1 \backslash \{b\}, j \in N^2 \backslash \{b\}$. Since no player of $N^1 \backslash \{b\}$ has played against any player in $N^2 \backslash \{b\}$, the connectedness of irreducible ranking problems with bridge players depends crucially on them, in the sense that the ranking problem obtained after removing a bridge player would not be connected. We denote by (N^1, A^1) and (N^2, A^2) the subproblems obtained from (N, A) by reducing A to the player sets N^1 and N^2, respectively. Further, note that the irreducibility of A implies that a bridge player has scored against at least one player in each of the subproblems, and that at least one player in each of the subproblems has scored against him.

One might argue that when a ranking problem consists of two subproblems connected by a bridge player, the relative rankings within each subproblem should not be influenced by the results in the other subproblem. This is formalised in the property BPI.

Bridge player independence (BPI): Let b be a bridge player with corresponding subproblems (N^1, A^1) and (N^2, A^2). Then for all $i, j \in N^1$, $\varphi(A)$ ranks i weakly above j if and only if $\varphi(A^1)$ ranks i weakly above j.

The main motivation to require BPI lies in the following example. Consider a ranking problem (N^1, A^1) with $A_{ij} = \frac{1}{2}$ for all $i \neq j$. Obviously, any symmetric ranking method ranks all players equally in A^1. Now suppose player i finds some (not very strong) friends outside N^1 who are willing to play against him and, subsequently, player i (who is now a bridge player between N^1 and N^2) has good results in A^2. BPI implies that player i, with respect to the players in N^1, should not benefit from having played the extra matches against the players in N^2; actually, these extra matches should have no bearing at all on the relative ranking of i within N^1.[24]

It immediately follows from Proposition 6.2 below that generalised row sum does not satisfy BPI. For the rest of the ranking methods, just consider again the ranking problems A and A' in Example 6.2. Player 1 is a bridge player in both A and A', with $N^1 = \{1, 4\}$ and $N^2 = \{1, 2, 3\}$. In ranking problem A, all players are tied according to all ranking methods. Yet, in A', players 1 and 3 are not tied anymore according to the scores, Neustadtl, recursive performance and Buchholz. Since the only difference between A and A' is in the subproblem (N^1, A^1), these rules do not satisfy BPI.

Proposition 6.1. *Fair bets, least squares and maximum likelihood satisfy* BPI.

Proof. Let b be a bridge player with respect to the subproblems (N^1, A^1) and (N^2, A^2). We start with the proof for fair bets.

Take $x^1 = r^{\text{fb}}(N^1, A^1)$ and $x^2 = r^{\text{fb}}(N^2, A^2)$ and define, for all $i \in N$,

$$y_i = \begin{cases} \frac{x_b^2}{x_b^1} x_i^1 & \text{if } i \in N^1, \\ x_i^2 & \text{if } i \in N^2. \end{cases}$$

Since the fair bets rating vector associated with an irreducible ranking problem is positive, the vector y is well defined. Then, for $i \in N^1 \backslash \{b\}$ we have that, for all $j \in N^2$, $A_{ij} = A_{ji} = 0$ and hence

$$\sum_{j \in N} A_{ij} y_j = \sum_{j \in N^1} A^1_{ij} \frac{x_b^2}{x_b^1} x_j^1 = \frac{x_b^2}{x_b^1} \sum_{j \in N^1} A^1_{ji} x_i^1 = \sum_{j \in N} A_{ji} y_i.$$

Similarly, for $i \in N^2 \backslash \{b\}$ we have

$$\sum_{j \in N} A_{ij} y_j = \sum_{j \in N^2} A^2_{ij} x_j^2 = \sum_{j \in N^2} A^2_{ji} x_i^2 = \sum_{j \in N} A_{ji} y_i.$$

Finally,

$$\sum_{j \in N} A_{bj} y_j = \sum_{j \in N^1} A^1_{bj} \frac{x_b^2}{x_b^1} x_j^1 + \sum_{j \in N^2} A^2_{bj} x_j^2 = \frac{x_b^2}{x_b^1} \sum_{j \in N^1} A^1_{jb} x_b^1 + \sum_{j \in N^2} A^2_{jb} x_b^2 = \sum_{j \in N} A_{jb} y_b.$$

Since the system given by the $\sum_{j \in N} A_{ji} y_j = \sum_{j \in N} A_{ij} y_i$ equations has a unique solution up to a positive scalar multiplication, $r^{\text{fb}}(N, A)$ and y induce the same rankings. From this, BPI follows.

The proof for maximum likelihood is analogous, but we use the vector π that solves the system of non-linear equations that are used to compute r^{ml}. Since r^{ml} is a strictly monotonic transformation of π, they induce the same ranking.

Finally, the proof for least squares goes along similar lines, but the vector y is defined, for all $i \in N$, by $y_i = \begin{cases} x_i^1 + x_b^2 & \text{if } i \in N^1, \\ x_b^1 + x_i^2 & \text{if } i \in N^2. \end{cases}$ □

[24]To what extent it is important to work with ranking methods that satisfy this property depends on how the matches matrix is constructed; for instance, on whether or not the players can choose their own opponents.

From Example 6.2 and our discussion on the merits of BPI, it is apparent that PRB and BPI are essentially incompatible. More precisely, the strictness on the rankings imposed by PRB is incompatible with BPI. The next result formalises this statement.

Proposition 6.2. *If a ranking method satisfies* SYM*, then it cannot satisfy both* PRB *and* BPI.

Proof. Let φ be a ranking satisfying SYM and consider the ranking problems A and A' in Example 6.2. Since A is a symmetric ranking problem, SYM implies that φ is flat on A. Now, by PRB player 1 should be ranked on top of any other player in ranking problem A'. Also, note that Player 1 is a bridge player in both A and A', with $N^1 = \{1,4\}$ and $N^2 = \{1,2,3\}$. Thus, by BPI, the relative ranking of player 1 with respect to players 2 and 3 should not change from A to A'. Thus, the requirements of PRB and BPI are incompatible. □

A natural solution to the above incompatibility is to relax the strict inequality in PRB, which gives NNRB below.

Nonnegative responsiveness to the beating relation (NNRB): Let A be a ranking problem such that $\varphi(A)$ ranks i weakly above j. Let A' be a ranking problem identical to A, except that there is $k \in N \backslash \{i\}$ such that $M'_{ik} = M_{ik}$, $A'_{ik} > A_{ik}$ and $A'_{ki} < A_{ki}$. Then, $\varphi(A')$ ranks i weakly above j.

Of course, PRB trivially implies NNRB. It follows from Example 6.2 that Neustadtl, recursive performance and Buchholz do not even satisfy this weaker responsiveness property, because all three methods rank player 1 lower than 2 and 3 in A', whereas in A they were ranked equal.

Next, we show that maximum likelihood, least squares and fair bets satisfy NNRB.

Proposition 6.3. *Maximum likelihood satisfies* NNRB.

Proof. For each $i \in N$, let $A_i = \sum_{j \in N} A_{ij}$. We defined maximum likelihood as the unique positive solution of the system of non-linear equations given by $\pi^\top e = 1$ and, for each $i \in N$,

$$\pi_i = \frac{m_i s_i}{\sum_{\ell \in N \backslash \{i\}} \frac{M_{i\ell}}{\pi_i + \pi_\ell}}, \quad \text{or, equivalently,} \quad A_i = \sum_{\ell \in N} M_{i\ell} \frac{\pi_i}{\pi_i + \pi_\ell}.$$

Let A and A' be two matrices as in the definition of NNRB:

- There are i and j in N such that $\pi_i \geq \pi_j$.
- $M = M'$ and A and A' are equal except that there is $k \in N \backslash \{i\}$ such that $A'_{ik} > A_{ik}$ and $A'_{ki} < A_{ki}$.

We want to show that $\pi'_i \geq \pi'_j$. First, note the following. Given t and ℓ in N,

$$\frac{\pi'_t}{\pi'_t + \pi'_\ell} \overset{(\geq)}{>} \frac{\pi_t}{\pi_t + \pi_\ell} \iff \frac{\pi_t + \pi_\ell}{\pi_t} \overset{(\geq)}{>} \frac{\pi'_t + \pi'_\ell}{\pi'_t} \iff 1 + \frac{\pi_\ell}{\pi_t} \overset{(\geq)}{>} 1 + \frac{\pi'_\ell}{\pi'_t}$$
$$\iff \frac{\pi_\ell}{\pi_t} \overset{(\geq)}{>} \frac{\pi'_\ell}{\pi'_t} \iff \frac{\pi'_t}{\pi_t} \overset{(\geq)}{>} \frac{\pi'_\ell}{\pi_\ell}. \tag{6.1}$$

Suppose now that $\pi'_j > \pi'_i$. We want to reach a contradiction. Let $V = \{t \in N : \frac{\pi'_t}{\pi_t} = \max_{\ell \in N} \frac{\pi'_\ell}{\pi_\ell}\}$. Note that, $\frac{\pi'_j}{\pi'_i} > 1 \geq \frac{\pi_j}{\pi_i}$ and, hence, $\frac{\pi'_j}{\pi_j} > \frac{\pi'_i}{\pi_i}$. Thus, $i \notin V$. By Eq (6.1), for each $t \in V$ and each $\ell \in N$, $\frac{\pi'_t}{\pi'_t+\pi'_\ell} \geq \frac{\pi_t}{\pi_t+\pi_\ell}$ and, if $\ell \notin V$, $\frac{\pi'_t}{\pi'_t+\pi'_\ell} > \frac{\pi_t}{\pi_t+\pi_\ell}$. Moreover, by irreducibility, there are $t \in V$ and $\ell \notin V$ such that $M_{t\ell} > 0$. Since $t \neq i$, $A_t \geq A'_t$, but

$$A'_t = \sum_{\ell \in N} M_{t\ell} \frac{\pi'_t}{\pi'_t + \pi'_\ell} > \sum_{\ell \in N} M_{t\ell} \frac{\pi_t}{\pi_t + \pi_\ell} = A_t,$$

and we have a contradiction. Therefore, $\pi'_i \geq \pi'_j$. □

The two results below establish that both least squares and fair bets satisfy NNRB. The result for fair bets extends the result in Levchenkov (1992) and Laslier (1997) for binary tournaments. In our proofs we build upon the arguments in the later reference and, for the sake of exposition, we relegate them to the Appendix.

Proposition 6.4. *Let ranking problem A be such that $r_i^{\mathrm{rb}}(A) \geq r_j^{\mathrm{rb}}(A)$. Let A' be a ranking problem identical to A, except that there is $k \in N \backslash \{i\}$ such that $M'_{ik} = M_{ik}$ and $A'_{ik} > A_{ik}$. Then, $r_i^{\mathrm{rb}}(A') \geq r_j^{\mathrm{rb}}(A')$ and, if $k = j$, then $r_i^{\mathrm{rb}}(A') > r_j^{\mathrm{rb}}(A')$. In particular, least squares satisfies* NNRB.

Proof. See Appendix. □

Proposition 6.5. *Let ranking problem A be such that $r_i^{\mathrm{fb}}(A) \geq r_j^{\mathrm{fb}}(A)$. Let A' be a ranking problem identical to A, except that there is $k \in N \backslash \{i\}$ such that $M'_{ik} = M_{ik}$ and $A'_{ik} > A_{ik}$. Then, $r_i^{\mathrm{fb}}(A') \geq r_j^{\mathrm{fb}}(A')$ and, if $k = j$, $r_i^{\mathrm{fb}}(A') > r_j^{\mathrm{fb}}(A')$. In particular, fair bets satisfies* NNRB.

Proof. See Appendix. □

We present now one last monotonicity property, introduced in Chebotarev and Shamis (1997) and deeply analized in Chebotarev and Shamis (1999) under the name of self-consistent monotonicity.[25] Informally, it says that if the results of two players i and j can be compared in such a way that it is clear that player i has performed better than player j, then player i should be ranked above player j. More precisely, given a ranking method φ, if we can decompose a ranking problem A in a series of subproblems such that, based on $\varphi(A)$, player i has obtained unquestionably better results than player j in each of them, then φ has to rank player i above player j in ranking problem A. This property was originally defined in a domain like the one described in the Introduction, where all the R^ℓ matrices are available; this is why need to artificially "decompose" ranking problem A into subproblems.

Self-consistent monotonicity (SCM): Let A be a ranking problem and φ a ranking method. Suppose there are two players i and j for which there is $p \in \mathbb{N}$ such that A can be decomposed as $A = \sum_{t=0}^{p} A^t$, where the (possibly reducible) A^t matrices satisfy:

[25] Two very similar properties are discussed in Chebotarev and Shamis (1998) and Conner and Grant (2009).

(i) $\sum_{k\in N} A^0_{ki} = 0$ and $\sum_{k\in N} A^0_{jk} = 0$, *i.e.*, player i has a perfect score in A^0 (provided that he has played some match in A^0) and player j has a zero score.

(ii) For each $t \in \{1,\dots,p\}$, the matches matrix M^t is a $0-1$ matrix. Let $O^t_i = \{k \in N : M_{ik} = 1\}$ denote the set of players against whom player i has played in A^t; similarly, define $O^t_j = \{k \in N : M_{jk} = 1\}$. Then, there is a one-to-one mapping $g^t : O^t_i \to O^t_j$ such that, if $g^t(k) = \ell$, then $A^t_{ik} \geq A^t_{j\ell}$ and $\varphi(A)$ ranks k weakly above ℓ.

Then, $\varphi(A)$ ranks i weakly above j. Further, if $\sum_{k\in N} A^0_{ik}$ is different from zero, or $\sum_{k\in N} A^0_{kj}$ is different from zero, or at least one weak inequality/comparison in (ii) is strict, then $\varphi(A)$ ranks i strictly above j.

Proposition 6.6. *Maximum likelihood and generalised row sum satisfy* SCM.

Proof. This result is an immediate consequence of Theorem 12 in Chebotarev and Shamis (1999); regarding generalised row sum, the result holds as long as $\varepsilon > 0$. ☐

Example 6.3. *Consider the ranking problem A described below:*

$$A = \begin{pmatrix} 0 & 0 & 0.9 & 1 \\ 0 & 0 & 0.9 & 0 \\ 0.1 & 0.1 & 0 & 0.9 \\ 0 & 0 & 0.1 & 0 \end{pmatrix}$$

Neus	least-sq	r-perf	Buch
0.19	0.2	1.361	1.158
0.33	0.3	1.665	1.267
0.077	−0.1	−0.462	1
0.018	−0.4	−2.565	0.708

In this ranking problem, players 1 and 2 have the same results, with the difference that player 1 has an extra match against player 4 in which he has achieved a perfect score. Thus, any ranking method satisfying SCM *should rank player 1 on top of player 2. Therefore, Neustadtl, least squares, recursive performance and Buchholz do not satisfy* SCM.

We consider that Example 6.3 illustrates a fairly undesirable behaviour of these four methods. First, note that the situation is completely different from what we had with PRB. Here, even if we had defined a weaker version of SCM without the strictness requirement included in the last part of the definition, the same example would work to show that the four methods still violate this property. It is interesting to interpret what is happening with least squares. Because of his extra victory compared to player 2, player 1 has a higher score. Yet, this comes at the price of lowering his "average opponent". When these two things come into play through the equation $\bar{M}x + \hat{s} = x$, it turns out that the latter effect dominates. From the point of view of the least squares minimisation, as Chebotarev and Shamis (1999) put it, this method punishes player 1 for the win over player 4 because, given that player 1 beats player 3 by a large margin and that player 3 beats player 4 also by a large margin, player 1 should beat player 4 with a greater intensity, which is not possible.

Example 6.4. *Consider the following ranking problems:*

$$A = \begin{pmatrix} 0 & 0 & 0.9 & 0 \\ 0 & 0 & 0 & 0.9 \\ 0.1 & 0 & 0 & 0.9 \\ 0 & 0.1 & 0.1 & 0 \end{pmatrix} \quad \text{scores: } \begin{matrix} 0.9 \\ 0.9 \\ 0.5 \\ 0.1 \end{matrix} \qquad \textit{and} \qquad A' = \begin{pmatrix} 0 & 0.6 & 0.5 & 1 & 0 \\ 0.4 & 0 & 0.5 & 0 & 1 \\ 0.5 & 0.5 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 \end{pmatrix} \quad \text{f-bets: } \begin{matrix} 0.378 \\ 0.289 \\ 0.2 \\ 0.067 \\ 0.067 \end{matrix}$$

In ranking problem A, players 1 *and* 2 *have beaten, respectively, players* 3 *and* 4 *with an identical result. They have not played more matches and, hence, since player* 3 *has a higher score than player* 4*,* SCM *requires that player* 1 *has a strictly greater rating than player* 2*;* i.e.*, according to* SCM *player* 1 *should be rewarded for having beaten a stronger opponent than player* 2*. Therefore, the score ranking method does not satisfy* SCM*.*

In ranking problem A', players 4 *and* 5 *have lost, respectively, against players* 1 *and* 2 *with an identical result. Further, they have achieved the same result against player* 3*, playing no more matches. In this case, since according to the fair bets ranking player* 1 *is strictly better than player* 2*,* SCM *requires that player* 4 *has a strictly greater rating than player* 5*;* i.e.*, according to* SCM *player* 4 *should be rewarded for having lost against a stronger opponent than player* 5*. Therefore, the fair bets ranking method does not satisfy* SCM

7 Discussion

Table 1 summarises the results of the ranking methods we have studied with respect to the different properties (with the addition of one extra property that we discuss below). From our point of view, maximum likelihood is the ranking method that looks most appealing, with generalised row sum also exhibiting a very good behaviour. Essentially, the main difference between maximum likelihood and generalised row sum hinges on PRB and BPI, two properties that we have shown to be essentially incompatible. Thus, to decide between these two ranking methods one may want to think of the desirability of PRB and BPI in the setting under study.

The main weakness of least squares is that it does not satisfy SCM and Example 6.3 shows a ranking problem where violating this property seems completely inappropriate. Concerning fair bets, although it also violates SCM, the behavior in the corresponding example does not seem as inappropriate as the example for least squares. Yet, we consider that its major weakness is that it violates INV, which imposes the natural requirement that if we reverse all the results in the ranking problem, then the corresponding ranking should be obtained by reversing the original ranking as well.

It is worth noting that one potential weakness of maximum likelihood is that it requires to solve a system of non-linear equations. Thus, it may be hard to compute in settings where there is a high number of players to be ranked. The difficulties to compute r^{ml} were already studied in Dykstra (1956), but it would definitely be interesting to reassess these difficulties in the light of all advances that computer science has experienced since then. All other methods under consideration are easy to calculate, since they are based on rating vectors that can be computed by solving some linear system of equations (indicated by the L-SOL property in Table 1).

The score ranking method satisfies most of the properties we have studied. However, although this ranking method is very natural when looking at round-robin ranking problems, in our more general setting it has the important drawback that it just looks at the aggregate score of each player, ignoring the opponents he has faced to obtain this score. All the other

	Scores	Maximum likelihood	Neustadtl	Fair bets	Least squares	Buchholz	Recursive performance	Generalised row sum ($\varepsilon = \frac{1}{\hat{m}(n-2)}$)
ANO	✓	✓	✓	✓	✓	✓	✓	✓
HOM	✓	✓	✓	✓	✓	✓	✓	✓
SYM	✓	✓	✓	✓	✓	✓	✓	✓
FP	✓	✓	X	✓	✓	X	✓	✓
OP	✓	X	X	X	X	X	X	X
INV	✓	✓	X	X	✓	✓	✓	✓
NRL	X	X	X	✓	X	X	X	X
SCC	✓	✓	X	X	✓	✓*	✓	✓
HTV	✓	✓	X	X	✓	✓*	✓	✓
IIM	✓	X	X	X	X	X	X	X
PRB	✓	X	X	X	X	X	X	✓
NNRB	✓	✓	X	✓	✓	X	X	✓
BPI	X	✓	X	✓	✓	X	X	X
SCM	X	✓	X	X	X	X	X	✓
L-SOL	✓	X	✓	✓	✓	✓	✓	✓

*Requires $|N| > 2$.

Table 1: Ranking methods and properties.

ranking methods we have considered use this information. That is, in one way or another, they are responsive to the strength of the opponents of each player. This is captured by the fact that the score ranking method is the only one satisfying IIM also outside the subdomain of round-robin ranking problems. So when players have different opponents (or face opponents with different intensities), IIM is a property one would rather not have.

References

Arrow, K. J. (1963). *Social choice and individual values.* John Wiley & Sons, Inc., New York, London, Sydney.

Berman, A. and R. Plemmons (1994). *Nonnegative matrices in the mathematical sciences* (2nd ed.). SIAM.

Birkhoff, G. (1946). Tres observaciones sobre el álgebra lineal. *Univ. Nac. Tucumán. Revista A*, **5**, 147–150.

Borm, P., G. Fiestras-Janeiro, H. Hamers, E. Sánchez-Rodríguez, and M. Voorneveld (2002). On the convexity of games corresponding with sequencing situations with due dates. *European Journal of Operational Research*, **136**, 616–634.

Bouyssou, D. (2004). Monotonicity of ranking by choosing: A progress report. *Social Choice and Welfare*, **23**, 249–273.

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

Brozos-Vázquez, M., M. A. Campo-Cabana, J. C. Díaz-Ramos, and J. González-Díaz (2008). Ranking participants in tournaments by means of rating functions. *Journal of Mathematical Economics*, **44**, 1246–1256.

Chebotarev, P. (1989). Generalization of the row sum method for incomplete paired comparisons. *Automation and Remote Control*, **50**, 1103–1113.

Chebotarev, P. (1994). Aggregation of preferences by the generalized row sum method. *Mathematical Social Sciences*, **27**, 293–320.

Chebotarev, P. and E. Shamis (1997). Constructing an objective function for aggregating incomplete preferences. In: A. Tangian and J. Gruber (Eds.), *Econometric Decision Models: Constructing Scalar-Valued Objective Functions*, Volume 453 of *Lecture Notes in Economics and Mathematical Systems*, pp. 100–124. Springer-Verlag.

Chebotarev, P. and E. Shamis (1998). Characterizations of scoring methods for preference aggregation. *Annals of Operations Research*, **80**, 299–332.

Chebotarev, P. and E. Shamis (1999). Preference fusion when the number of alternatives exceeds two: indirect scoring procedures. *Journal of the Franklin Institute*, **336**, 205–226.

Conner, G. R. and C. P. Grant (2009). Neighborhood monotonicity, the extended Zermelo model, and symmetric knockout tournaments. *Discrete Mathematics*, **309**, 3998–4010.

Copeland, A. H. (1951). A reasonable social welfare function. In: Seminar on Applications of Mathematics to the Social sciences. Univ. of Michigan.

Daniels, H. (1969). Round-robin tournament scores. *Biometrika*, **56**, 295–299.

David, H. A. (1988). *The method of paired comparisons.* London: Chapman and Hall.

Dutta, B. and J. Laslier (1999). Comparison functions and choice correspondences. *Social Choice and Welfare*, **16**, 513–552.

Dykstra, O. (1956). A note on the rank analysis of incomplete block designs – Applications beyond the scope of existing tables. *Biometrics*, **3**, 301–306.

Ford, L. R. (1957). Solution of a ranking problem from binary comparisons. *American Mathematical Monthly*, **64**, 28–33.

Gulliksen, H. (1956). A least squares solution for paired comparisons with incomplete data. *Psychometrika*, **21**, 125–134.

Herings, P., G. van der Laan, and D. Talman (2005). The positional power nodes in digraphs. *Social Choice and Welfare*, **24**, 439–454.

Horst, P. (1932). A method for determining the absolute affective value of a series of stimulus situations. *Journal of Educational Psychology*, **23**, 418–440.

Kaiser, H. F. and R. C. Serlin (1978). Contributions to the method of paired comparisons. *Applied Psychological Measurement*, **2**, 421–430.

Kendall, M. (1955). Further contributions to the theory of paired comparisons. *Biometrics*, **11**, 43–62.

Langville, A. N. and C. D. Meyer (2012). *Who is#1? The science of rating and ranking.* Princeton University Press.

Laslier, J. (1992). Solutions de tournois. Technical report, Habilitation à diriger les recherches en Economie, Université de Cergy-Pontoise.

Laslier, J. (1997). *Tournament Solutions and Majority Voting.* Springer-Verlag.

Levchenkov, V. (1992). Social choice theory: a new insight. Technical report, Institute of Systems Analysis, Moscow.

Liebowitz, S. and J. Palmer (1984). Assesing the relative impacts of economic journals. *Journal of Economic Literature*, **22**, 77–88.

Moon, J. W. and N. Pullman (1970). On generalized tournament matrices. *SIAM Review*, **12**, 384–399.

Mosteller, F. (1951). Remarks on the method of paired comparisons. I. The least squares solution assuming equal standard deviations and equal correlations. *Psychometrika*, **16**, 3–11.

Neustadtl, H. (1882). Letter to The Chess Monhtly, London (Eds. L. Hoffer and J. Zukertort).

Palacios-Huerta, I. and O. Volij (2004). The measurement of intellectual influence. *Econometrica*, **72**, 963–977.

Rubinstein, A. (1980). Ranking the participants in a tournament. *SIAM Journal on Applied Mathematics*, **38**, 108–111.

Slikker, M., P. Borm, and R. van den Brink (2010). Internal slackening scoring methods. Beta working paper, WP-306. Technische Universiteit Eindhoven. Eindhoven, The Netherlands.

Slutzki, G. and O. Volij (2005). Ranking participants in generalized tournaments. *International Journal of Game Theory*, **33**, 255–270.

Slutzki, G. and O. Volij (2006). Scoring of web pages and tournaments- Axiomatizations. *Social Choice and Welfare*, **26**, 75–92.

Wei, T. (1952). *The algebraic foundations of ranking theory.* Ph. D. thesis, Cambridge University.

World Chess Federation (2012). Handbook. C: General rules and recommendations for tournaments. Annex to the FIDE tournament regulations regarding tiebreaks. `http://www.fide.com/component/handbook`.

Zermelo, E. (1929). Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, **29**, 436–460.

A Proofs of Propositions 6.4 and 6.5

We start with an auxiliary result that is crucial in the proof for both fair bets and least squares.

Lemma A.1. *Let $B \in \mathbb{R}^{n\times n}$ be such that*

(i) *B is invertible,*

(ii) *for all $i \neq j$, $B_{ij} \leq 0$, and*

(iii) *$\sum_{j=1}^{n} B_{ji} \geq 0$ for all $i \in \{1,\dots,n\}$.*

If $\gamma, \lambda \in \mathbb{R}^n$ are two vectors such that λ is nonnegative and $B\gamma = \lambda$, then γ is nonnegative.

Proof. It follows from the assumptions that B is an M-matrix (see, for instance, Berman and Plemmons (1994)). As a result, B^{-1} is non-negative, from which the result follows immediately. □

Proof of Proposition 6.4. Let $i, j, k \in N$ be as in the statement. Below we explicitly characterise how the recursive Buchholz ratings vary as a function of A_{ik} and A_{ki}, provided that M_{ik} stays constant. Recall that r^{rb} is the unique solution of $\bar{M}x + \hat{s} = x$ such that $(r^{\text{rb}})^\top e = 0$. Hence, $(I - \bar{M})r^{\text{rb}} = \hat{s}$. Define $B = I - \bar{M}$ and $\breve{N} = N\backslash\{i,k\}$. Then, equation ℓ of the system $Br^{\text{rb}} = \hat{s}$ can be written as

$$\sum_{h\in\breve{N}} B_{\ell h} r_h^{\text{rb}} = \hat{s}_\ell - B_{\ell i} r_i^{\text{rb}} - B_{\ell k} r_k^{\text{rb}}. \tag{A.1}$$

with $\ell \in \breve{N}$. Define $\breve{B} \in \mathbb{R}^{(n-2)\times(n-2)}$ to be the matrix obtained from B by deleting the rows and columns corresponding to players i and k.

We prove now that $\breve{B}$ is invertible. Suppose, on the contrary, that there is an $y \in \mathbb{R}^{n-2}$, $y \neq 0$, such that $y^\top \breve{B}^\top = 0$. Let $\ell \in \breve{N}$ be such that $y_\ell = \max_{h\in\breve{N}} y_h$. We assume, without loss of generality, that $y_\ell > 0$. For each $h \neq \ell$, $B_{h\ell}^\top \leq 0$ and, hence, $-y_h B_{h\ell}^\top \leq -y_\ell B_{h\ell}^\top$, with equality only if $y_h = y_\ell$ or $B_{h\ell}^\top = 0$. Since $y^\top\breve{B}^\top = 0$, $\sum_{h\in\breve{N}\backslash\{\ell\}} -y_h B_{h\ell}^\top = y_\ell B_{\ell\ell}^\top$. Further, since $\sum_{h\in N} B_{h\ell}^\top = 0$, we have $\sum_{h\in\breve{N}\backslash\{\ell\}} -B_{h\ell}^\top = B_{\ell\ell}^\top + B_{i\ell}^\top + B_{k\ell}^\top \leq B_{\ell\ell}^\top$, with equality only if $B_{i\ell}^\top = B_{k\ell}^\top = 0$. Then, we have

$$y_\ell B_{\ell\ell}^\top = \sum_{h\in\breve{N}\backslash\{\ell\}} -y_h B_{h\ell}^\top \leq y_\ell \sum_{h\in\breve{N}\backslash\{\ell\}} -B_{h\ell}^\top \leq y_\ell B_{\ell\ell}^\top$$

and, hence, all the inequalities are indeed equalities. Therefore, $B_{i\ell}^\top = B_{k\ell}^\top = 0$ and, for each $h \in \breve{N}\backslash\{\ell\}$, $y_h = y_\ell$ or $B_{h\ell}^\top = 0$. Define $\bar{N} = \{m \in \breve{N} \,|\, y_m = \max_{h\in\breve{N}} y_h\}$. Now, for each $m \in \bar{N}$, we have $B_{im}^\top = B_{km}^\top = 0$ and, further, for each $h \in \breve{N}\backslash\bar{N}$, $B_{hm}^\top = 0$. That is, no player outside $\bar{N}$ has played against players inside $\bar{N}$, which contradicts the irreducibility of A.

Define $C = (\breve{B})^{-1}$, $\breve{r}^{\text{rb}} = (r_h^{\text{rb}})_{h\in\breve{N}}$, $B^i = (B_{hi})_{h\in\breve{N}}$ and $B^k = (B_{hk})_{h\in\breve{N}}$. Then, using (A.1) we have $\breve{B}\breve{r}^{\text{rb}} = \breve{s} - B^i r_i^{\text{rb}} - B^k r_k^{\text{rb}}$ and hence, $\breve{r}^{\text{rb}} = C(\breve{s} - B^i r_i^{\text{rb}} - B^k r_k^{\text{rb}})$. So, for all $\ell \in \breve{N}$,

$$r_\ell^{\text{rb}} = \breve{r}_\ell^{\text{rb}} = \sum_{h\in\breve{N}} C_{\ell h}(\breve{s}_h - B_{hi}r_i^{\text{rb}} - B_{hk}r_k^{\text{rb}}).$$

Define $\gamma_\ell^s = \sum_{h\in\breve{N}} C_{\ell h}\breve{s}_h$, $\gamma_\ell^i = -\sum_{h\in\breve{N}} C_{\ell h}B_{hi}$ and $\gamma_\ell^k = -\sum_{h\in\breve{N}} C_{\ell h}B_{hk}$. Then, for each $\ell \in \breve{N}$,

$$r_\ell^{\text{rb}} = \gamma_\ell^s + \gamma_\ell^i r_i^{\text{rb}} + \gamma_\ell^k r_k^{\text{rb}}. \tag{A.2}$$

Furthermore, equation i in $Br^{\text{rb}} = \hat{s}$ is

$$B_{ii}r_i^{\text{rb}} + B_{ik}r_k^{\text{rb}} + \sum_{\ell\in\breve{N}} B_{i\ell}r_\ell^{\text{rb}} = \hat{s}_i. \tag{A.3}$$

Define $\Gamma^{i,i} = -\sum_{\ell\in\breve{N}} B_{i\ell}\gamma_\ell^i$ and $\Gamma^{i,k} = -\sum_{\ell\in\breve{N}} B_{i\ell}\gamma_\ell^k$. Then, plugging in the expression of each r_ℓ^{rb} (A.2) into (A.3) we get

$$(B_{ii} - \Gamma^{i,i})r_i^{\text{rb}} + (B_{ik} - \Gamma^{i,k})r_k^{\text{rb}} = \hat{s}_i - \sum_{\ell\in\breve{N}} \gamma_\ell^s. \tag{A.4}$$

Now, adding up (A.2) over all $\ell \in \breve{N}$ and using that $\sum_{h\in N} r_h^{\text{rb}} = 0$,

$$(1 + \sum_{\ell\in\breve{N}} \gamma_\ell^i)r_i^{\text{rb}} + (1 + \sum_{\ell\in\breve{N}} \gamma_\ell^k)r_k^{\text{rb}} = -\sum_{\ell\in\breve{N}} \gamma_\ell^s. \tag{A.5}$$

Define $\sigma_i = \sum_{\ell\in\breve{N}} \gamma_\ell^i$ and $\sigma_k = \sum_{\ell\in\breve{N}} \gamma_\ell^k$. Then, solving equations (A.4) and (A.5), we get

$$r_i^{\text{rb}} = \frac{\hat{s}_i - (1 - \frac{B_{ik}-\Gamma^{i,k}}{1+\sigma_k})\sum_{\ell\in\breve{N}}\gamma_\ell^s}{(B_{ii} - \Gamma^{i,i}) - (B_{ik} - \Gamma^{i,k})\frac{1+\sigma_k}{1+\sigma_i}} \quad \text{and} \quad r_k^{\text{rb}} = \frac{-\sum_{\ell\in\breve{N}}\gamma_\ell^s}{1+\sigma_k} - \frac{1+\sigma_i}{1+\sigma_k}r_i^{\text{rb}}. \tag{A.6}$$

To understand how r_i^{rb} and r_k^{rb} vary with $\hat{s}_i$, it is convenient to know the signs of γ^i and γ^k. We claim that both γ^i and γ^k are nonnegative vectors. By definition, $\gamma^i = -CB^i$ and, since $C^{-1} = \breve{B}$, $\breve{B}\gamma = -B^i$. Furthermore, $-B^i \geq 0$. Since matrix $\breve{B}$ and vectors γ^i and $-B^i$ satisfy the conditions of Lemma A.1, γ^i is nonnegative. The argument for γ^k is analogous using $-B^k$ instead of $-B^i$. The nonnegativity of γ^i and γ^k implies that σ_i and σ_k are also nonnegative. Since γ^k is nonnegative, also $\Gamma^{i,k}$ is nonnegative and $B_{ik} - \Gamma^{i,k}$ is negative. Furthermore,

$$B_{ii} - \Gamma^{i,i} = B_{ii} + \sum_{\ell\in\breve{N}} B_{i\ell}\gamma_\ell^i \geq B_{ii} + \sum_{\ell\in\breve{N}} B_{i\ell} \geq 0.$$

We reexamine now equation (A.6). Note that γ^s, γ^i, γ^k, $\Gamma^{i,i}$, $\Gamma^{i,k}$, B_{ii} and B_{ik} only depend on $\breve{B}$. Then, the denominator of the expression for r_i^{rb} is positive and so r_i^{rb} is strictly increasing in $\hat{s}_i$. Further, since r_k^{rb} is strictly decreasing in r_i^{rb}, it is strictly decreasing in $\hat{s}_i$.

Now, because of (A.2), r_ℓ^{rb} is weakly increasing in r_i^{rb} and r_k^{rb}. Yet, since r_i^{rb} and r_k^{rb} are strictly increasing and decreasing, respectively, in $\hat{s}_i$, some extra work is needed to understand how r_ℓ^{rb} varies with $\hat{s}_i$. To do so, we first show that all the components of γ^i and γ^k are no larger than 1. We prove it for γ^i, the proof for γ^k being analogous.

$$\breve{B}(e-\gamma^i) = \breve{B}e - \breve{B}\gamma^i = \breve{B}e - B^i,$$

and, for each $\ell \in \breve{N}$,

$$(\breve{B}e - B^i)_\ell = \sum_{h\in\breve{N}} B_{\ell h} + B_{\ell i} \geq \sum_{h\in\breve{N}} B_{\ell h} + B_{\ell i} + B_{ki} = 0.$$

Then, since $\breve{B}e - B^i$ is a nonnegative vector, matrix $\breve{B}$ and vectors $e-\gamma^i$ and $\breve{B}e - B^i$ are in the conditions of Lemma A.1 and, hence, $e-\gamma^i$ is nonnegative.

Therefore, we know that all the components of γ^i and γ^k are no larger than 1. Looking again at equation (A.2), we have that r_ℓ^{rb} cannot increase with $\hat{s}_i$ faster than r_i^{rb} so $r_\ell^{\mathrm{rb}}/r_i^{\mathrm{rb}}$ is weakly decreasing in $\hat{s}_i$. Similarly, $r_\ell^{\mathrm{rb}}/r_k^{\mathrm{rb}}$ is weakly increasing in $\hat{s}_i$. From this, the statement follows. □

Proof of Proposition 6.5. The proof of is analogous to the proof of Proposition 6.4, with $B = L_A - A$ instead of $B = I - \bar{M}$. □