On Tanking and Competitive Balance: Reconciling Conflicting Incentives*

Aleksandr M. Kazachkov and Shai Vardi

March 22, 2020

Abstract

Sports leagues aim to promote competitive balance among teams by giving worse teams the opportunity to pick better incoming players in an end-of-season draft. This creates a perverse incentive for teams to misrepresent their true quality by purposefully losing games. Though many proposals exist to reduce tanking, these mostly ignore the effect on long-term competitive balance, an important consideration as attempts to disincentive tanking can lead to a more inaccurate ranking of teams, inhibiting the success of the draft. We introduce a stylized model of a sports league to simultaneously assess the effects of the draft on both tanking and competitive balance. In addition, we propose and analyze a new bilevel ranking mechanism, in which the ranking of non-playoff teams is based on their relative order after a preset breakpoint game. We precisely characterize team tanking strategies under the bilevel ranking and present simulation results comparing it to the system currently used by the National Basketball Association (NBA) and a proposal based on mathematical elimination ordering. We show that the bilevel ranking not only reduces tanking, but that it can also *increase* the competitive balance in the league relative to other ranking systems, including the one currently used by the NBA.

1 Introduction

Look, losing is our best option.
— Mark Cuban, Dallas Mavericks Owner, 2018

Sports leagues are decentralized markets in which classic principal-agent problems frequently arise, due to conflicting objectives of the teams and league management. In this paper, we focus on the league's pursuit of long-term competitive balance, or parity, which leads to perverse incentives for teams to tank—purposely lose games.

A prominent mismatch between league and team incentives occurs as a result of the *entry draft*, in which teams select players that are incoming to the league, typically in round-robin fashion. In the interest of competitive balance, the league seeks to assign better draft positions to worse teams. Unfortunately, the league cannot know which teams are truly better or worse, and so, as a proxy, uses a ranking of the teams based on their performance.

*An earlier version of this paper, under the title "On Tanking and Competitive Balance", was presented at the *Conference on Algorithmic Decision Theory* in October 2019.

Thus, a team finishing in a worse position in the league has a higher chance at drafting better players, giving teams an incentive to tank.

Although it is difficult to prove that a team loses intentionally, it is widely accepted that tanking is pervasive in the major sports leagues in the United States: football [Bar17], hockey [McI16], baseball [Pay18], and basketball [Abb12]. A plethora of statistical evidence supports these popular opinions on the prevalence of tanking in these leagues [For18]. The effect appears to be especially pronounced in the National Basketball Association (NBA) [TT02, PSBH10, SH13, KL17] because single players can disproportionately bring success to the franchise for many years, creating economic value through larger fanbases, increased sponsor interest, and more sales revenue [WW12, Sau18].

While tanking may benefit teams acting in their own selfish interests, it is detrimental to the league. In addition to being fundamentally unethical [McM18], it can decrease fan and sponsor engagement, and league revenues as a consequence [Bin13, Fri12, Soe11]. Somewhat ironically, tanking confounds the league's ability to correctly rank the teams, undermining the draft's purpose of promoting competitive balance, which hinges on the league obtaining an accurate ranking. These reasons have motivated leagues to try to reduce tanking by reforming their drafts. The NBA has reformed the draft multiple times, yet each time reforms were not considered sufficiently successful at mitigating tanking. The most recent change was made in 2019, yet even the NBA commissioner believes that these "new tanking reforms may not be enough to address the issue" [Beg18].

Many draft reforms have been suggested in order to eliminate or reduce tanking. Unfortunately, it is unclear how to compare them. How can we objectively decide if one proposal is "better" than another? To remedy this, we introduce a flexible model by which to compare the different proposals with respect to their effects on tanking and competitive balance.

Contributions. Our first contribution is a stylized model of a sports league that allows us to objectively compare the effect of different draft mechanisms on tanking and competitive balance. Assuming that there is a true full ranking of the teams (from best to worst) and the optimal draft (in terms of achieving competitive balance) uses the reverse of this order, we can quantify the effect of a draft order on competitive balance as the statistical distance between the two orders: the closer the draft order is to the reverse of the true ranking, the more competitive the league is expected to be in the long run. To quantify the amount of tanking, we consider both the total number of games tanked in a season and the number of the teams that tank.

Our second contribution is a *bilevel ranking* system that separates the ranking of the playoff teams from the ranking of the non-playoff teams. The teams that advance to the playoffs are ranked the same way that they are currently ranked, by win percentage at the end of the season. The ranking of the non-playoff teams is decided based on their relative ranking at a *breakpoint* in the season. Optionally, this ranking can then be used as an input to a *draft lottery* used to decide the final draft positions. The bilevel ranking system captures many of the suggested mechanisms for setting draft positions, including the system currently used by the league, in which the breakpoint is simply the end of the season.

Intuitively, the appeal of the bilevel ranking is reducing the incentive of teams to tank after the breakpoint, as teams may not be able to improve their draft position by losing in games after the breakpoint. However, as we show in our analysis, there is a clear tradeoff for the league: setting the breakpoint earlier in the season will result in fewer tanked games, but setting it too early would prevent the league from having a sufficiently accurate ranking of the teams, due to the league having fewer samples of games between every pair of teams.

To evaluate our proposal, we compare the bilevel ranking system under different breakpoints to the current draft lottery system in the NBA, as well as to the mathematical elimination ordering due to Lenten [Len16]. We measure the effect on competitive balance and tanking via a Monte Carlo simulation implementing our stylized sports league model. A key aspect of our model is the decision of each team to try to win or lose a game. While some teams tank when presented the opportunity, it is believed that others consistently choose not to, due to, e.g., moral reasons [Abb12, Dee13]. We therefore partition the teams into those that *never* tank (moral teams) and ones that do tank if they feel it will help them attain a better draft pick (selfish teams).

While moral teams never tank, it is not clear a priori how selfish teams should make their decision. The third, and main theoretical, contribution of this paper is a complete characterization of the optimal behavior of selfish teams throughout the season. This includes the counterintuitive result that there exist circumstances in which a team's optimal behavior is to tank even after the breakpoint. These theoretical results enable our simulation experiments.

In Section 5, we present the results of our simulations, as well as validating our assumptions and parameters chosen using real NBA data. Our main insights are the following.

- Setting the breakpoint between 5/6 and 7/8 of the season reduces the number of tanked games by 57–72% compared to setting it at the end of the season.
- Setting the breakpoint between 5/6 and 7/8 of the season gives strictly better competitive balance than ordering based on mathematical elimination ordering [Len16] or the current NBA system, regardless of the number of selfish teams.
- Setting the breakpoint between 5/6 and 7/8 of the season gives strictly better competitive balance than setting it at the end of the season when around 1/4 to 3/4 of the teams are selfish.

Our experiments show that the current system used by the NBA hurts competitive balance compared to our proposed bilevel ranking. With an appropriately set breakpoint, which the league can identify based on, e.g., the estimated strength of the subsequent year's draft, the bilevel ranking can lead to both long-term improvements in competitive balance and reduced tanking. Our bilevel ranking mechanism is not only theoretically sound, but also practical and quickly deployable by sports leagues.

1.1 Related literature

There is a long line of literature on competitive balance in sports, from both theoretical and empirical perspectives. The theoretical work raises questions such as whether equalizing the strength of teams is consistent with profit maximization [EHQ71], and whether revenue sharing leads to more competitive balance [Kes00]. Empirical works include the effects of the changes in the business practices of leagues on competitive balance [FQ95] and the effects of competitive balance on, e.g., attendance [SB01] and revenue [Bin13]. For reviews of the vast literature on competitive balance, see [FM03, LvAAM18, Soe11]; Soebbing [Soe11] in particular extensively reviews the literature on the NBA draft and tanking.

Several papers have analyzed the effects on tanking of the systems the NBA has used for the draft. Taylor and Trogdon [TT02] find evidence for tanking in the NBA under the reverse order draft (no lottery) and the weighted lottery, which is similar to the current system. They find no evidence of tanking when the non-playoff teams were allocated equal weighting in the draft, which is consistent with our theoretical results. Price et al. [PSBH10] find that teams are more likely to tanks at the end of the regular season when the incentives to finish last were the largest.

Over the years, a myriad of draft reforms have been suggested in order to eliminate or reduce tanking. Arguably the simplest method to eliminate tanking is to let the order in the draft be uniformly at random. If finishing lower in the ranking offers no advantage, this completely eliminates the incentive to tank for this reason. There may still be other incentives to tank, e.g., a playoff team may tank in order to play against a specific rival in the first round of the playoffs. Two popular recent proposals are based on *mathematical elimination ordering* [Gol12, Len16]; a team is mathematically eliminated when there is no scenario in which it makes the playoffs.

Lenten [Len16] proposes to rank teams in the order that they are mathematically eliminated, which may sound compelling, but it suffers from several drawbacks: (i) it is computationally challenging to decide whether a team is mathematically eliminated (see, e.g., [HR70, SC18]). As a result, this method lacks transparency, and it will be unappealing to fans to not have a rule they can easily verify; (ii) teams will most likely tank when they believe their chances of making the playoffs are negligible, which often occurs well before they have been mathematically eliminated; and (iii) it is unclear what effect this has on competitive balance. Gold [Gol12] suggests to rank teams eliminated from the playoffs based on the number of wins they have had since being mathematically eliminated. This proposal shares all the shortcomings of [Len16], but seems much less likely to promote competitive balance. Consider, for instance, a team that is not good enough to win any games at all in a season. It would be likely to receive the worst draft pick under this system.

Most recently, contemporaneously with the present paper, Banchio and Munro [BM20] propose a new dynamic lottery system in which draft odds are adjusted after each game in a way that incentivizes teams to exert full effort in each game. The promising approach of [BM20] shares several of the motivations and advantages of the bilevel ranking mechanism, but may be significantly more complex to convey to fans and other stakeholders. In future work, we intend to add this new reformed lottery to our simulations.

Many other ideas to eliminate or reduce tanking have been proposed (e.g., [DLB14, Sil15, Cas19]; see also [Avi15]). While none have been analyzed with respect to their effect on competitive balance, and many of them have not been rigorously analyzed at all, they all appear to suffer from at least one of the following three drawbacks: (1) they fail to adequately prevent tanking, (2) they do not sufficiently promote competitive balance, or (3) they are difficult to compute or verify.

Our bilevel mechanism has several distinct advantages, in comparison. It is effective at reducing tanking incentives, without sacrificing the quality of the ranking of the non-playoff teams, and, in fact, it even improves the ranking over the status quo under modest assumptions. Moreover, it is simple to implement and understand.

Although our focus is on teams that lose in order to obtain a better draft postion, there has been a substantial amount of recent work on other scenarios where tanking occurs. One example is the badminton controversy from the 2012 Olympics [HK12], which inspired theoretical results showing that, under broad conditions, two-stage tournaments cannot be entirely strategyproof (incentive compatible) unless exactly one team is allowed to qualify to the second stage from each group in the first stage [Pau14, Von17]. In other cases, tanking may occur when it can benefit a larger group [CDL11]. In the present paper, we ignore these other possibilities and exclusively consider tanking with respect to the draft, but we refer the reader to recent surveys of the broad subject of such market failures [Wri14, KL17].

2 Model

We describe a flexible model of a generic sports league with a regular season, playoffs, and a draft. Although our model is stylized, our simulation experiments with the model in Section 5 illustrate that it is sufficiently rich to capture key characteristics of a real sports league, by comparing the simulations with historical NBA data. At the same time, in Section 4, we are able to theoretically analyze team strategies under this model. We denote the set $\{1, \ldots, n\}$ by $[n]$.

Structure of the league. There are n teams, labeled $1, \ldots, n$, that play a regular season of T games, with each team playing an equal number of games, and $n^* < n$ teams make the playoffs. We do not model the playoff games, only the regular season. Games are played one at a time, and the regular-season game order is fixed before the season starts and known by all teams. Exactly one team wins each game; there are no ties.

A *ranking* π of the teams is a mapping of the teams to positions, i.e., $\pi_i \in [n]$ is the position of team i. If $\pi_i < \pi_j$, then we say team i is ranked *better* than team j under π. We assume that there exists a *true ranking* of the teams from best to worst, i.e., a static total ordering π^{TRUE} that is unknown to the league. After the season, the league sets a *league ranking* π^{LEAG} of the teams and hosts a draft, in which teams select players in a round-robin fashion. Teams pick in the draft in the reverse order of their standings in π^{LEAG}. The league can use randomization when setting π^{LEAG}, e.g., by imposing a *lottery* (randomization of the order) after setting an initial ranking. The model therefore captures the current NBA system, which we describe in detail in Section 5.2.

League objectives. In order to ensure competitive balance across seasons, the league aims to determine a draft order that is close to the optimal draft order, the reverse of π^{TRUE}. A secondary goal of the league is to minimize the expected number of games tanked per season.

Formally, for the first objective, the league seeks to set π^{LEAG} to minimize the (expected) distance between π^{LEAG} and π^{TRUE}, where distance is computed via the standard *Kendall tau distance* [KG90],[1] the number of pairs of teams that are ordered differently by the two

[1] Alternatively referred to as bubble sort distance, swap distance, or inversion distance.

rankings. More precisely, the Kendall tau distance between two rankings π and ρ is

$$d_K(\pi, \rho) := |\{(i, j) : (\pi_i > \rho_j) \wedge (\pi_i < \rho_j)\}|. \tag{1}$$

As an example, in a league with four teams in which the true ranking is $\pi^{\text{TRUE}} = \mathbf{a} \succ \mathbf{b} \succ \mathbf{c} \succ \mathbf{d}$ and the league ranking is $\pi^{\text{LEAG}} = \mathbf{b} \succ \mathbf{c} \succ \mathbf{a} \succ \mathbf{d}$, the Kendall tau distance $d_K(\pi^{\text{LEAG}}, \pi^{\text{TRUE}}) = 2$, as the pairs $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{a}, \mathbf{c})$ are ordered differently in π^{LEAG} and π^{TRUE}.

Probabilities and tanking. Every team privately decides whether to exert *high* (H) or *low* (L) effort in each game, representing whether the team tries to win or lose (tank) that game. For every pair of teams i and j, there is some probability $p_{ij}(e_i, e_j)$ that team i beats team j when the teams effort levels are $e_i, e_j \in \{L, H\}$. As there are no ties in our model, $p_{ji}(\cdot, \cdot) = 1 - p_{ij}(\cdot, \cdot)$. If team i is better than team j under π^{TRUE}, i.e., $\pi_i^{\text{TRUE}} < \pi_j^{\text{TRUE}}$, then $p_{ij}(H, H) \geq 1/2$. In addition, we assume that

$$p_{ij}(L, H) < p_{ij}(L, L) < p_{ij}(H, L), \text{ and}$$
$$p_{ij}(L, H) < p_{ij}(H, H) < p_{ij}(H, L).$$

Properly calibrating these values is critical for the model to have high fidelity to real leagues. We focus on two of the most widely-used options among a vast literature on generating rankings from pairwise comparisons [Cat12]: the (Zermelo-)Bradley-Terry [Zer29, BT52, Luc59] and noisy comparison [BM08, RV17] models. Both models are described in Section 5.1, with further details in Appendix A on how we choose between the models for our simulation experiments and how we validate our simulations using historical NBA data.

Team objectives. We assume that each team's utility is simply a function of their rank in π^{LEAG}. Let $u_i(r)$ be the (positive-valued) utility team i has for being ranked r in π^{LEAG}. We assume that team i's utility satisfies

$$u_i(1) > u_i(2) > \cdots > u_i(n^*) > u_i(n) > u_i(n-1) > \cdots > u_i(n^* + 1) > 0. \tag{2}$$

Throughout the paper, we will assume that $u_i(n^*) \gg u_i(n)$; that is, teams would greatly prefer making the playoffs to obtaining a better draft pick. We elaborate om this below.

Each team decides whether to exert high or low effort in a game based on the team's *subjective expected utility* of winning or losing the game. This is based on the team's subjective perceptions of the relative likelihood of the possible *scenarios* for the rest of the season. Formally, let *scenario* $\mathcal{O}$ denote a fixed set of outcomes for a sequence of games. Team i assigns every scenario $\mathcal{O}$ a (subjective) probability $\mathbf{p}^i(\mathcal{O}) > 0$.

We denote the set of all possible scenarios for games t' through $t'' \geq t'$ by $\Omega^{t',t''}$. Suppose team i is playing in game t with the outcomes of games 1 through $t-1$ fixed. For a scenario $\mathcal{O} \in \Omega^{t,T}$ for the rest of the season, let $\Pi_i(\mathcal{O})$ denote the set of possible ranks team i can attain under the league ranking π^{LEAG} when $\mathcal{O}$ is realized.

Let $\mathbf{Pr}[\pi_i^{\text{LEAG}} = r \mid \mathcal{O}]$ be the probability (over a tie-breaking rule) that $\pi_i^{\text{LEAG}} = r$ under scenario $\mathcal{O}$. Define the *scenario expected utility* of scenario $\mathcal{O}$ for team i to be

$$u_i(\mathcal{O}) := \sum_{r \in \Pi_i(\mathcal{O})} \mathbf{Pr}[\pi_i^{\text{LEAG}} = r \mid \mathcal{O}] \cdot u_i(r).$$

Denote the outcome of a game by W_i when team i wins, and L_i when team i loses. Given a scenario $\mathcal{O}$, we define scenario $\mathcal{O}^W$ as $\{W_i\} \cup \mathcal{O}$ and $\mathcal{O}^L$ by $\{L_i\} \cup \mathcal{O}$. We can now define team i's *subjective expected utility* of winning game t to be

$$U_i^{\mathbf{p}}(W_i) := \sum_{\mathcal{O} \in \Omega^{t,T}} \mathbf{p}^i(\mathcal{O}) \cdot u_i(\mathcal{O}^W),$$

The definition of $U_i^{\mathbf{p}}(L_i)$ is analogous.

There are two types of teams: *selfish* teams, whose goal is to maximize their subjective expected utility, and *moral* teams, whose goal is always to win as many games as possible, or "play their best". The league does not know which teams are selfish or moral.

Choosing when to tank. We assume that if the subjective expected utility of winning or losing a game is identical, then a selfish team will prefer to win. When team i plays in game t after the first $t-1$ game outcomes are fixed, we have the following definition for when a team will prefer to exert low or high effort.

Definition 1. *If $U_i^{\mathbf{p}}(W_i) \geq U_i^{\mathbf{p}}(L_i)$ for all possible probabilities $\mathbf{p}$, then exerting high effort is dominant for team i in game t. If $U_i^{\mathbf{p}}(W_i) < U_i^{\mathbf{p}}(L_i)$ for all possible probabilities $\mathbf{p}$, then exerting low effort is dominant for team i in game t.*

This model is general in that it puts very few restrictions on team beliefs. We focus on two special cases, which we call *optimistic* and *conservative* decision-making processes:

- The *optimistic* model. Under this model, teams are optimistic when there is a chance of them making the playoffs. Formally, let

  $$p_i^* := \min\{\mathbf{p}^i(\mathcal{O}) : \mathcal{O} \in \Omega^{t,T}, \min\{\Pi_i(\mathcal{O})\} \leq n^*\}$$

  denote the minimum probability that team $i \in [n]$ assigns to any scenario in which it can make the playoffs, over the entire season. Set $\epsilon_i := u_i(n)/u_i(n^*)$. We assume that $\epsilon_i \leq p_i^*/(n + p_i^*)$, i.e., that

  $$p_i^* \geq n\epsilon_i/(1-\epsilon_i). \tag{3}$$

  The optimistic model is strongly connected to mathematical elimination, as we show in Section 4. This is particularly useful because mathematical elimination is the basis for analyzing team decision-making in much of the literature [Gol12, Len16].

- The *conservative* model. In this model, we assume more structure on how teams update their beliefs based on partial season standings. Specifically, rather than waiting for mathematical elimination, we assume a team checks after each game whether it is *effectively eliminated*: we say a team team is effectively eliminated if the win percentage it would have after winning all of its remaining games is less than the current win percentage of the team ranked n^*. In the conservative model, when a team is effectively eliminated, we assume that it assigns $\mathbf{p}^i(\mathcal{O}) = 0$ for any scenario $\mathcal{O}$ in which the team may make the playoffs.

We note that it is possible a team is effectively eliminated but not mathematically eliminated, implying that the status of being effectively eliminated is not necessary permanent once it has been attained, in contrast to mathematical elimination. However, this occurs rarely, as we discuss in Section 5.

3 Bilevel Ranking

We introduce the *bilevel ranking* to set π^{LEAG} in the model of Section 2. A publicly known *breakpoint* game δ is set. The *win-rank* of a team is its position in an ordering of the teams based on their win percentage (where the highest win percentage is first in the order). Denote the win-rank of the teams after game δ by π^{δ}.

> **Bilevel ranking**
>
> The n^* teams with best win-rank at the end of the regular season, i.e., the playoff teams, are ranked in π^{LEAG} by their win-rank at in π^T. The teams that do not make the playoffs are assigned ranks $n^* + 1, \ldots, n$ according to their relative win-rank in π^{δ}.

We refer to the *bilevel-rank* of a team as its position in π^{LEAG} when the bilevel ranking is used. Next, we give an example of the bilevel ranking on a 6-team league when 3 teams make the playoffs.

Example 1. Consider $n = 6$ teams, labeled $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}$, and let $n^* = 3$ teams make the playoffs.

Win-rank at end of season (π^T)	**a b c d e f**
Win-rank at breakpoint (π^{δ})	**b e a f c d**
Relative win-rank of teams that make the playoffs	**a b c**
Relative win-rank of teams that do not make the playoffs	**e f d**
Bilevel ranking	**a b c e f d**

The bilevel ranking does not change the ranks of the playoff teams, whereas the non-playoff teams are ranked in the same relative order that they were ranked at the breakpoint. ■

As mentioned in Section 2, the league may choose to use the bilevel ranking as an input into a draft lottery, in which draft positions are allocated probabilistically, where team i's expected draft position is at least as good as team j's if $\pi_i^{\text{LEAG}} > \pi_j^{\text{LEAG}}$. The current system employed by the NBA corresponds to a lottery combined with the bilevel ranking when $\delta = T$. Setting a different δ can be used with or without the lottery, though we focus our analysis on the case of no lottery.

The mathematical elimination ordering of **(author?)** [Len16] can be seen as a straightforward extension of the bilevel ranking into a multilevel ranking, in which each team has its own breakpoint δ set to the specific game in the season in which the team is mathematically eliminated. For simplicity, we concentrate on the model as defined and discuss these possible extensions in the conclusion.

If π^{LEAG} is set using the bilevel ranking, the league's objective becomes to optimize δ to trade off its effect on competitive balance and tanking. Intuitively, the incentive to tank in order to obtain a better position in the draft is virtually eliminated after the breakpoint. However, though setting the draft breakpoint earlier will reduce tanking, it will also result in lower accuracy of the ranking of non-playoff teams, as their rank is then determined based on playing fewer games. On the other hand, setting δ closer to the end of the season will lead to more tanking incentives. The challenge we address in the subsequent sections is how the league should optimize over δ while considering competitive balance and tanking simultaneously.

4 Selfish Team Strategies

In this section, we prove our main theoretical result, Theorem 2, which characterizes whether a selfish team's optimal strategy is to exert high or low effort in any given game. We adopt the optimistic decision-making model from Section 2. Our results can be extended to the conservative model, at the expense of clarity in the theoretical statements (and with few, if any, additional insights).

4.1 Notation

We analyze a selfish team i playing in game t against team j, when the outcomes of games 1 through $t-1$ are known, and the outcomes of games t through T are not. Our analysis implicitly conditions on the first $t-1$ outcomes being fixed, and dependence on t is also typically implicit in the notation we introduce. Recall that a scenario $\mathcal{O}$ is a fixed set of outcomes for a sequence of games. We use the shorthand Ω^T for $\Omega^{t+1,T}$, the set of scenarios for games $t+1$ through T, as these will be the most common scenarios we consider. Let Π_i^{best} and Π_i^{worst} be the best and worst rank in π^{LEAG} that team i can attain across all possible scenarios in $\Omega^{t,T}$ for games t through T, respectively.[2]

Define τ_i as the first game in which team i is mathematically eliminated.[3] Formally, team i is mathematically eliminated before game t, i.e., $\tau_i \leq t$, if and only if $\Pi_i^{\text{best}} > n^*$. The set of teams can be partitioned into three categories: those that make the playoffs in every scenario, those that are mathematically eliminated, and those whose playoff chances are not settled. We define this last group of teams as *contenders* for the playoffs. Formally, team i is a *contender under scenario* $\mathcal{O}$ in game t when, fixing the outcomes of the games given by scenario $\mathcal{O}$, there exists a scenario for the remaining games in the season in which team i makes the playoffs and a scenario in which it does not.

For a scenario $\mathcal{O} \in \Omega^{t,T}$, let $\Pi_i(\mathcal{O})$ denote the set of possible ranks team i can attain under the bilevel ranking π^{LEAG} when $\mathcal{O}$ is realized. More generally, when $\mathcal{O}$ is a scenario for a *strict* subset of games t through T, define $\Pi_i(\mathcal{O})$ as the union over all scenarios in $\Omega^{t,T}$ that are supersets of $\mathcal{O}$:

$$\Pi_i(\mathcal{O}) := \bigcup_{\substack{\hat{\mathcal{O}} \in \Omega^{t,T}: \\ \mathcal{O} \subset \hat{\mathcal{O}}}} \Pi_i(\hat{\mathcal{O}}).$$

Overloading notation, define $\Pi_i := \bigcup_{\mathcal{O} \in \Omega^{t,T}} \Pi_i(\mathcal{O})$. In other words, Π_i is all the possible bilevel ranks of team i at the end of the season.

Due to the bilevel nature of π^{LEAG}, reasoning directly about Π_i is complicated. Instead, we first analyze π^δ and π^T, the first-level rankings by win-rank after games δ and T, which are the building blocks of Π_i. For any $t' \geq t$ and $\mathcal{O} \in \Omega^{t,t'}$, let $\Psi_i(\mathcal{O})$ denote the set of possible win-ranks that team i can attain after realization of scenario $\mathcal{O}$, under all possible

[2] Recall that the best possible rank is 1; the worst possible rank is n.

[3] For a team that is never eliminated, we set $\tau_i = T + 1$.

T	Number of games in the season
n	Number of teams
n^*	Number playoff teams
π^{LEAG}	The bilevel ranking of the teams at the end of the season
π^t	Ranking of the teams in order of decreasing winning percentage after game t
δ	Breakpoint game, used for ranking non-playoff teams in π^{LEAG}
τ_i	The first game in which team i is mathematically eliminated
$\mathcal{O}$	A set of outcomes (*scenario*)
$\Omega^{t',t''}$	The set of all possible scenarios for games t' through $t'' \geq t'$
Ω^T	Equivalent to $\Omega^{t+1,T}$, the set of scenarios for games $t+1$ through T
$\mathcal{O}^W$	A scenario $\mathcal{O}$ appended with the outcome that team i wins game t
$\mathcal{O}^L$	A scenario $\mathcal{O}$ appended with the outcome that team i loses game t
$\Pi_i(\mathcal{O})$	The set of all possible bilevel-ranks in π^{LEAG} that team i can attain under $\mathcal{O}$
Π_i	The set of ranks in π^{LEAG} attainable by team i at the end of the season
Π_i^{best}	Best bilevel-rank in π^{LEAG} that team i can achieve
Π_i^{worst}	Worst bilevel-rank in π^{LEAG} that team i can achieve
$\Psi_i(\mathcal{O})$	Positions team i can attain under scenario $\mathcal{O}$ and ranking by win totals
$\Psi_i^{\text{best}}(\mathcal{O})$	Best win-rank in $\Psi_i(\mathcal{O})$ that team i can achieve under scenario $\mathcal{O}$
$\Psi_i^{\text{worst}}(\mathcal{O})$	Worst win-rank in $\Psi_i(\mathcal{O})$ that team i can achieve under scenario $\mathcal{O}$

Table 1: Summary of main notation

tie-breaks. As all teams are assumed to have played the same number of games after games δ and T, ranking by number of wins is the same as win-rank. Let $\Psi_i^{\text{best}}(\mathcal{O})$ and $\Psi_i^{\text{worst}}(\mathcal{O})$ be the best and worst win-ranks of team i at π^T respectively.

We assume that all teams have played the same number of games after the conclusion of game δ. It is easy for the league to set a schedule that satisfies this assumption.

For reference, Table 1 contains a summary of the notation.

4.2 Theorem statement

We will prove the following theorem characterizing selfish team strategies, in order to understand how varying δ will affect team incentives. The theorem is divided into two cases, depending on whether the team is eliminated before or after δ.

Theorem 2. *Consider selfish team i playing against team j in game t. Exerting low effort is dominant for team i if and only if either $\tau_i \leq t \leq \delta$ and $\Pi_i^{\text{best}} \neq \Pi_i^{\text{worst}}$, or $\delta < \tau_i \leq t$ and (1) $\pi_i^\delta < \pi_j^\delta$, i.e., team j was ranked worse than team i at game δ, and (2) there exists a team k with $\pi_k^\delta < \pi_i^\delta$ and a scenario $\mathcal{O} \in \Omega^T$ such that teams j and k are both contenders under scenario $\mathcal{O}$.*

Theorem 2 follows from Lemmas 9, 12, 13, and 14, which cover the different cases, as summarized in Figure 1.

Counterintuitively, we show that there exist scenarios in which a team can benefit from exerting low effort in a game after δ. To give some insight into why a team may exert low

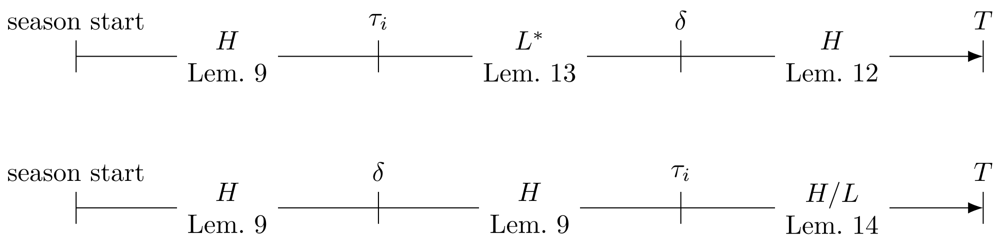

Figure 1: Effort levels (H: high, L: low) of selfish teams, depending on the timing of τ_i and δ with respect to t. *When $\tau_i \leq t \leq \delta$, teams exert low effort unless $\Pi_i^{\text{best}} = \Pi_i^{\text{worst}}$.

effort in a game after δ, suppose team i could still have made the playoffs at game δ but is subsequently eliminated. There might exist a team j that was ranked worse than team i at game δ, but, unlike team i, team j can still make the playoffs at game t. Then team i is incentivized to have team j make the playoffs, giving team i a worse bilevel-rank and higher utility. The theorem shows that, save for the realization of this unlikely scenario when $\delta < \tau_i \leq t$, teams will exert high effort as a dominant strategy after the breakpoint game δ. It is important to note that for any game $t > \delta$ in which a selfish team would exert low effort, that team would also have exerted low effort if δ were after game t (and the outcomes of the first $t-1$ games remained unchanged).

4.3 The Crossing Lemma and preliminaries

Before proving Theorem 2, we state and prove some results that will be repeatedly invoked, the primary one being Lemma 6, which concerns scenarios that *cross* a win-rank r with respect to team i, by which we mean that team i is neither guaranteed to have win-rank r or better, nor is it guaranteed to have a strictly worse win-rank than r.

Definition 3. Scenario $\mathcal{O}$ *crosses* r with respect to team i if $\Psi_i^{\text{best}}(\mathcal{O}^W) \leq r < \Psi_i^{\text{worst}}(\mathcal{O}^L)$.

Note that team i is a contender under scenario $\mathcal{O} \in \Omega^T$ if and only if scenario $\mathcal{O}$ crosses n^* with respect to team i.

Before stating the Crossing Lemma, we give some intuition to its importance to proving Theorem 2. Part of Theorem 2 states that before a team has been eliminated, it will exert high effort. While this may seem intuitive, we need to prove that this is *always* the case. Consider the following hypothetical situation, in which there are two games remaining: team i against team j, and team u against team v. Suppose that if team u wins the second game, then team i makes the playoffs, and finishes in position n^*, regardless of whether team i wins or loses the first game. If team v wins the second game, then suppose that team i will finish in position n^*+1 when it wins the first game, and in position n^*+2 when it loses that game. It is easy to see that in this situation, team i should exert low effort, in contradiction to Theorem 2. The Crossing Lemma shows that this type of situation cannot occur.

We need the following simple but important lemma, which states that when all other game outcomes are fixed, team i's worst possible rank after winning game t is at least as good as its best possible rank after losing game t.

Lemma 4. *For any scenario* $\mathcal{O} \in \Omega^{t+1,t'}$, $t+1 \leq t'$, $\Psi_i^{\text{worst}}(\mathcal{O}^W) \leq \Psi_i^{\text{best}}(\mathcal{O}^L)$.

Proof. As all game outcomes except t are fixed by scenario $\mathcal{O}$, we can assume without loss of generality that t is the last game (i.e., switch t and t'). Let ℓ be the number of teams that have strictly more wins than team i before game t. Hence, $\Psi_i^{\text{worst}}(\mathcal{O}^W) = \ell + 1 \leq \Psi_i^{\text{best}}(\mathcal{O}^L)$. □

The following observation is useful to track the possibilities for team i in a scenario $\mathcal{O}$.

Claim 5. *For any scenario $\mathcal{O}$ and $r \in [n]$, exactly one of the following holds:*

1. $\mathcal{O}$ crosses r *with respect to team i, or*
2. $\Psi_i^{\text{worst}}(\mathcal{O}^L) \leq r$*, or*
3. $r < \Psi_i^{\text{best}}(\mathcal{O}^W)$*.*

Proof. From Lemma 4, the second and third cases cannot occur together, as

$$\Psi_i^{\text{best}}(\mathcal{O}^W) \leq \Psi_i^{\text{worst}}(\mathcal{O}^W) \leq \Psi_i^{\text{best}}(\mathcal{O}^L) \leq \Psi_i^{\text{worst}}(\mathcal{O}^L).$$

The claim then follows, as $\mathcal{O}$ crosses r if and only if the other two cases both do not hold. □

We now state our main technical lemma, which shows that, if team i can attain win-rank $\pi_i^T \leq r$ in one scenario, and team i's best possible win-rank is $\pi_i^T > r$ in another scenario, then there exists a third scenario that crosses r with respect to i.

Lemma 6 (Crossing Lemma)**.** *Let Ω denote $\Omega^{t+1,t'}$, $t' \geq t+1$. If there exist $\dot{\mathcal{O}}, \ddot{\mathcal{O}} \in \Omega$ and $r \in [n]$ such that $\Psi_i^{\text{worst}}(\dot{\mathcal{O}}^L) \leq r < \Psi_i^{\text{best}}(\ddot{\mathcal{O}}^W)$, then there exists some $\widehat{\mathcal{O}} \in \Omega$ such that $\widehat{\mathcal{O}}$ crosses r.*

Proof. Assume the contrary, i.e., that for every $\widehat{\mathcal{O}} \in \Omega$, it holds that $\Psi_i^{\text{worst}}(\widehat{\mathcal{O}}^L) \leq r$ or $r < \Psi_i^{\text{best}}(\widehat{\mathcal{O}}^W)$. If this assumption holds, there must exist two scenarios $\mathcal{O}_a$ and $\mathcal{O}_b$ from Ω that differ in the outcome of single game, for which $\Psi_i^{\text{worst}}(\mathcal{O}_a^L) \leq r$ and $r < \Psi_i^{\text{best}}(\mathcal{O}_b^W)$. (To see this, lexicographically sort the scenarios in Ω, such that adjacent scenarios differ in the outcome of exactly one game.) Hence, we can assume w.l.o.g. that $\dot{\mathcal{O}}$ and $\ddot{\mathcal{O}}$ differ in the outcome of a single game, $\bar{t}$. We show that this yields a contradiction. Specifically, we show that for any two scenarios $\dot{\mathcal{O}}, \ddot{\mathcal{O}} \in \Omega$ such that $\dot{\mathcal{O}}$ and $\ddot{\mathcal{O}}$ differ in the outcome of a single game, $\Psi_i^{\text{best}}(\ddot{\mathcal{O}}^W) \leq \Psi_i^{\text{worst}}(\dot{\mathcal{O}}^L)$. Scenarios $\dot{\mathcal{O}}$ and $\ddot{\mathcal{O}}$ have the same set of outcomes for all games except $\bar{t}$. Similarly to Lemma 4, we rearrange the order of games t through t' so that game $\bar{t}$ is the last game to be played and game t is the penultimate game. Let w_k denote the number of wins that team k has before games t and $\bar{t}$ have been played, in this rearranged order. For any $q \in \mathbb{R}$, denote the set of teams with exactly q more wins than team i by $\mathcal{W}^q := \{k \in [n] : w_k = w_i + q\}$. Let $\mathcal{W}^{\geq q}$ denote the teams that have at least q more wins than team i.

Let team j be team i's opponent in game t, and suppose game $\bar{t}$ is played between teams u and v (possibly $i \in \{u, v\}$). Without loss of generality, assume that team u wins game $\bar{t}$ in scenario $\dot{\mathcal{O}}$ and v wins it in $\ddot{\mathcal{O}}$. Let $\mathbb{1}[\cdot]$ denote the indicator function (evaluating to 1 if the argument is true, and 0 otherwise).

Case 1: $u, v \neq i$**.** First, consider scenario $\dot{\mathcal{O}}^L$, when team i loses game t and team u wins game $\bar{t}$. The number of teams that have at least w_i wins determines the worst-case rank for team i. This can include teams j and u if those belong to $\mathcal{W}^{-1}$. Therefore,

$$\Psi_i^{\text{worst}}(\dot{\mathcal{O}}^L) = |\mathcal{W}^{\geq 0}| + |\{j, u\} \cap \mathcal{W}^{-1}| \geq |\mathcal{W}^{\geq 0}|.$$

Next, consider scenario $\ddot{\mathcal{O}}^W$, when team i wins game t and team v wins game $\bar{t}$. In this case, the best attainable position for team i is one more than the number of teams that end with at least $w_i + 2$ wins. Therefore,

$$\Psi_i^{\text{best}}(\ddot{\mathcal{O}}^W) = |\mathcal{W}^{\geq 2}| + \mathbb{1}[v \in \mathcal{W}^1] + 1.$$

We now have

$$\begin{aligned}
\Psi_i^{\text{worst}}(\dot{\mathcal{O}}^L) &\geq |\mathcal{W}^{\geq 0}| \\
&= |\mathcal{W}^{\geq 2}| + |\mathcal{W}^1| + |\mathcal{W}^0| \\
&\geq |\mathcal{W}^{\geq 2}| + \mathbb{1}[v \in \mathcal{W}^1] + 1 \\
&= \Psi_i^{\text{best}}(\ddot{\mathcal{O}}^W),
\end{aligned}$$

as required. The (second) inequality uses that $\mathcal{W}^1$ is of cardinality at least 1 if $v \in \mathcal{W}^1$ and at least 0 otherwise, and $|\mathcal{W}^0| \geq 1$ because $i \in \mathcal{W}^0$.

Case 2: $v = i$**.** The proof is similar to Case 1, only now, in scenario $\ddot{\mathcal{O}}^W$, team i has two additional wins from games t and $\bar{t}$:

$$\Psi_i^{\text{best}}(\ddot{\mathcal{O}}^W) = |\mathcal{W}^{\geq 3}| + 1.$$

We get that $\Psi_i^{\text{worst}}(\dot{\mathcal{O}}^L) \geq |\mathcal{W}^{\geq 0}| \geq |\mathcal{W}^{\geq 3}| + 1 = \Psi_i^{\text{best}}(\ddot{\mathcal{O}}^W)$, as desired.

Case 3: $u = i$**.** In this case, in scenario $\ddot{\mathcal{O}}^W$, team i beats team j in game t and loses to team v in game $\bar{t}$. The characterization of $\Psi_i^{\text{best}}(\ddot{\mathcal{O}}^W)$ is therefore identical to Case 1.

Under scenario $\dot{\mathcal{O}}^L$, team j wins game t and team i wins game $\bar{t}$, ending the scenario with $w_i + 1$ wins. The worst-case ranking $\Psi_i^{\text{worst}}(\dot{\mathcal{O}}^L)$ is determined by the number of teams that end with at least $w_i + 1$ wins.

$$\Psi_i^{\text{worst}}(\dot{\mathcal{O}}^L) = |\mathcal{W}^{\geq 1}| + \mathbb{1}[j \in \mathcal{W}^0] + 1.$$

Then

$$\Psi_i^{\text{worst}}(\dot{\mathcal{O}}^L) \geq |\mathcal{W}^{\geq 1}| + 1 = |\mathcal{W}^{\geq 2}| + |\mathcal{W}^1| + 1 \geq |\mathcal{W}^{\geq 2}| + \mathbb{1}[v \in \mathcal{W}^1] + 1 = \Psi_i^{\text{best}}(\ddot{\mathcal{O}}^W),$$

completing the proof. □

The following are some simple results that will be useful in the lemmas used to prove Theorem 2.

Lemma 7. *If* $\Pi_i^{\text{best}} = \Pi_i^{\text{worst}}$*, exerting high effort is dominant for team* i*.*

Proof. In this case, team i's position in the bilevel ranking is fixed no matter the scenario for the rest of the season; therefore, by Definition 1, exerting high effort is dominant. $\square$

Lemma 8. *If $\mathcal{O}$ crosses n^* with respect to team i, then $u_i(\mathcal{O}^W) - u_i(\mathcal{O}^L) \geq (1-\epsilon_i)u_i(n^*)/n$.*

Proof. From the definition of crossing, $\Psi_i^{\text{best}}(\mathcal{O}^W) \leq n^* < \Psi_i^{\text{worst}}(\mathcal{O}^L)$. Exactly one of the following is true: (i) $\Psi_i^{\text{best}}(\mathcal{O}^L) \leq n^*$, or (ii) $\Psi_i^{\text{best}}(\mathcal{O}^L) > n^*$.

Suppose first that (i) is true. Consider scenario $\mathcal{O}^L$. Let $\bar{\mathbf{p}}$ denote the probability that team i makes the playoffs in this scenario. Since $\Psi_i^{\text{worst}}(\mathcal{O}^L) > n^*$ and $n^* \leq n-1$, it holds that $\bar{\mathbf{p}} \leq (n-1)/n$. Therefore, when team i makes the playoffs, its utility is at most $u_i(\Psi_i^{\text{best}}(\mathcal{O}^L))$; when team i does not make the playoffs, its utility is at most $u_i(n)$. Hence,

$$u_i(\mathcal{O}^L) \leq \bar{\mathbf{p}} \cdot u_i(\Psi_i^{\text{best}}(\mathcal{O}^L)) + (1-\bar{\mathbf{p}}) \cdot u_i(n).$$

For scenario $\mathcal{O}^W$, using Lemma 4, we have $\Psi_i^{\text{best}}(\mathcal{O}^W) \leq \Psi_i^{\text{worst}}(\mathcal{O}^W) \leq \Psi_i^{\text{best}}(\mathcal{O}^L) \leq n^*$ and

$$u_i(\mathcal{O}^W) \geq u_i(\Psi_i^{\text{worst}}(\mathcal{O}^W)) \geq u_i(\Psi_i^{\text{best}}(\mathcal{O}^L)).$$

From the definition of ϵ_i, $u_i(n) = \epsilon_i \cdot u_i(n^*) \leq \epsilon_i \cdot u_i(\Psi_i^{\text{best}}(\mathcal{O}^L))$.

$$\begin{aligned} u_i(\mathcal{O}^W) - u_i(\mathcal{O}^L) &\geq u_i(\Psi_i^{\text{best}}(\mathcal{O}^L)) - \bar{\mathbf{p}} \cdot u_i(\Psi_i^{\text{best}}(\mathcal{O}^L)) - (1-\bar{\mathbf{p}}) \cdot u_i(n) \\ &= (1-\bar{\mathbf{p}}) \cdot (u_i(\Psi_i^{\text{best}}(\mathcal{O}^L)) - \epsilon_i \cdot u_i(n^*)) \\ &\geq \frac{1}{n} \cdot (u_i(n^*) - \epsilon_i \cdot u_i(n^*)) = \frac{(1-\epsilon_i)}{n} \cdot u_i(n^*). \end{aligned}$$

The proof for (ii) is analogous and omitted. $\square$

4.4 Proof of Theorem 2

We are now ready to prove the various cases of Theorem 2.

Lemma 9. *For all $t < \tau_i$, exerting high effort is dominant for team i.*

Proof. By Claim 5, every scenario $\mathcal{O} \in \Omega^T$ falls into exactly one of three categories: (i) team i is a contender under $\mathcal{O}$, i.e., $\mathcal{O}$ crosses n^*, (ii) team i is sure to make the playoffs, i.e., $\Psi_i^{\text{worst}}(\mathcal{O}^L) \leq n^*$, and (iii) team i is sure not to make the playoffs, i.e., $n^* < \Psi_i^{\text{best}}(\mathcal{O}^W)$.

For a type (i) scenario $\mathcal{O}$ that crosses n^*, $u_i(\mathcal{O}^W) - u_i(\mathcal{O}^L) \geq (1-\epsilon_i)u_i(n^*)/n$ by Lemma 8. For every scenario $\mathcal{O}$ of type (ii), when team i is assured to make the playoffs, Lemma 4 implies that $u_i(\mathcal{O}^W) \geq u_i(\mathcal{O}^L)$. Next, consider a scenario $\mathcal{O}$ of type (iii), when team i is eliminated. By Lemma 4 (and indeed, intuitively), team i may have incentive to exert low effort in this scenario. However, we can upper bound the extra utility from losing as follows:

$$u_i(\mathcal{O}^L) - u_i(\mathcal{O}^W) \leq u_i(n) - u_i(n^*+1) = \epsilon_i \cdot u_i(n^*) - u_i(n^*+1).$$

If there are no scenarios of type (iii), then clearly $U_i^{\mathbf{p}}(W_i) \geq U_i^{\mathbf{p}}(L_i)$ and exerting high effort is dominant for team i. Thus, assume that scenarios of type (iii) exist. By Lemma 6 and the fact that $t < \tau_i$, a scenario $\bar{\mathcal{O}}$ of type (i) exists. By 3, the probability that team i assigns to scenario $\bar{\mathcal{O}}$ satisfies $\mathbf{p}^i(\bar{\mathcal{O}}) \geq p_i^* \geq n\epsilon_i/(1-\epsilon_i)$. Hence,

$$\begin{aligned}
U_i^{\mathbf{p}}(W_i) - U_i^{\mathbf{p}}(L_i) &= \sum_{\mathcal{O}\in\Omega^T} \mathbf{p}^i(\mathcal{O}) \cdot \big(u_i(\mathcal{O}^W) - u_i(\mathcal{O}^L)\big) \\
&\geq p_i^* \cdot \big(u_i(\bar{\mathcal{O}}^W) - u_i(\bar{\mathcal{O}}^L)\big) - (1-p_i^*) \cdot (\epsilon_i \cdot u_i(n^*) - u_i(n^*+1)) \\
&\geq \frac{n\epsilon_i}{1-\epsilon_i} \cdot \frac{(1-\epsilon_i)\cdot u_i(n^*)}{n} - (\epsilon_i \cdot u_i(n^*) - u_i(n^*+1)) \\
&= u_i(n^*+1) > 0.
\end{aligned}$$

Therefore, team i's dominant strategy is to exert high effort in game t. □

The next lemma implies that if team i is eliminated by the breakpoint δ, then every team with win-rank worse than team i after game δ is also eliminated. This lemma is proved in [AEHO02]; we include a proof for completeness.

Lemma 10. *If $\tau_i \leq \delta$, then, for every team j with at most as many wins as team i after game δ, it holds that $\tau_j \leq \delta$.*

Proof. All teams are assumed to have the same number of games, say $\dot{T}$, left to play after δ. Let w_i denote the number of wins team i has after game δ. Suppose, for the sake of contradiction, that some team j has at most w_i wins after game δ, but makes the playoffs in scenario $\mathcal{O}$. Without loss of generality, we can assume that team j wins all of its remaining $\dot{T}$ games, ending the season with at most $w_i + \dot{T}$ wins under scenario $\mathcal{O}$, which we have now assumed is enough wins to make the playoffs. Let $\mathcal{O}'$ be identical to scenario $\mathcal{O}$, except that team i wins all of its remaining games in $\mathcal{O}'$. Relative to scenario $\mathcal{O}$, under scenario $\mathcal{O}'$, all teams except team i have the same or fewer wins at the end of the season, and team i has $w_i + \dot{T}$ wins. It follows that team i will be at least tied for the last playoff position, a contradiction to $\tau_i \leq \delta$. □

As a corollary, when $\tau_i \leq \delta$, team i's bilevel-rank is determined at the end of game δ.

Corollary 11. *If $\tau_i \leq \delta < t$, then $\Pi_i^{\text{best}} = \Pi_i^{\text{worst}}$.*

Lemma 12. *For $\tau_i \leq \delta < t$, exerting high effort is dominant for team i.*

Proof. Direct from Corollary 11 and Lemma 7. □

Lemma 13. *If $\tau_i \leq t \leq \delta$ and $\Pi_i^{\text{best}} \neq \Pi_i^{\text{worst}}$, exerting low effort is dominant for team i.*

Proof. For any $\mathcal{O} \in \Omega^{t+1,\delta}$, Lemma 4 implies that $\Psi_i^{\text{worst}}(\mathcal{O}^W) \leq \Psi_i^{\text{best}}(\mathcal{O}^L)$. By Corollary 11, team i's bilevel-rank is finalized at game δ. Hence, for every scenario, team i's bilevel-rank after winning game t is at most its bilevel-rank after losing game t. Therefore, by 3, $u_i(\mathcal{O}^W) \leq u_i(\mathcal{O}^L)$. Using Definition 1, it is insufficient to argue that losing weakly improves team i's utility. We need to show that there exists a scenario in which losing game t *strictly* improves team i's expected utility.

Let $r := \Pi_i^{\text{best}}$. We prove that when $\Pi_i^{\text{best}} \neq \Pi_i^{\text{worst}}$, there exists some $\bar{\mathcal{O}} \in \Omega^{t+1,\delta}$ such that $\bar{\mathcal{O}}$ crosses r. The proof follows from Claim 5 and Lemma 6, from which we know that it cannot be that in all scenarios, either $\Psi_i^{\text{worst}}(\mathcal{O}^L) \le r$ or $r < \Psi_i^{\text{best}}(\mathcal{O}^W)$, so there exists the desired crossing scenario $\bar{\mathcal{O}}$, in which $\Psi_i^{\text{best}}(\bar{\mathcal{O}}^W) \le r < \Psi_i^{\text{worst}}(\bar{\mathcal{O}}^L)$. Hence, for this scenario $\bar{\mathcal{O}}$, team i's utility is strictly greater if it loses game t, completing the proof. □

The last case is when $\delta < \tau_i \le t$.

Lemma 14. *Let $\delta < \tau_i \le t$, such that in game t, team i plays against against team j. The dominant strategy for team i is to exert low effort in game t if and only if the following conditions hold:*

(1) $\pi_i^\delta < \pi_j^\delta$, i.e., team j's win-rank was worse than team i's at game δ, and

(2) there exists a team k with $\pi_k^\delta < \pi_i^\delta$ and a scenario $\bar{\mathcal{O}} \in \Omega^T$ such that teams j and k are both contenders under scenario $\bar{\mathcal{O}}$.

Proof. Let $\mathcal{I}$ denote the set of teams whose win-rank is worse than team i's at game δ, i.e., $\mathcal{I} := \{\ell \in [n] : \pi_i^\delta < \pi_\ell^\delta\}$. Team i will have bilevel-rank $\pi_i^{\text{LEAG}} = r$ if and only if $r - \pi_i^\delta$ teams from $\mathcal{I}$ make the playoffs. Let $\mathcal{C}(\mathcal{O})$ denote the set of teams that are contenders under a scenario $\mathcal{O} \in \Omega^T$, with game t undecided (this is equivalent to rearranging the season so that game t is last, and all games but t have been played). Recall that $u_i(\mathcal{O}) = \sum_{r\in\Pi_i(\mathcal{O})} \mathbf{Pr}[\pi_i^{\text{LEAG}} = r \mid \mathcal{O}] \cdot u_i(r)$. Observe that, under scenario $\mathcal{O}$, if $\bar{n}$ of the teams from $\mathcal{I}$ are sure to be eliminated, then $\Pi_i(\mathcal{O}) = \{\pi_i^\delta + \bar{n}, \ldots, \pi_i^\delta + \bar{n} + |\mathcal{C}(\mathcal{O}) \cap \mathcal{I}|\}$, and the utility for team i increases as more teams from $\mathcal{C}(\mathcal{O}) \cap \mathcal{I}$ make the playoffs. Our analysis involves comparing $\mathbf{Pr}[\pi_i^{\text{LEAG}} = r \mid \mathcal{O}^W]$ with $\mathbf{Pr}[\pi_i^{\text{LEAG}} = r \mid \mathcal{O}^L]$ for all $r \in \Pi_i(\mathcal{O})$.

Only if: We assume that team i's dominant strategy is to exert low effort in game t, and show that both conditions in the theorem then necessarily hold.

We start with condition (2), and show that $u_i(\mathcal{O}^W) = u_i(\mathcal{O}^L)$ for every $\mathcal{O}$ that does not satisfy condition (2). If $j \notin \mathcal{C}(\mathcal{O})$, then either team j is sure to make the playoffs or guaranteed to be eliminated, so the set of contenders will not change after game t is decided, which implies that $u_i(\mathcal{O}^W) = u_i(\mathcal{O}^L)$. Similarly, if $\mathcal{C}(\mathcal{O}) \cap \mathcal{I} = \emptyset$ or $\mathcal{C}(\mathcal{O}) \cap \mathcal{I} = \mathcal{C}(\mathcal{O})$, i.e., either none or all of the contenders were worse-ranked than team i after game δ, then, for all ranks r, $\mathbf{Pr}[\pi_i^{\text{LEAG}} = r \mid \mathcal{O}^W] = \mathbf{Pr}[\pi_i^{\text{LEAG}} = r \mid \mathcal{O}^L]$, so $u_i(\mathcal{O}^W) = u_i(\mathcal{O}^L)$. Thus, if exerting low effort is dominant for team i, then there exists a scenario $\bar{\mathcal{O}} \in \Omega^T$ in which $j \in \mathcal{C}(\mathcal{O})$ and $\emptyset \neq \mathcal{C}(\mathcal{O}) \cap \mathcal{I} \subsetneq \mathcal{C}(\mathcal{O})$. This implies that there necessarily exists $k \in \mathcal{C}(\mathcal{O}) \setminus \mathcal{I}$, a team that was better-ranked than team i at game δ and that is also a contender under scenario $\bar{\mathcal{O}}$, completing the proof of the necessity of condition (2).

Next, consider any scenario $\mathcal{O}$ satisfying condition (2), and suppose $j \notin \mathcal{I}$ for the sake of contradiction. Observe that the teams in $\mathcal{C}(\mathcal{O}) \setminus \{j\}$ all have the same number of wins, as the undecided game t can only change team i or j's win total. For any $\ell \in [n]$, let w_ℓ equal the number of wins team ℓ has in $\mathcal{O}$ (with game t undecided). As $j \in \mathcal{C}(\mathcal{O})$, it follows that strictly fewer than n^* teams have $w_j + 2$ or more wins, as otherwise, team j cannot make the playoffs even after winning game t. Hence, the teams with $w_j + 2$ or more wins are sure to make the playoffs, so they are not contenders. Similarly, strictly more than n^* teams have w_j or more wins, as otherwise team j is sure to make the playoffs; thus, all teams with at most $w_j - 1$ wins are eliminated. This leaves two cases to consider: (i) $w_k = w_j$ for all $k \in \mathcal{C}(\mathcal{O})$, or (ii) $w_k = w_j + 1$ for all $k \in \mathcal{C}(\mathcal{O}) \setminus \{j\}$.

Denote the number of remaining (undecided) playoff positions after fixing the outcomes of scenario $\mathcal{O}$ by q, to be allocated among the teams in $\mathcal{C}(\mathcal{O})$. We show that, for all $0 < m \leq |\mathcal{C}(\mathcal{O}) \cap \mathcal{I}|$, the probability that exactly m teams from $\mathcal{C}(\mathcal{O}) \cap \mathcal{I}$ make the playoffs is strictly higher when team i wins game t than when it loses, implying that $u_i(\mathcal{O}^W) > u_i(\mathcal{O}^L)$.

In case (i), if team i loses game t, team j will have $w_j + 1$ wins and is assured to make the playoffs, reducing the probability that $m > 0$ teams from $\mathcal{C}(\mathcal{O}) \cap \mathcal{I}$ make the playoffs within the remaining $q - 1$ playoff positions. In case (ii), if team i wins game t, then team j is eliminated; if team i loses game t, then team j may make the playoffs, if it is chosen among the teams with $w_j + 1$ wins. As $j \notin \mathcal{I}$, the probability that $m > 0$ teams from $\mathcal{C}(\mathcal{O}) \cap \mathcal{I}$ make the playoffs is strictly higher when team i wins game t. In either case, $u_i(\mathcal{O}^W) > u_i(\mathcal{O}^L)$, as required.

If: We show that when conditions (1) and (2) hold, the expected utility from winning is strictly lower than the expected utility of losing game t. It suffices to show that $u_i(\mathcal{O}^W) \leq u_i(\mathcal{O}^L)$ for every scenario $\mathcal{O} \in \Omega^T$, and this holds strictly for at least one scenario.

Given conditions (1) and (2) hold, we can partition scenarios into those in which $j \notin \mathcal{C}(\mathcal{O})$ or $\mathcal{C}(\mathcal{O}) \subseteq \mathcal{I}$, or those in which condition (2) is satisfied. For the former, we already proved that $u_i(\mathcal{O}^W) = u_i(\mathcal{O}^L)$.

Thus, consider any scenario that satisfies condition (2), i.e., $j \in \mathcal{C}(\mathcal{O}) \cap \mathcal{I}$ and $\mathcal{C}(\mathcal{O}) \setminus \mathcal{I} \neq \emptyset$. For such a scenario, similarly to the proof of the **only if** case, for all $0 < m \leq |\mathcal{C}(\mathcal{O}) \cap \mathcal{I}|$, the probability that exactly m teams from $\mathcal{C}(\mathcal{O}) \cap \mathcal{I}$ make the playoffs is strictly higher when team i loses game t compared to when it wins, so $u_i(\mathcal{O}^W) < u_i(\mathcal{O}^L)$. Hence, if conditions (1) and (2) hold, then for any possible $\mathbf{p}$, $U_i^{\mathbf{p}}(W_i) < U_i^{\mathbf{p}}(L_i)$. □

5 Computational Results

We present our computational evaluation of the bilevel ranking mechanism we proposed in Section 3, using the theoretical model of a sports league from Section 2. Our experiments complement Section 4's theoretical characterization of dominant team strategies under the bilevel ranking. By Theorem 2, teams have less incentive to tank when the breakpoint game δ is set earlier in the season. Implicit in this theorem is the effect of the breakpoint δ on competitive balance. Intuitively, when δ is set earlier, the league receives a worse signal of each team's strength due to fewer games being played at the the time the draft ranking is determined, and the bilevel ranking would therefore be expected to be less accurate. The less accurate the league's ranking, the worse the league is at assigning draft positions in the correct order, which in turn negatively impacts long-term competitive balance. The question is then: *what is an appropriate value for δ to trade off the league's two objectives: to improve long-term competitive balance and to reduce tanking?*

Our simulations show that setting δ to be somewhere around 5/6ths to 7/8ths through the season gives a reasonable balance of these two objectives. In particular, by setting δ like this, we find that:

- The number of tanked games decreases by 50–70%.

- When no teams are selfish, the accuracy of the bilevel ranking is only slightly worse than the accuracy of the ranking at the end of the season.
- The bilevel ranking can actually be strictly *better* than the end-of-season ranking when some teams are selfish.

The last of these conclusions is perhaps most relevant to sports leagues, as it seems to reflect the empirical reality that some, but not all, teams will engage in tanking [Abb12, Dee13]. Specifically, when about 1/4 to 3/4 of the teams are selfish, the bilevel ranking turns out to be strictly better both in reducing tanking *and* in improving competitive balance. Such a result may seem counterintuitive, but it is explained by the fact that tanking adds noise to rankings, and the effect is particularly pronounced close to the end of the season, because this is when most teams are eliminated from the playoffs.

5.1 Computational setup

We simulate a one-division sports league using the model defined in Section 2. Our stylized league has $n = 30$ teams, of which $n^* = 16$ teams make the playoffs, and every pair of teams plays each other 3 times, which results in each team playing 87 games in the simulated (regular) season, a total of $T = 1305$ games.[4] We fix the true ranking of the 30 teams, labeled $1, \ldots, 30$, as $\pi_1^{\text{TRUE}} < \cdots < \pi_{30}^{\text{TRUE}}$.

In our experiments, one parameter is the number of selfish teams, which we test with values from 0 to 30. The other primary parameter that we vary is δ, for which we use values $\{(1/2)T, (2/3)T, (3/4)T, (5/6)T, (7/8)T, T\}$ (rounded to the nearest integer).

In theory, each data point should be a season in which δ is set to a value and the number of selfish teams is set to some value. Then we simulate that data point 100K times. However, there are $6 \times 31 = 186$ of these possibilities, and it would mean that experiments that take 1 week would instead take a month in half. In our results, we assume selfish teams don't stop tanking after δ, which may mess things up, in that different teams might make the playoffs.

As opposed to the optimistic model that we use for the theoretical results, in our simulations, we use the conservative decision-making model described in Section 2, i.e., that selfish teams will tank once they have been effectively eliminated from the playoffs, as it is more realistic. A team will typically be effectively eliminated before it is mathematically eliminated, so there is a chance that an effectively eliminated team eventually makes the playoffs. In our simulations, when no teams are selfish, this happens on average for one team every 3 to 4 seasons, decreasing to one team every 8 seasons when all teams are selfish. In addition to being a more realistic tanking criterion, effective elimination has the advantage that it is easy to check, whereas mathematical elimination is $\mathcal{NP}$-Hard to verify [McC99].

Nevertheless, we do calculate mathematical elimination to compute the ranking proposed by Lenten [Len16]. To decide mathematical elimination, after each game and for each team, we solve the mixed-integer program (MIP) discussed in Appendix B.1, which we implement using the JuMP framework [DHL17, LDGL20]. As solving a MIP is computationally demanding, we design several heuristics, described in the appendix, which reduce the number of MIPs solved to around 40–45 per replication of a simulated season (compared to a naïve

[4] These parameters are motivated by the NBA, in which there are 30 teams and each team plays 82 games.

implementation, which would involve nT MIPs). Nonetheless, performing this step causes our computational experiments to be slower by about an order of magnitude compared to running the simulations without any calculation of mathematical elimination.

To determine the winner of each game, we must first decide how to set $p_{ij}(e_i, e_j)$, the probability that team i beats team j, for all $i, j \in [n]$ and effort levels $e_i, e_j \in \{L, H\}$. We first make a simplifying assumption, due to which we only need to set $p_{ij}(H, H)$:

$$0 = p_{ij}(L, H) < p_{ij}(L, L) = p_{ij}(H, H) < p_{ij}(H, L) = 1.$$

That is, a team that exerts low effort will always lose to a team that exerts high effort, and $p_{ij}(H, H) = p_{ij}(L, L)$ means that if both teams exert low effort, the win probability is the same as when they both exert high effort. Without these assumptions, e.g., if we consider the realistic case that a team's ability to tank is not related to its true strength, the outcomes of the games would be noisier, but we would expect to see qualitatively similar results.

As discussed in Section 2, we consider two models for determining the probability $p_{ij}(H, H)$. The first is the (Zermelo)-Bradley-Terry model [Zer29, BT52, Luc59]. The second is the *noisy comparison* model; see, e.g., [BM08, RV17]. Our evaluation of these alternative models is discussed in Appendix A, calibrated using data from the 2004–2018 NBA seasons,[5] excluding the strike-shortened 2011 season. This leads us to adopt the noisy comparison model with $\gamma = 0.71425$.

To give a sense of how closely our simulation reflects real NBA data, we plot in Figure 2 the average win percentage of teams based on rank, assuming that half the teams are selfish. The black solid line gives the win percentage for each rank, averaged over the 2004–2018 end-of-season NBA data, while the vertical black lines give the range of these historical values. Both the Bradley-Terry and noisy comparison models have roughly the same distribution of winning percentages as the historical NBA data, though the noisy comparison model clearly is more accurate at the tails.

Unless otherwise stated, the results are the average of 10,000 replications for each data point. All our code is implemented in Julia 1.3.1 and open source.[6]

5.2 Evaluation of the bilevel ranking

To evaluate the bilevel ranking in context of other options, we calculate several alternative league rankings. The three systems we test are the draft lottery, the "Lenten" system based on mathematical elimination ordering, and our bilevel ranking.

NBA draft lottery. The draft lottery is the system currently in place in the NBA, used to select the first 4 positions in the draft, while the remaining teams pick in reverse order of their end-of-season standings. The lottery involves picking at random from the non-playoff teams, where teams ranked worse at the end of the regular season receive higher odds of being picked. Teams 1–3 are given a 14% chance of being picked, and the odds steadily decrease for the subsequent positions.[7] After a team is selected, the odds for the remaining teams are renormalized.

[5] Data is obtained from `http://www.basketball-reference.com`.

[6] The code is publicly available at `https://github.com/akazachk/tanking`.

[7] See `https://en.wikipedia.org/wiki/NBA_draft_lottery` for the values used for the draft lottery.

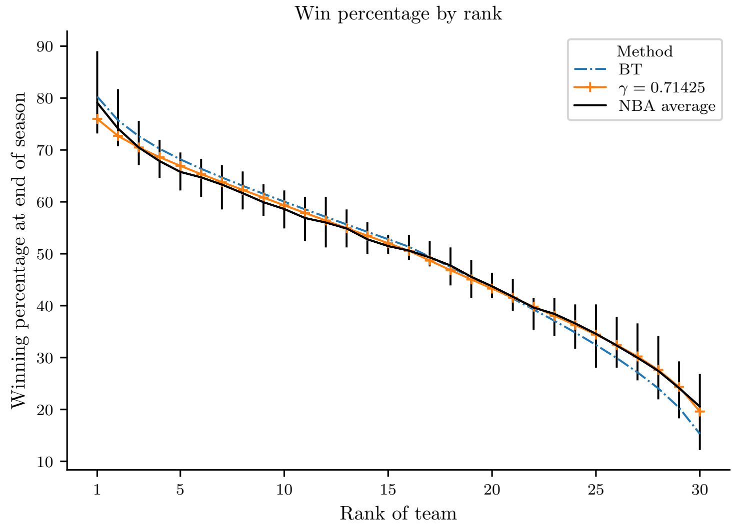

Figure 2: Win percentage of teams achieving each rank, compared across the Bradley-Terry model, noisy comparison model, and historical NBA data. The vertical lines show the minimum / maximum win percentages for each rank across the 2004–2018 NBA seasons. Each data point is the average of 100,000 replications.

The draft lottery was recently changed, in 2019, so we also compare the current system to the previous one. In the previous draft lottery, only the first 3 positions of the draft were selected by the lottery, and the odds were different, e.g., the last-place team at the end of the regular season had a 25% chance of being picked.

Though the lottery was introduced and continues to be modified in order to reduce tanking incentives created by the draft, we know of no theoretical treatment on whether tanking is actually likely to be mitigated by the lottery.

Mathematical elimination ordering. Lenten [Len16] proposes to rank teams in the order that they are mathematically eliminated. The advantage of this system is that it is dynamic and teams have no incentive to tank after being mathematically eliminated. However, teams are likely to start tanking before this point, e.g., based on effective elimination. In addition, mathematical elimination is difficult to compute and therefore to convey to fans and other stakeholders.

Bilevel ranking. The bilevel ranking requires setting the parameter δ. We assume δ is fixed and publicly announced before the season. Extensions not analyzed in this section include a randomized or dynamically set δ, as well as the effect of combining the draft lottery and the bilevel ranking.

5.3 Simulation results

Effect of δ on tanking. Our simulations provide a sense of how many tanked games can be prevented by using the bilevel ranking. Figure 3 plots the number of games that teams tank in our simulations as a function of the number of selfish teams and the breakpoint

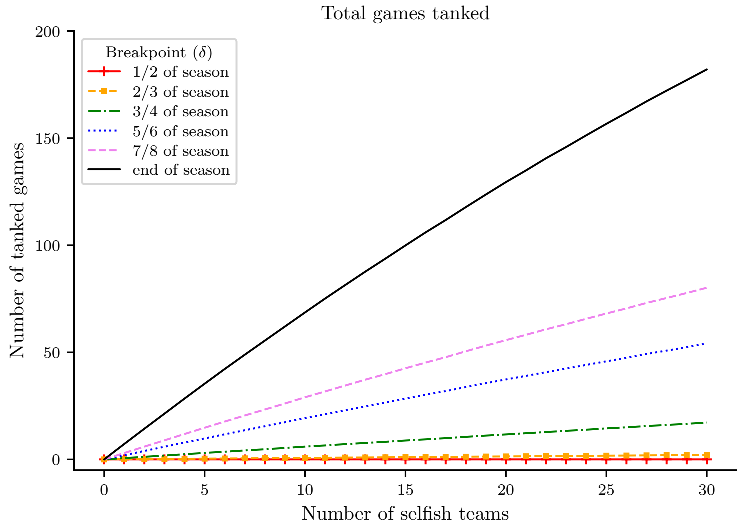

Figure 3: The average number of games tanked in a season for different breakpoints.

δ. For example, if half of the teams are selfish (the relative reductions remain similar for other values), then when $\delta = (5/6)T$, on average around 29 games are tanked, and when $\delta = (7/8)T$, on average around 43 games are tanked. Compare this to $\delta = T$, when on average around 100 games are tanked.

We remark that the results in Figure 3 assume that teams will not tank after game δ. This is not the case, as we proved in Lemma 14, but preliminary experiments showed that the conditions in Lemma 14 are only infrequently satisfied, occurring on average for 4–5 games per season. The effect of our simplifying assumption is therefore almost negligible, and the extra computational effort required to frequently check whether a team is a contender is not warranted.

Effect of δ on ranking accuracy. Recall that the accuracy of the bilevel ranking is our proxy for competitive balance: if the league uses a ranking that is "closer" to the true ranking, then the draft will be more effective at allocating better players to worse teams. Figure 4 shows the Kendall tau distance, as defined in Equation (1), between the true and bilevel rankings of the non-playoff teams, plotted for several possible breakpoints, as well as for the Lenten ranking based on mathematical elimination ordering, and the ranking obtained after randomizing the order of teams based on the draft lottery (using the pre-2019 and post-2019 odds). On the horizontal axis, we vary the number of selfish teams. 14 teams do not make the playoffs, hence the maximum Kendall tau distance from their true ranking is $\binom{14}{2} = 91$. Note that when almost all teams are tanking, the ranking accuracy is similar to when no teams are tanking—this is because of our assumption that $p_{ij}(L, L) = p_{ij}(H, H)$.

We see in Figure 4 that when $\delta = T$ and there are no selfish teams, on average around 17 team pairs can be expected to be incorrectly relatively ranked. This is due to the noisiness of the outcome of each game and the small number of times each pair of teams plays. When nearly no teams or all teams are tanking, and setting $\delta = (5/6)T$ or $(7/8)T$, only 19 or 20 pairs of teams would be incorrectly ordered, which compares favorably to the accuracy of the ranking when $\delta = T$. As a result, we conclude the bilevel mechanism would yield, in the

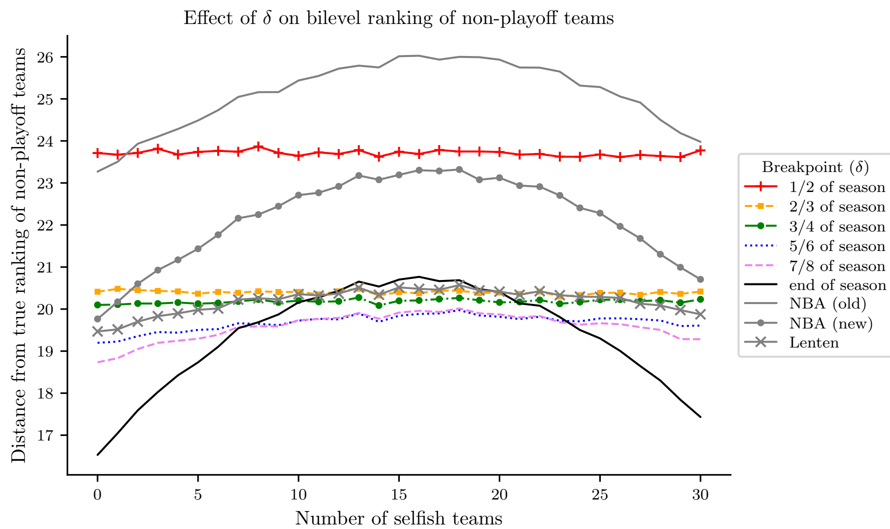

Figure 4: The average Kendall tau distance of the bilevel ranking from the true ranking as a function of the proportion of selfish teams.

worst case, great reductions in tanking incentives with only small losses in ranking accuracy.

When some (between 8 and 23) teams are selfish, we can make an even more striking conclusion: in this case, the accuracy of the ranking of the non-playoff teams at $\delta = T$ can actually be *worse* than their ranking at an earlier breakpoint. This may seem counterintuitive, but a moment's reflection explains the phenomenon: most teams are eliminated close to the end of the season (see, e.g., Figure 7 in the appendix), so this is when most tanking occurs, and the tanked games add significant noise to the ranking. The upshot is that, even if the league only cared about promoting competitive balance and placed zero emphasis on reducing tanking, it would still benefit from using the bilevel ranking with $\delta < T$.

Comparison to Lenten ranking and draft lottery. Figure 4 also indicates that the bilevel ranking is better in terms of competitive balance than the current draft lottery system, which itself is better than the previous lottery odds. Using the new lottery, the distance of the current NBA draft order to the true ranking is always worse, on 1 to 3 pairs of teams, than when using $\delta \in \{(5/6)T, (7/8)T\}$. Moreover, while we are able to quantify how the bilevel ranking affects team tanking strategies, we know of no conclusive theoretical analysis of the effect of the draft lottery on tanking, despite the clear loss in ranking accuracy. Comparing to the Lenten ranking, we see that the bilevel ranking with $\delta \in \{(5/6)T, (7/8)T\}$ may slightly improve ranking accuracy in this case as well.

Effect of tanking on rank. Lastly, through our simulations, we can also analyze how effective tanking is at improving a team's standing in the draft, which is a way of understanding how competitive balance is hurt by tanking. In Figure 5, we plot the average rank of teams that do not make the playoffs, where we vary the number of selfish teams on the horizontal axis. The plot only averages those simulated seasons in which there is at least one moral and one selfish non-playoff team; e.g., if there is one selfish team but it makes the playoffs, then that team never tanked and we get no signal of the effect of tanking on rank.

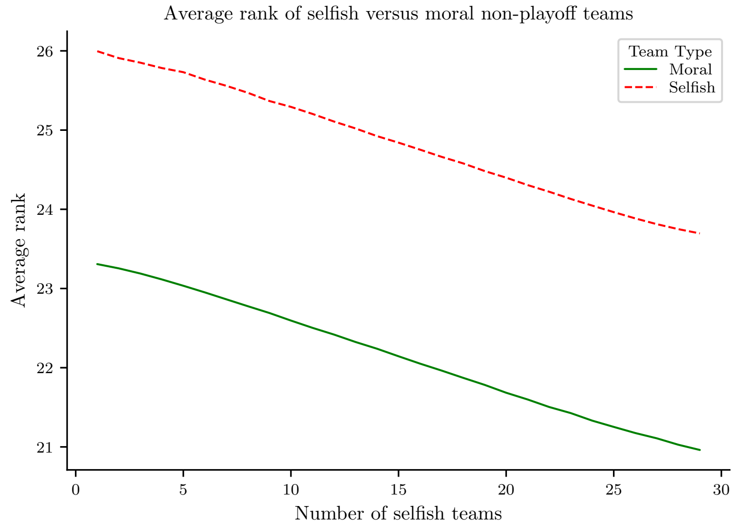

Figure 5: Average win-rank in π^T of non-playoff teams.

We see from the figure that, on average, if teams' picks in the draft are related to the reverse of their end-of-season ranking, as is the case in the current system, then a selfish team may improve its draft ranking by 2.5 positions compared to a moral team, which underlines how significantly tanking can distort ranking accuracy and thereby inhibit long-term competitive balance.

6 Conclusion

In this paper, we introduce a flexible model of a sports league to capture the effect of draft reforms on team incentives to tank games, while also accounting for the impact of these reforms on competitive balance in the league. We introduce a novel bilevel ranking scheme and characterize the dominant tanking strategies for teams under this mechanism. Our theoretical and simulation results demonstrate that the bilevel ranking can significantly diminish tanking incentives, while at the same time improving the ability of the league to rank teams accurately, corresponding to improvements in long-term competitive balance. In addition, the bilevel ranking is transparent and easily implementable.

Our model makes it possible to quantify the effect of draft reforms on tanking and competitive balance. Although our simulations are thorough, in our theoretical results, we do not analyze the draft lottery specifically. Throughout, we have also made several simplifying assumptions. Perhaps the most important is that there exists a true ranking that stays fixed throughout the season. In reality, team strength is dynamic and subject to injuries, trades, and quality of management or coaching. Furthermore, it is not necessarily true that teams can be totally ordered.

We also do not consider the effect of other forms of tanking, such as a team that decides to tank an entire season due to "betting on the draft" or teams that purposefully lose games at the end of the season in order to achieve a more favorable playoff matchup. A substantially more involved simulation would be required to adequately capture these phenomena.

Finally, it would be worthwhile to place our study in a larger context, considering more of the league's objectives and decisions. While the draft is one of the primary ways the league controls for competitive balance, there exist other methods, such as through limits on salaries, contracts, and trades. Indeed, there has been empirical work suggesting the draft may not necessarily be the best path to improving long-term success of teams [MRLL16].

Perhaps more importantly, it is reasonable to ask and analyze whether the league should even pursue competitive balance in the first place. Despite this uncertainty, it is clear that reducing tanking and having an accurate ranking of teams are worthwhile, in and of themselves, and the bilevel ranking we introduce is substantially more effective at these goals than the systems currently in place and others proposed in the literature.

References

[Abb12] Henry Abbott. Tanking is the tip of the iceberg. http://www.espn.com/blog/truehoop/post/_/id/39318/tanking-is-the-tip-of-the-iceberg, April 10, 2012. Retrieved December 7, 2018.

[AEHO02] Ilan Adler, Alan L. Erera, Dorit S. Hochbaum, and Eli V. Olinick. Baseball, optimization, and the world wide web. *Interfaces*, 32(2):12–22, 2002.

[Avi15] Jody Avirgan. How To Stop NBA Tanking: Tie Your Fate To Another Team's Record. https://fivethirtyeight.com/features/how-to-stop-nba-tanking-tie-your-fate-to-another-teams-record/, May 13, 2015. Retrieved December 9, 2018.

[Bar17] Bill Barnwell. A tanking guide to the NFL, and a warning. http://www.espn.com/nfl/story/_/id/19652764/tanking-nfl-advantages-disadvantages-worth-new-york-jets-cleveland-browns-2017, June 21, 2017. Retrieved December 8, 2018.

[Beg18] Ian Begley. Adam Silver: New tanking reforms may not be enough to address issue. http://www.espn.com/espn/print?id=23159490, April 13, 2018. Retrieved August 23, 2018.

[Bin13] Burhan Biner. Is parity good? externalities in professional sports. *Econ. Model.*, 30:715–720, 2013.

[BM08] Mark Braverman and Elchanan Mossel. Noisy sorting without resampling. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 268–276, 2008.

[BM20] Martino Banchio and Evan Munro. An incentive-compatible draft allocation policy. In *MIT Sloan Sports Analytics Conference*, March 2020.

[BT52] Ralph Allan Bradley and Milton E. Terry. Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika*, 39:324–345, 1952.

[Cas19] Jarod Castillo. Fixing the NBA draft lottery with tournament-style playoffs. https://sportsnaut.com/2019/05/nba-draft-lottery-replace-tournament-style-playoffs/, May 2019. Retrieved February 6, 2020.

[Cat12] Manuela Cattelan. Models for paired comparison data: A review with emphasis on dependent data. *Statist. Sci.*, 27(3):412–433, 08 2012.

[CDL11] Xi Chen, Xiaotie Deng, and Becky Jie Liu. On incentive compatible competitive selection protocols. *Algorithmica*, 61(2):447–462, 2011.

[Dee13] Mark Deeks. The right and wrong way to float in the NBA's middle class. https://www.sbnation.com/nba/2013/9/25/4770108/milwaukee-bucks-washington-wizards-nba-middle-class, September 2013. Retrieved December 8, 2018.

[DHL17] Iain Dunning, Joey Huchette, and Miles Lubin. JuMP: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320, 2017.

[DLB14] Michael De Lorenzo and Anthony Bedford. An alternative draft system for the allocation of player draft selections in the National Basketball Association. In Anthony Bedford and Tim Heazlewood, editors, *Proceedings of the 12th Australasian Conference on Mathematics and Computers in Sport*, pages 60–65. MathSport (ANZIAM), 2014.

[EHQ71] Mohamed El-Hodiri and James Quirk. An economic model of a professional sports league. *J. Polit. Econ.*, 79(6):1302–1319, 1971.

[FM03] Rodney Fort and Joel Maxcy. Competitive balance in sports leagues: An introduction. *J. Sport. Econ.*, 4(2):154–160, 2003.

[For18] Helena Fornwagner. Incentives to lose revisited: The NHL and its tournament incentives. *J. Econ. Psychol.*, 2018.

[FQ95] Rodney Fort and James Quirk. Cross-subsidization, incentives, and outcomes in professional team sports leagues. *J. Econ. Lit.*, 33(3):1265–1299, 1995.

[Fri12] John Friel. 5 reasons tanking must be eliminated in the NBA. https://bleacherreport.com/articles/1160643-5-reasons-tanking-must-be-eliminated-in-the-nba, April 25, 2012. Retrieved December 7, 2018.

[GM02] D. Gusfield and C. Martel. The structure and complexity of sports elimination numbers. *Algorithmica*, 32(1):73–86, 2002.

[Gol12] Adam M. Gold. How to cure tanking. In *MIT Sloan Sports Analytics Conference*, March 2012.

[Gur18] Gurobi Optimization, Inc. Gurobi Optimizer Reference Manual. http://www.gurobi.com, 2018. Version 8.1.0.

[HK12] Jason Hartline and Robert Kleinberg. Badminton and the science of rule making. `https://www.huffingtonpost.com/jason-hartline/badminton-and-the-science-of-rule-making_b_1773988.html`, August 13, 2012. Retrieved August 28, 2018.

[HR70] A. J. Hoffman and T. J. Rivlin. When is a team "mathematically" eliminated? In *Proceedings of the Princeton Symposium on Mathematical Programming (Princeton Univ., 1967)*, pages 391–401. Princeton Univ. Press, Princeton, N.J., 1970.

[Hun04] David R. Hunter. MM algorithms for generalized Bradley-Terry models. *Ann. Stat.*, 32(1):384–406, 2004.

[Kes00] Stefan Kesenne. Revenue sharing and competitive balance in professional team sports. *Journal of Sports Economics*, 1(1):56–65, 2000.

[KG90] Maurice Kendall and Jean Dickinson Gibbons. *Rank correlation methods*. A Charles Griffin Title. Edward Arnold, London, fifth edition, 1990.

[KL17] Graham Kendall and Liam J. A. Lenten. When sports rules go awry. *Eur. J. Oper. Res.*, 257(2):377–394, 2017.

[LDGL20] Benoit Legat, Oscar Dowson, Joaquim Dias Garcia, and Miles Lubin. Mathoptinterface: a data structure for mathematical optimization problems, 2020.

[Len16] Liam J. A. Lenten. Mitigation of perverse incentives in professional sports leagues with reverse-order drafts. *Rev. Ind. Org.*, 49(1):25–41, Aug 2016.

[LMB18] Michael J. Lopez, Gregory J. Matthews, and Benjamin S. Baumer. How often does the best team win? A unified approach to understanding randomness in North American sport. *Ann. Appl. Stat.*, 12(4):2483–2516, 12 2018.

[Luc59] R. Duncan Luce. *Individual Choice Behavior: A Theoretical Analysis*. John Wiley, 1959.

[LvAAM18] M. Leeds, P. von Allmen, and V. A. Matheson. *The Economics of Sports*. New York: Routledge, 2018.

[McC99] S. Thomas McCormick. Fast algorithms for parametric scheduling come from extensions to parametric maximum flow. *Oper. Res.*, 47(5):744–756, 1999.

[McI16] Sean McIndoe. Does the NHL have a tanking problem? `https://www.theguardian.com/sport/2016/mar/23/does-the-nhl-have-a-tanking-problem`, March 2016. Retrieved December 8, 2018.

[McM18] Joseph McManus. Ethical considerations & the practice of tanking in sport management. *Sport, Ethics and Philosophy*, 13(2):145–160, 2018.

[MRLL16] Akira Motomura, Kelsey V. Roberts, Daniel M. Leeds, and Michael A. Leeds. Does it pay to build through the draft in the national basketball association? *J. Sports Econ.*, 17(5):501–516, 2016.

[Pau14] Marc Pauly. Can strategizing in round-robin subtournaments be avoided? *Soc. Choice Welf.*, 43(1):29–46, 2014.

[Pay18] Neil Payne. This year's MLB tankfest is epic (and maybe pointless). `https://fivethirtyeight.com/features/this-years-mlb-tankfest-is-epic-and-maybe-pointless/`, April 2018. Retrieved December 8, 2018.

[PSBH10] Joseph Price, Brian P. Soebbing, David Berri, and Brad R. Humphreys. Tournament incentives, league policy, and NBA team performance revisited. *J. Sport. Econ.*, 11(2):117–135, 2010.

[RV17] Aviad Rubinstein and Shai Vardi. Sorting from noisier samples. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 960–972, 2017.

[Sau18] David Sauriol. Is tanking a viable team building strategy? `http://www.rawbw.com/~deano/articles/20180101_SSM_Intro.htm`, January 1, 2018. Retrieved August 28, 2018.

[SB01] Martin B. Schmidt and David J. Berri. Competitive balance and attendance: The case of major league baseball. *Journal of Sports Economics*, 2(2):145–167, 2001.

[SC18] Ildikó Schlotter and Katarína Cechlárová. A connection between sports and matroids: how many teams can we beat? *Algorithmica*, 80(1):258–278, 2018.

[SH13] Brian P. Soebbing and Brad R. Humphreys. Do gamblers think that teams tank? Evidence from the NBA. *Contemp. Econ. Policy*, 31(2):301–313, 2013.

[Sil15] Nate Silver. A simple, Knicks-proof proposal to improve the NBA's draft lottery. `https://fivethirtyeight.com/features/a-simple-knicks-proof-proposal-to-improve-the-nbas-draft-lottery/`, April 2015. Retrieved December 8, 2018.

[Soe11] Brian P. Soebbing. *The Amateur Draft, Competitive Balance, and Tanking in the National Basketball Association*. PhD thesis, University of Alberta. Faculty of Physical Education and Recreation, 2011.

[TT02] Beck A. Taylor and Justin G. Trogdon. Losing to win: tournament incentives in the National Basketball Association. *J. Labor Econ.*, 20(1):23–41, 2002.

[Von17] Allen I. K. Vong. Strategic manipulation in tournament games. *Game. Econ. Behav.*, 102:562–567, 2017.

[Way01] Kevin D. Wayne. A new property and a faster algorithm for baseball elimination. *SIAM J. Discrete Math.*, 14(2):223–229, 2001.

[Wri14] Mike Wright. OR analysis of sporting rules — a survey. *Eur. J. Oper. Res.*, 232(1):1–8, 2014.

[WW12] Christopher Walters and Tyler Williams. To tank or not to tank? Evidence from the NBA. In *MIT Sloan Sports Analytics Conference*, March 2012.

[Zer29] E. Zermelo. Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. *Math. Z.*, 29(1):436–460, 1929.

A Determining win probabilities based on NBA data

This section describes how we determine the winner of each game in our simulation. We assume that $p_{ij}(H, H) = p_{ij}(L, L)$, i.e., that two tanking teams that play each other will have the same win probability as two teams exerting high effort in the game. As a result, we need only to set $p_{ij}(H, H)$. Our target is to make our simulated league resemble the NBA on a high level. In particular, we evaluate the distribution of end-of-season win percentages of teams in our simulation compared to 14 years of NBA data, 2004–2018, excluding the strike-shortened 2011 season.

We consider two models for determining the probability $p_{ij}(H, H)$. The first is the (Zermelo)-Bradley-Terry model [Zer29, BT52, Luc59]. The second is the *noisy comparison* model; see, e.g., [BM08, RV17].

The Bradley-Terry model involves setting a strength parameter p_i for each team $i \in [n]$, and based on this, the probability $p_{ij}(H, H)$ that team i beats team j is set to $p_i/(p_i + p_j)$. We estimate these strength parameters using a variant on maximum likelihood estimation for the Bradley-Terry model [Zer29, Hun04], described in Algorithm 1. The input to the algorithm takes an $n \times n$ matrix H and n-vector w. The entries H_{ij} of H denote the number of times a team ranked i at the end of the season beat a team ranked j at the end of the season, based on the historical NBA data. For w, for each $i \in [n]$, define

$$w_i := \sum_{j \neq i} \frac{H_{ij}}{H_{ij} + H_{ji}}.$$

The output of Algorithm 1 is the strength of teams 1 through n, sorted from strongest to weakest team.

In contrast to the Bradley-Terry model, in the noisy comparison model, there is a single fixed value $\gamma = p_{ij}(H, H) \geq 1/2$ for all $i, j \in [n]$ such that $\pi_i^{\text{TRUE}} < \pi_j^{\text{TRUE}}$. The main advantage of the noisy comparison model is that it has a single parameter, and hence it is easy to analyze and compare the effects of varying this parameter.

To compare these options, we calculate the mean squared error of the win percentage distribution obtained by our simulation compared to end-of-season NBA data. Specifically, let $\mathcal{Y}$ denote the years of data that we use, and for $\text{yr} \in \mathcal{Y}$ and $r \in [n]$, let $\text{winpct}_r^{\text{yr}}$ denote the win percentage at the end of the season by the team with rank r. Similarly, let $\overline{\text{winpct}_r}$

Algorithm 1 Variant of Maximum Likelihood Estimation for Bradley-Terry Parameters

1: **function** MLE(H, w)
2: $\quad p_i \leftarrow 1/n$ for $i \in [n]$
3: $\quad$ **for** step $= 1, \dots, 1000$ **do**
4: $\qquad p'_i \leftarrow p_i$ for $i \in [n]$
5: $\qquad$ **for** $i = 1, \dots, n$ **do**
6: $\qquad\quad v \leftarrow \sum_{j \neq i} (H_{ij} + H_{ji})/(p_i + p_j)$
7: $\qquad\quad p'_i \leftarrow w_i / v$
8: $\qquad$ **if** $\|p - p'\|_1 < 10^{-5}$ **then**
9: $\qquad\quad$ Go to step 11
10: $\qquad p_i \leftarrow p'_i / \sum_{j \in [n]} p_j$ for $i \in [n]$
11: $\quad$ **return** p, sorted from largest to smallest

denote the average win percentage of a team ranked r at the end of the simulated season. The error is

$$\text{mean squared error} = \frac{1}{n \cdot |\mathcal{Y}|} \sum_{r \in [n]} \sum_{\text{yr} \in \mathcal{Y}} (\overline{\text{winpct}}_r - \text{winpct}_r^{\text{yr}})^2.$$

For setting the value of $p_{ij}(H, H)$, we perform 100,000 replications of each data point, i.e., each simulated regular season. Figure 6 shows the mean squared error for the estimated Bradley-Terry model and for various values of γ, when the number of selfish teams is either 0, 15, or 30. We see that the Bradley-Terry model would be the most accurate when there are no selfish teams, and $\gamma = 0.7$ would be the most accurate when all teams are selfish. Using binary search over γ, we find that $\gamma = 0.71425$ gives a reasonable loss, regardless of the number of selfish teams. This value is similar to the parameters used in Lopez et al. [LMB18], who statistically estimate the value of a quantity akin to γ (namely, they estimate the probability that a better team will win at a neutral site) to be 0.67.

As another way to validate of our model, Figure 7 plots the number of teams that are effectively eliminated throughout the season.[8] This represents the maximum number of teams that could tank under our model. We compare these historical numbers to the expected number of eliminated teams from our simulation, averaged across all tanking probabilities, shown as a dashed line in the figure. Note that a real NBA season does not play a full three rounds as in our simulated season, but we normalize for this on the horizontal axis. The simulation appears faithful to the real data, as it is close to the average of the historical numbers. Hence, the results regarding tanking and competitive balance that we will now present in Section 5.3 can be reasonably expected to also apply in practice.

Lastly, we note that we experimented with several values γ, and our results remained qualitatively similar, except that the breakpoint δ should be set later for smaller γ. This is intuitively obvious: when γ is larger, weak teams are eliminated from the playoffs earlier.

[8] Some seasons may end with fewer than 14 teams eliminated, e.g., when there is a tie between the 16th and 17th place teams.

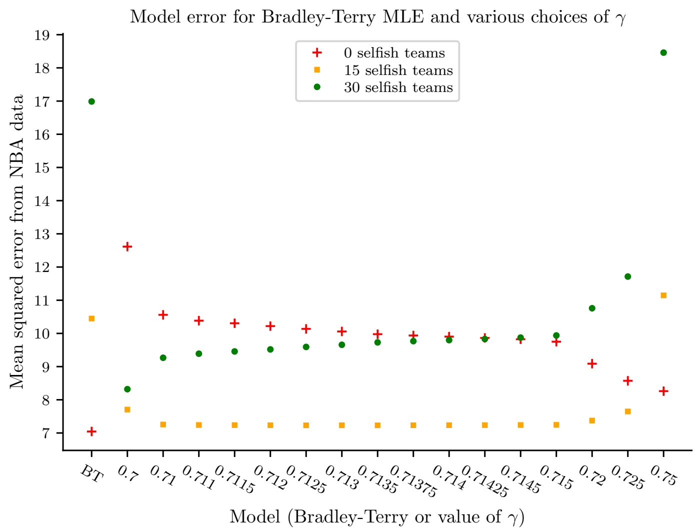

Figure 6: Mean squared error from 2004–2018 NBA seasons, comparing the Bradley-Terry model and various values for γ in the noisy comparison model, plotted when there are 0, 15, or 30 selfish teams. Setting $\gamma = 0.71425$ for $p_{ij}(H, H)$ when $\pi_i^{\text{TRUE}} < \pi_j^{\text{TRUE}}$ is balances the error across these possibilities.

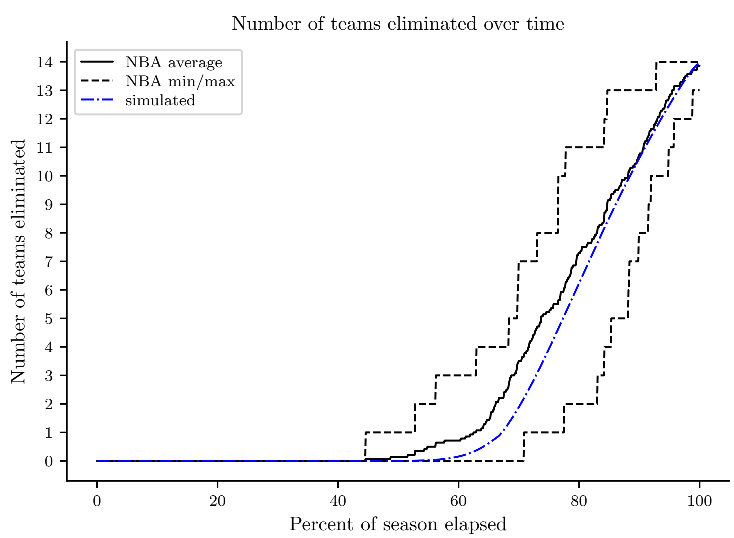

Figure 7: Number teams eliminated as a function of the percent of the season played.

B Mixed-Integer Programming Formulation for Mathematical Elimination from Playoffs

Assume that games 1 through $t'-1$ have been played and we would like to determine whether team k has been eliminated from position n^*. We are given the outcomes of the first $t'-1$ games and the schedule of the remainder of the season. Let $\mathcal{G}_t$ denote the set of teams playing in game $t \in [T]$, and set M as the total number of games each team plays in a season. We assume that team k wins its remaining games and ends the season with $\bar{W}$ wins.

The method we present is based on mixed-integer programming. The general version of this problem was proven $\mathcal{NP}$-Hard by [McC99], but whether that computational complexity is maintained for our specific model remains an open problem. The integer programs are solved in Julia via the JuMP [DHL17] interface with Gurobi version 8.1.0 [Gur18].

B.1 Binary formulation

We define several variables and auxiliary variables. For each game t, say between teams i and j, i.e., $\mathcal{G}_t = \{i,j\}$, we introduce binary variables x_{it} and x_{jt}, exactly one of which will take value one (representing the team that wins game t). Let $\overline{\mathcal{T}} \subseteq [T]$ denote the set of games for which outcomes are fixed, e.g., $\overline{\mathcal{T}} \supseteq [t'-1]$ and $\overline{\mathcal{T}}$ contains all games $t \in [t', T]$ such that $k \in \mathcal{G}_t$; for each $t \in \overline{\mathcal{T}}$, we set $x_{it} = \bar{x}_{it}$, assigned based on the fixed outcomes. We use w_i for the final number of wins for team i. The objective is to find the best possible rank team k can achieve, which will be one more than the number of teams that finish with at least $w_i \geq \bar{W}+1$ wins. For all $i \neq k$, let z_i be a binary variable that takes value 1 if and only if $w_i \geq \bar{W}+1$. This is achieved by a set of constraints modeling $(w_i - \bar{W})/M \leq z_i \leq \max\{0, w_i - \bar{W}\}$, where the lower bound is valid because $w_i - \bar{W} \leq M$ and forces $z_i = 1$ when appropriate, and we apply a standard linearization trick on the upper bound.

$$
\begin{aligned}
\min_{w,x,y,z} \quad & 1 + \sum_{i \in [n]} z_i \\
& w_i = \sum_{\substack{t \in [T]: \\ i \in \mathcal{G}_t}} x_{it} && \forall\, i \in [n] && \text{(number wins for team } i\text{)} \\
& x_{it} = \bar{x}_{it} && \forall\, t \in \overline{\mathcal{T}}, i \in \mathcal{G}_t && \text{(fixed outcomes of games in set } \overline{\mathcal{T}}\text{)} \\
& x_{it} + x_{jt} = 1 && \forall\, t \in [t', T], \mathcal{G}_t = \{i,j\} && \text{(exactly one team wins game } t\text{)} \\
& z_i \geq (w_i - \bar{W})/M && \forall\, i \in [n] && \text{(force } z_i = 1 \text{ if } i \text{ has more wins)} \\
& z_i \leq y_i && \forall\, i \in [n] && (z_i = 0 \text{ if } i \text{ does not have more wins)} \\
& y_i \geq w_i - \bar{W} && \forall\, i \in [n] && (y_i \geq \max\{0, w_i - \bar{W}\}) \\
& y_i \geq 0 && \forall\, i \in [n] && (y_i \geq \max\{0, w_i - \bar{W}\}) \\
& x_{it} \in \{0,1\} && \forall\, t \in [T], i \in \mathcal{G}_t \\
& z_i \in \{0,1\} && \forall\, i \in [n]
\end{aligned}
$$

The output of this mixed-integer program is the best rank that team k can possibly achieve and the set of outcomes for the remainder of the season in which team k can attain that rank. Note that variables $\{w_i\}_{i\in[n]}$ can be projected out from the formulation, but this can be handled by solver presolve routines. Observe that we can reuse the same formulation to calculate the minimum rank for other teams than k by modifying the value of $\bar{W}$ and using a different set of fixed games $\overline{\mathcal{T}}$.

B.2 Speeding up the optimization process

For all $i, j \in [n]$, let g_{ij} be the number of remaining games between teams i and j for which the outcome is not yet fixed (i.e., $g_{ij} = |\{t \in [T] \setminus \overline{\mathcal{T}} : i \in \mathcal{G}_t\}|$).

- If $\bar{w}_i > \bar{W}$ or $\bar{w}_i + g_i \leq \bar{W}$, then for each t such that x_{it} is not yet fixed, set $x_{it} = 1$ for all t such that $i \in \mathcal{G}_t$, and set $x_{jt} = 0$ for the other team j playing in game t. (These teams can win all of their remaining games without affecting the rank of team k.) We can add those games whose outcomes we fixed to $\overline{\mathcal{T}}$ and repeat until no further updates are possible.
- We add an early stopping criterion for the optimization, to exit when a feasible integer solution is found with value n^* or better, or when the dual bound is greater than n^*.

B.3 Alternative general integer formulation

An equivalent mixed-integer programming formulation for the minimum rank problem is as follows. It has the advantage of having fewer variables, but the disadvantage that these variables can take general integer values rather than only binary values.

Let x_{ij} be an integer variable representing the number of the remaining games that team i wins against team j. We use $\bar{x}_{ij}$ to denote a prescribed set of wins by team i over team j, such as based on outcomes of earlier games or some heuristic solution. We let $\bar{g}_{ij}$ be the number of total games between teams i and j.

$$
\begin{aligned}
\min_{w,x,y,z} \quad & 1 + \sum_{i\in[n]} z_i \\
& w_i = \sum_{j\neq i} x_{ij} && \forall\, i \in [n] && \text{(number wins for team } i\text{)} \\
& x_{ij} \geq \bar{x}_{ij} && \forall\, i, j \in [n] : i \neq j && \text{(account for settled games)} \\
& x_{ij} + x_{ji} = \bar{g}_{ij} && \forall\, i, j \in [n] : i < j && \text{(all games have a winner)} \\
& z_i \geq (w_i - \bar{W})/M && \forall\, i \in [n] && \text{(force } z_i = 1 \text{ if } i \text{ has more wins)} \\
& z_i \leq y_i && \forall\, i \in [n] && (z_i = 0 \text{ if } i \text{ does not have more wins)} \\
& y_i \geq w_i - \bar{W} && \forall\, i \in [n] && (y_i \geq \max\{0, w_i - \bar{W}\}) \\
& y_i \geq 0 && \forall\, i \in [n] && (y_i \geq \max\{0, w_i - \bar{W}\}) \\
& 0 \leq x_{ij} \leq \bar{g}_{ij},\ x_{ij} \in \mathbb{Z} && \forall\, i, j \in [n] : i \neq j \\
& z_i \in \{0, 1\} && \forall\, i \in [n]
\end{aligned}
$$

B.4 Solving fewer mixed-integer programs

As observed by Wayne [Way01] and Adler et al. [AEHO02] (see also [GM02]), it actually suffices to solve one mixed-integer program after each game to simultaneously determine all of the teams that are eliminated. This hinges on the following simple, yet extremely useful, proposition. We let $\bar{w}_i$ denote the number of wins team i currently has, after game $t' - 1$.

Proposition 15. *There exists a team-independent value W after each game t such that any team i can achieve position n^* or better if and only if the number of games it has remaining is at least $W - \bar{w}_i$.*

Proof. Among all possible outcomes for the remainder of the season, consider the scenario $\mathcal{O}$ in which the last team to make the playoffs has the fewest number of wins. Let W equal this value.

Observe that if a team i has at least $W - \bar{w}_i$ games remaining, then we can find a new scenario $\mathcal{O}'$, which is obtained from $\mathcal{O}$ but with team i winning all its remaining games. Relative to scenario $\mathcal{O}$, in scenario $\mathcal{O}'$, all teams except team i have the same or fewer wins at the end of the season, and team i has at least W wins. It follows that team i will be at least tied for the last playoff position.

Conversely, if team i has fewer than $W - \bar{w}_i$ games remaining, then it cannot make the playoffs in any scenario by assumption on how $\mathcal{O}$ was chosen. □

Adler et al. also provide a mixed-integer programming formulation using Proposition 15, which we adapt to our setting below.

$$
\begin{aligned}
\min_{W,x,\alpha} \quad & W \\
& x_{ij} \geq \bar{x}_{ij} && \forall\, i,j \in [n] : i \neq j && \text{(account for settled games)} \\
& x_{ij} + x_{ji} = \bar{g}_{ij} && \forall\, i,j \in [n] : i < j && \text{(all games have a winner)} \\
& W \geq \sum_{j \neq i} x_{ij} - M\alpha_i && \forall\, i \in [n] && \text{(bound on } W\text{)} \\
& \sum_{i \in [n]} \alpha_i = n^* && && (n^* \text{ teams make playoffs}) \\
& 0 \leq x_{ij} \leq \bar{g}_{ij},\ x_{ij} \in \mathbb{Z} && \forall\, i,j \in [n] : i \neq j \\
& \alpha_i \in \{0,1\} && \forall\, i \in [n]
\end{aligned}
$$

Note that the above problem can be reformulated using only binary variables, if desired. The advantage of this formulation is that it has fewer variables, and that only one mixed-integer program needs to be solved after each game, rather than as many as one per each team (though, in practice, one rarely needs to solve more than one program per iteration even with the team-based approach). The disadvantage of this formulation is that we are unable to provide an a priori value of the objective at which the optimization can exit early (as in the team-based formulation, in which one can stop after the maximum best rank is n^*). For this reason, we do not use this formulation in our experiments.

B.5 Open questions on computation of mathematical elimination

Although the generic problem of identifying whether team k can achieve rank n^* or better at the end of the season is $\mathcal{NP}$-Hard, in a season of length T, how many times will an $\mathcal{NP}$-Hard problem need to be solved? For example, if we have computed the value W from Proposition 15 after game $t' - 1$, then team i beats team j in game t' and the solution corresponding to W had $x_{ij} - \bar{x}_{ij} > 0$, then no recomputation is necessary.