CESifo WORKING PAPERS

7474
2019

January 2019

Strategic Investment Decisions in Multi-stage Contests with Heterogeneous Players

Christian Deutscher, Marco Sahm, Sandra Schneemann, Hendrik Sonnabend

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7474*
Category 11: Industrial Organisation

Strategic Investment Decisions in Multi-stage Contests with Heterogeneous Players

Abstract

We study strategic investment decisions in multi-stage contests with heterogeneous players. Our theoretical model of a round-robin rank-order tournament predicts that players conserve resources in a current contest to spend more in the subsequent contest if the degree of heterogeneity in the current (subsequent) contest is sufficiently large (small). We confirm these predictions using data from German professional soccer, where players are subject to a one-match ban if they accumulate five yellow cards. We find that players with four yellow cards facing the risk of being suspended for the next match are (i) less likely to be fielded when the heterogeneity in the current match increases and (ii) more likely to receive a fifth yellow card in the current match when the heterogeneity in the next match increases or when the heterogeneity in the next match but one (for which they return from their ban) decreases.

JEL-Codes: C730, D840, L830, M510, M540.

Keywords: tournaments, multi-stage contests, heterogeneity, anticipating behavior, shadow effects.

Christian Deutscher
Department of Sport Science
Bielefeld University
Postfach 100131
Germany - 33501 Bielefeld
christian.deutscher@uni-bielefeld.de

Marco Sahm
Otto-Friedrich-Universität Bamberg
Feldkirchenstraße 21
Germany - 96052 Bamberg
marco.sahm@uni-bamberg.de

Sandra Schneemann
Department of Sport Science
Bielefeld University
Postfach 100131
Germany - 33501 Bielefeld
sandra.schneemann@uni-bielefeld.de

Hendrik Sonnabend
University of Hagen
Department of Economics
Universitätsstraße 11
Germany - 58097 Hagen
hendrik.sonnabend@fernuni-hagen.de

1 Introduction

Contests are common in real-life settings such as labor markets, industrial economics, sports, and public choice, especially when ability levels vary across contestants. Ability heterogeneity has an adverse effect on contestants' willingness to invest resources, often referred to as the *discouragement effect* (e.g. Konrad, 2009).[1] In a dynamic (tournament) setting with multiple rounds and asymmetric abilities, a player's investment in a current stage might also be affected by the strength of the next-round opponents.

To study anticipating behavior that results in the strategic allocation of resources over the course of rank-order tournaments, we propose a round-robin tournament model in which three heterogeneous contestants have information about upcoming opponents. The players can *self-restrict* and conserve resources in the first round, such that they can contribute more resources in the second round. Our model establishes two main results. We find that contestants self-restrict in the first round, as long as the gap in abilities at this stage is sufficiently large. This result corresponds to the discouragement effect. We also note a spillover effect, such that contestants also self-restrict when heterogeneity in the subsequent contest (second round) is sufficiently small.

To establish empirical evidence, we use data from the German *Bundesliga*, a professional association soccer league with a double round-robin tournament structure. Professional soccer leagues offer a formalized structure and full information about the number of rounds and opponents at the beginning of the tournament, they also entail contests with high incentives. To test our theoretical predictions, we leverage a unique rule stating that a player who accumulates a critical number of yellow cards (also called 'bookings') will be suspended for the next match. Teams thus may self-restrict, by not fielding a player threatened by a ban or by strategically provoking a booking to obtain a suspension for the subsequent match.

In line with the proposed model, we find that teams systematically self-restrict in round t if the difference in ability, or *heterogeneity*, in t is sufficiently large. Results regarding the effect of heterogeneity in $t+1$ on the decision to hold back resources in t are mixed, but they still provide strong evidence of anticipating behavior.

The article proceeds as follows. Section 2 gives a brief review of the related literature. Section 3 presents the theoretical framework. In Section 4 we empirically test the implications of our analytical model with field data. Finally, Section 5 concludes.

2 Previous Studies

This paper contributes to three strands of literature. First, it relates to the work on the impact of heterogeneity in contests in general, starting with Rosen (1986), who clearly suggests that unbalanced match-ups result in lower effort levels. Theoretical work strongly supports this 'discouragement effect'; see, e.g., Baik (1994), Stein (2002), Szymanski (2003), Konrad (2009), or March and Sahm (2018). Intuitively, underdogs lower their willingness to invest when winning

[1] We define resources as an input needed to improve a contestant's winning probabilities, given the opponent's ability and investment level. Several synonyms appear in tournament theory and economic analyses of conflict, such as effort, resources, endowments, or force size.

becomes too costly, while favorites respond by lowering their investment as the outcome appears predefined. Empirical evidence obtained in laboratory settings (e.g. Dechenaux et al., 2015; Hart et al., 2015; March and Sahm, 2017) and from the field (e.g. Ehrenberg and Bognanno, 1990; Sunde, 2009; Schneemann and Deutscher, 2017) similarly affirms these discouragement effects. We therefore expect that the investment of resources depends critically on heterogeneity in the competition.

Second, this paper contributes to a growing body of research on anticipating behavior and inter-temporal effort provision in multi-period contests with fixed price structures. For instance, Altmann et al. (2012) document that participants exert excess effort in the initial round of a two-stated contest in a laboratory setup, seemingly due to their limited forward-looking capabilities.

Third, and most directly, our study also relates to 'spillover' or 'carryover' effects in tournaments with heterogeneous players, such that the strength of past or future opponents influences the current effort level. Brown and Minor (2014), using data from top-level tennis, find that the probability that the favorite wins the current stage decreases with the strength of an expected future opponent. They argue that taking the competitor's ability in the next round into account changes the favorite's valuation of the tournament and hence optimal effort provision.[2] Lackner et al. (2015) also find that the intensity of play (measured by personal fouls) increases when the expected relative ability of the next stage contestants decreases in NBA and NCAA playoffs. Similar to Brown and Minor (2014), their approach builds on the idea that future opponents affect the probability of winning the tournament, or rather the continuation value, and thus the incentives to exert effort at the current stage. Harbaugh and Klumpp (2005) also consider budget constraints in their tournament model, such that they show that both underdogs and favorites have distinct incentives to reserve resources for upcoming battles; the underdogs benefit more from investing more of their resources in initial stages. The introduction of a 'rest day' also improves the performance of the favorites in NCAA basketball, indicating a shift in the incentives to conserve resources. However, direct evidence of such strategic behavior is missing.

In turn, the main novelty of our empirical approach is that we aim to provide direct evidence for strategic action to achieve an optimal allocation of resources *across* the tournament. In addition, rather than using scores or personal fouls, which might be rather fuzzy measures of effort provision, we argue that picking up deliberate yellow cards to trigger a one-game suspension represents a direct measure of strategic behavior.

3 Theoretical Model

We consider round-robin tournaments with three heterogeneous players to build a formal model of simple rank-order tournaments. Players may restrict their resources in the first match to increase the probability of having unrestricted resources in their second match. Whether a player uses this option in equilibrium depends on the extent of heterogeneity between that player and the opponents in his first and second match.

[2] Hill (2018) offers a replication study with data from NBA playoff matches.

3.1 Assumptions

The structure of the tournament

A *round-robin tournament*, also referred to as an all-play-all tournament, matches each participant with each other participant in a pairwise contest, ranks them according to the number of matches won, and awards prizes according to this ranking. For simplicity, we focus on round-robin tournaments with three risk-neutral players and consider an exogenous sequence in which player 1 is matched with player 2 in the first match, player 1 is matched with player 3 in the second match, and player 2 is matched with player 3 in the third match (Krumer et al., 2017a; Sahm, 2019).[3]

If a player earns two victories, he is ranked first and the player with one victory is ranked second; if there is a tie because each player has won one match, the ranks are assigned randomly with equal probabilities of 1/3 for each player and rank. The player ranked first (last) receives the first (last) prize, the value of which is identical for all players and normalized to 1 (0). The value of the second prize is assumed to be identical for all players as well, and fixed at 1/2, i.e., half of the first prize. This prize structure (is the only one that) ensures a fair tournament in the sense that all potential differences in players' equilibrium winning probabilities and expected payoffs result from differences in their abilities, not from their position in the sequence of matches (Laica et al., 2017).

The structure of the resulting sequential game with its $2^3 = 8$ potential courses is depicted in Figure 1. The seven nodes $k \in \{A, \dots, F\}$ represent all combinations for which the ranking of the tournament has not been fully determined when the respective match starts.

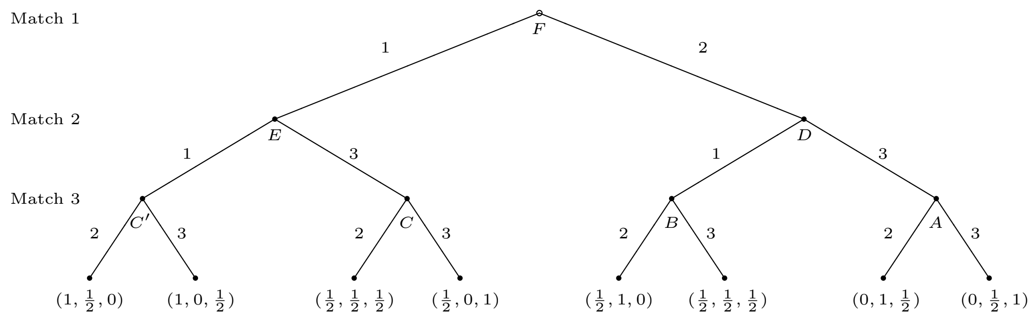


Figure 1: 3-player round-robin tournament with exogenous sequence

The matches of the tournament

Each match of the tournament is organized as a lottery contest between two potentially heterogeneous players, A and B, with linear costs of effort.[4] Specifically, player A's probability of

[3] Apart from renumbering players, this exogenous sequence is unique. The use of exogenous sequences is prevalent in sports tournaments. Sahm (2019), Krumer et al. (2017b), and Laica et al. (2017) also discuss endogenous sequences in which the outcome of the first match determines the order of the two remaining matches.

[4] The non-perfectly discriminating lottery contest offers concreteness and accounts for random elements that appear in most sport competition. Sahm (2019) and Laica et al. (2017) consider round-robin tournaments with general Tullock contests, including the limit case of all-pay auctions, as studied by Krumer et al. (2017a,b). The alternative assumption of zero effort costs but limited resources in round-robin tournaments with three players implies a quasi-simultaneous structure with multiple equilibria (Dagaev and Zubanov, 2017) and, thus, is not a fruitful modeling option in our context.

winning match $k \in \{A, \dots, F\}$ is

$$p_A^k = \begin{cases} 1/2 & \text{if } x_A^k = x_B^k = 0, \\ \frac{\theta_A^k x_A^k}{\theta_A^k x_A^k + \theta_B^k x_B^k} & \text{else,} \end{cases}$$

where θ_i^k describes ability, and x_i^k denotes the effort of player $i \in \{A, B\}$ in match k (Leininger, 1993; Baik, 1994).

Player A chooses $x_A^k \geq 0$ to maximize his expected payoff

$$E_A^k = p_A^k(w_A^k - x_A^k) + \left(1 - p_A^k\right)(\ell_A^k - x_A^k), \tag{1}$$

where w_i^k denotes player i's expected continuation payoff from winning match k, and ℓ_i^k denotes his expected continuation payoff from losing match k, with $w_i^k \geq \ell_i^k \geq 0$ for $i \in \{A, B\}$. For $w_A^k = \ell_A^k$, the optimal choice is $x_A^k = 0$ for any $x_B^k \geq 0$. If $x_A^k = 0$ and $w_B^k > \ell_B^k$, player B has no best reply unless there is a smallest monetary unit $\varepsilon > 0$, in which case the best reply is $x_B^k = \varepsilon$. As $\varepsilon \to 0$, in the limit, $x_B^k \to 0$, and $p_B^k \to 1$. Otherwise, there is a unique Nash equilibrium in pure strategies (see, e.g., Cornes and Hartley, 2005). The equilibrium effort levels can then be derived from the following necessary conditions

$$\frac{\partial E_i^k}{\partial x_i^k} = \frac{\theta_i^k \theta_j^k x_j^k}{(\theta_i^k x_i^k + \theta_j^k x_j^k)^2}(w_i^k - \ell_i^k) - 1 = 0$$

yielding

$$x_i^k = \frac{\theta_i^k \theta_j^k (w_i^k - \ell_i^k)^2 (w_j^k - \ell_j^k)}{[\theta_i^k (w_i^k - \ell_i^k) + \theta_j^k (w_j^k - \ell_j^k)]^2} \tag{2}$$

for $i, j \in \{A, B\}$ with $i \neq j$. The resulting equilibrium winning probabilities equal

$$p_i^k = \frac{\theta_i^k (w_i^k - \ell_i^k)}{\theta_i^k (w_i^k - \ell_i^k) + \theta_j^k (w_j^k - \ell_j^k)}. \tag{3}$$

Inserting (2) and (3) into (1) yields the expected equilibrium payoffs

$$E_i^k = \ell_i^k + \frac{(\theta_i^k)^2 (w_i^k - \ell_i^k)^3}{[\theta_i^k (w_i^k - \ell_i^k) + \theta_j^k (w_j^k - \ell_j^k)]^2}. \tag{4}$$

The possibility of self-restriction

In practice, most tournaments offer the possibility that players restrict their current resources in order to increase the probability of having unrestricted resources available in future matches. For example in team sports, such self-restriction might involve giving particular, important players a break, so they can rest for a future contest. To keep the analysis tractable, we model this self-restriction opportunity in a simple manner, making the following assumptions.

Players may differ in their basic abilities.[5] Let $\theta_i > 0$ denote the basic ability of player

[5] In the present framework of a Tullock contest with linear effort costs, heterogeneity in players' abilities is equivalent to heterogeneity in their (linear) costs of effort, as well as to heterogeneity in their valuations of the prizes (see, e.g., Cornes and Hartley, 2005; Ryvkin, 2013).

$i \in \{1,2,3\}$. Before the tournament starts, players 1 and 2 make a binary choice to restrict themselves (R) or not (NR) in their first match (match 1). If player $i \in \{1,2\}$ chooses R, his ability in his first match will be restricted to $\theta_i^F = r\theta_i$ with some $r \in (0,1)$, but he will have unrestricted ability $\theta_i^k = \theta_i$ in his second match ($k \in \{D,E\}$ if $i = 1$ and $k \in \{A,B,C,C'\}$ if $i = 2$). In contrast, if player $i \in \{1,2\}$ chooses NR, his ability in his first match will be unrestricted, i.e. $\theta_i^F = \theta_i$, but he will face some positive probability $\pi \in (0,1)$ of having to compete with restricted ability in his second match. To capture the related uncertainty in reduced form, we assume that his (expected) ability in his second match equals $\theta_i^k = q\theta_i$ with $q = 1 - \pi(1-r) \in (r,1)$. For simplicity, we assume that player 3 has unrestricted ability in both of his matches[6] which is normalized to one[7], $\theta_3^k = \theta_3 = 1$ for all $k \in \{A,\ldots,E\}$.

This round-robin tournament with three players and an opportunity of self-restriction (for players 1 and 2) constitutes a sequential game Γ with four stages: In Stage 0, players 1 and 2 decide simultaneously whether to self-restrict or not in match 1; their choices are observed by all players. In Stage 1, players 1 and 2 decide simultaneously about their effort in match 1, and the outcome of match 1 is observed by all players. In Stage 2, players 1 and 3 decide simultaneously about their effort in match 2, and the outcome of match 2 is observed by all players. Finally, in Stage 3, players 2 and 3 decide simultaneously about their effort in match 3.

3.2 Results

For each feasible combination of exogenous parameters $(\theta_1, \theta_2, r, q)$, the game Γ can be solved by backward induction for its subgame perfect equilibrium (SPE) through repeated use of Equations (2) – (4). Appendix A illustrates this procedure for the example of $(\theta_1, \theta_2, r, q) = (1, 1, 1/2, 3/4)$.

We are particularly interested in identifying the conditions in which players 1 or 2 choose self-restriction in the first match as part of an equilibrium strategy. Because the comparative statics are analytically not tractable, we study the effect of variations of the exogenous parameters $(\theta_1, \theta_2, r, q)$ on equilibrium behavior numerically. We use the spreadsheet program *Microsoft Excel* to compute the SPE of Γ on a grid with more than $200,000$ grid points, increasing θ_i from 0.1 to 10 in steps of 1% (and from 0.001 to 1000 in steps of 10%, respectively) for $i \in \{1,2\}$, and r as well as q from 0.5 to 0.9 in steps of 0.1 subject to $r < q$.[8] The calculations demonstrate that, depending on the parameters, all kinds of equilibrium behavior by players 1 and 2 may arise in the first stage.

Proposition 1 *For each of the following combinations of first-stage behavior by players 1 and 2, there are parameters $(\theta_1, \theta_2, r, q)$, such that the respective combination is part of the players' strategies in the SPE of game Γ:*

(a) no player chooses self-restriction R,

(b) only the weaker player chooses self-restriction R,

[6] Allowing player 3 to self-restrict as well considerably complicates the analysis, due to second-order strategic incentives: The self-restriction of one or both players in the first match may trigger/prevent the self-restriction of player 3 in the second match. The first-order strategic incentives of players 1 and 2 for restricting themselves remain present and do not fundamentally change though, because in the SPE of the game, players 1 and 2 correctly anticipate the (expected) ability of player 3 as their upcoming opponent.

[7] Fixing one of the three abilities is without loss of generality, because only relative abilities matter.

[8] The respective spreadsheets are available on request.

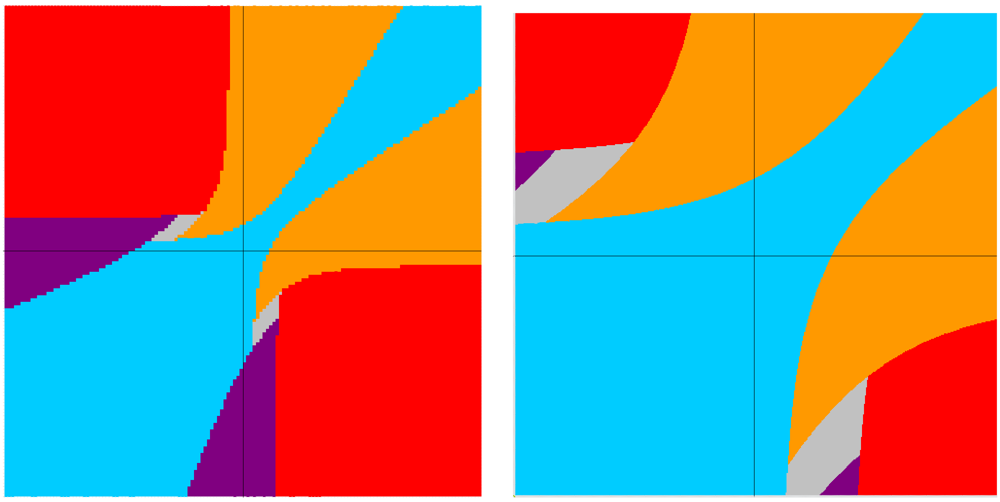
(c) both players choose self-restriction R,

(d) only the stronger player chooses self-restriction R,

(e) both players choose self-restriction R with positive probability (equilibrium in mixed strategies).

Figure 2 illustrates Proposition 1. The players' behavior only depends on their relative basic abilities, such that the diagrams are symmetric around the 45°-line on which $\theta_1 = \theta_2$. We thus can focus on cases in which player 1 is at least as able as player 2, i.e. the grid points on and below the 45°-line. Moreover, the left panel of Figure 2 encompasses a range of parameters in which the maximum difference in relative abilities of players 1 and 2 is 1:1,000,000, whereas the right panel zooms into the practically more relevant range in which this maximum difference is 1:100.

Figure 2: Comparative Statics w.r.t. θ_1 and θ_2



Note: This figure illustrates the effects of variations in θ_1 and θ_2. On the horizontal (vertical) axis of the left diagram, θ_1 (θ_2) increases from 0.001 to 1000 in steps of 10% for normalized $\theta_3 = 1$. The right diagram zooms in, such that θ_1 (θ_2) increases from 0.1 to 10 in steps of 1%, again for normalized $\theta_3 = 1$. In both diagrams, the vertical (horizontal) bold line represents $\theta_1 = 1$ ($\theta_2 = 1$). Each color represents a certain type of equilibrium behavior of players 1 and 2 in the first stage:

- blue ≙ no player chooses R,
- orange ≙ only the weaker player chooses R,
- red ≙ both players choose R,
- violet ≙ only the stronger player chooses R,
- grey ≙ both players choose R with positive probability (equilibrium in mixed strategies).

Around the 45°-line, such that players 1 and 2 have similar basic abilities, they never self-restrict. Self-restriction by one or even both players occurs only if the difference between their abilities θ_1 and θ_2 is sufficiently large.[9] Therefore,

Hypothesis 1 *The larger the difference in basic abilities θ_1 and θ_2 in the current match, the more likely a player chooses self-restriction.*

[9]More precisely, the diagrams suggest that, ceteris paribus, the greater the ability of the stronger (weaker) player, the more (less) likely it is that the weaker (stronger) player self-restricts.

The intuition is that players cannot afford to restrict themselves in their first match if the match will be close. Only if they are sufficiently sure they will win or lose, due to a large gap in abilities will players conserve their strength for their second match.

The scale on both axes of the diagrams in Figure 2 also is exponential, such that the ratio of the basic abilities of players 1 and 2 is constant along any parallel to the 45°-line. Moving from the lower left to the upper right on any such parallel, the ability of player $i \in \{1, 2\}$ increases relative to the ability of player 3. Because the region below the 45°-line in which player $i \in \{1, 2\}$ self-restricts is convex in the relevant range (see right panel of Figure 2), there is some ratio α_i of the basic abilities of players 1 and 2 for which the corresponding parallel $\theta_1 = \alpha_i \theta_2$ is tangent to the region in which player $i \in \{1, 2\}$ self-restricts. In this sense, the relative ability $\delta_i = \theta_i / \theta_3$ that characterizes the tangent point is that for which player i's probability of self-restriction is maximal (see also Figure B.1 in Appendix B). These observations suggest

Hypothesis 2 *The larger the difference in (weighted) basic abilities θ_i and $\delta_i \theta_3$ in the future match, the less likely player $i \in \{1, 2\}$ chooses self-restriction.*

Here, the intuition is that it will not pay off for players in the first match to conserve their strength for their second match if they expect that first match to be closer than the second match or more formally, if the (weighted) heterogeneity in the second match is too pronounced.[10]

Finally, if we vary the (expected) levels of restriction r and q, the number of grid points (θ_1, θ_2) for which the players choose self-restriction R in the SPE of game Γ increases in r and decreases in q. Figure B.2 in Appendix B illustrates this finding. Again, the intuition is straightforward: An increase in r means that the restriction in the players' first match is less severe and, therefore, choosing R is less expensive. An increase in q instead means that the possible restriction in the players' second match is less severe or less likely, so the incentive to conserve their full strength by choosing R is less pronounced. Moreover, the variations of r and q in Figure B.2 make clear that our hypotheses are stable across diverse specifications of the model.

4 Empirical Analysis

The empirical analysis uses a rich data set from men's German top-level (Bundesliga) soccer, which comprises detailed information for all players on a game-by-game level and covers five seasons from 2011/12 to 2015/16. Each season is organized as a double round-robin tournament among 18 teams, or 34 games per team and season. The order of games is publicly known in advance of the season.

To test our theoretical predictions, we leverage a unique rule in soccer that mandates players who accumulate a critical number of yellow cards will be suspended for the upcoming round.[11] After serving this suspension, the player regains a 'clean slate' in terms of yellow cards.

[10] In our theoretical model, $\delta_i > 1$ which may be due to the assumption that player 3 has the advantage of unrestricted ability $\theta_3 = 1$ in both matches. In our subsequent discussion of the baseline estimations for the empirical analysis we set $\delta_i = 1$ and check the robustness of our results for alternative values of δ_i.

[11] A player is banned from a (single) game for every five yellow cards he accumulates during the ongoing season. See Fédération International de Football Association (FIFA) (2017) for a list of offenses that trigger a yellow card. After each season, the number of cards and penalties reset. Furthermore, the accumulation of yellow cards counts within competition only, a ban received in league games can be served in league games only.

The strategic use of the rule can be exemplified by the case of two players from the Werder Bremen team who, in a March 2016 hearing, were accused of *intentionally* picking up yellow cards. Their team –struggling against relegation in the end of the season– just has claimed an important home victory against Hannover 96. In the final phase of this match, the two midfielders were booked for 'tactical' fouls. Since it was their 5th and 10th yellow card in season, it was assumed that the players, one of whom was the team captain, provoked the bookings in the hope of avoiding the subsequent game against the league's dominant leader Bayern Munich (which Werder Bremen ultimately lost by a crushing 5-0 score). They later admitted that their plan had been arranged in advance.

The example reveals three key takeaways. First, teams have incentives to let players at risk of a ban deliberately pick up a critical yellow card and take a pause when their resources are needed less.[12] Such a scenario likely arises if an upcoming match is sufficiently heterogeneous, such that the outcome is sufficiently certain.[13] Second, if the current match is homogeneous, such that its outcome is uncertain, players at risk of picking up a one-match ban for five yellow cards may continue to be fielded if the current match is sufficiently balanced. Third, a 'strategic' yellow card also might be more likely near the end of the game, to minimize the risk that the player receives yet another yellow card within the same match, which would lead to an immediate suspension without the count being 'reset'.[14]

Applying the theoretical model established in Section 3, we refer to the strategic use of a one-match ban for five yellow cards as a case of *self-restriction*. Likewise, the decision to not field a player at risk of a ban is as a case of self-restriction, because not fielding such a player ensures that he is available in the upcoming match. The formal analysis implies that a team's decision for or against self-restriction crucially depends on the heterogeneity of competition at time t and $t+1$, such we expect the following two hypotheses to hold:

(1) Teams self-restrict in t if the heterogeneity in match t is sufficiently strong.

(2) Teams self-restrict in t if the heterogeneity in match $t+1$ is sufficiently low.

Our empirical approach is twofold. First, we test the two hypotheses by examining a team's decision, as manifested in the starting line-up. Thus we can analyze the effect of heterogeneity in match t and $t+1$ on the likelihood that teams field players with a pending yellow-card suspension (in t). Second, we study the player's tendency to receive the fifth yellow card in round t followed by a suspension in round $t+1$. We expect to find this tendency to be positively influenced by

[12] About 72% of the players who received a fourth yellow card in match t receive the fifth yellow card within the following five matches. On average, it takes 4.5 matches for a player to receive a fifth yellow card (after a referee has shown him the fourth yellow card). VanDerwerken et al. (2018) indicate that, on balance, players who are one yellow card away from the suspension limit significantly reduce the number of fouls they commit per game, compared with when they are being two cards away.

[13] The use of strategic resting in multi-level contest in sports has been examined by Balsdon et al. (2007) and Raya (2015).

[14] In addition to the accumulation of certain thresholds of yellow cards, a player is unavailable to the team in $t+1$ if he was dismissed (red card) in match t. Dismissals considerably reduce a team's winning probability in game t and often result in players' suspension for multiple matches, so they do not represent advantageous strategic decisions to increase future success. Straight red cards and second yellow cards in the same game thus are not included in our analysis.

the anticipated heterogeneity in match $t+1$, but mitigated by the anticipated heterogeneity in $t+2$.[15]

4.1 Measuring heterogeneity

Our analysis focuses on the impact of heterogeneity (in match t and $t+1$) on a team's decisions to conserve resources in t and $t+1$. We measure the heterogeneity of match t by means of betting odds, which are superior to other indicators of contest heterogeneity, such as differences in league position, because they incorporate information on current form, home field advantages, or coach changes. As Malkiel and Fama (1970) show, financial markets, as represented by betting markets in our study context, digest all information available and are hence efficient (Deutscher et al., 2018). We generate winning probabilities per team and match derived from these odds[16] and calculate the (absolute) difference between the probabilities of victory for the opposing teams (Het_t).[17]

Prior to match t, teams decide whether or not to restrict resources in match t and whether they want to risk them for match $t+1$. That is, the decision about saving resources for match t relies on information prior to match t. Ideally, heterogeneity in our data would capture betting odds for all upcoming matches. Unfortunately, betting odds are only available for the respective next match, that is, for t prior to t and for $t+1$ prior to $t+1$. Yet we consider this concern negligible, because odds do not change considerably after being set, as e.g. noted by Spann and Skiera (2009).

4.2 The starting line-up decision

In this section, we test our hypotheses by analyzing a team's decision to choose a player for the starting lineup who is at risk of a ban due to a critical number of yellow cards. The roster is announced before the game starts. According to Hypothesis 1, we expect this player *not* to be fielded if heterogeneity in game t is sufficiently large. In other words, teams may self-restrict and conserve restricted resources if they are sufficiently sure they will win or lose. According to Hypothesis 2, players threatened by a suspension may be more likely to be fielded if the subsequent match $t+1$ is sufficiently unbalanced, so their absence caused by a fifth yellow card in t is less harmful to the team. We evaluate roster data for each team and match, which refer to both players in the starting lineup and players on the bench. After excluding the first four rounds and the final round of each season, as well as goalkeepers and missing observations, we are left with 39,604 player-game observations.[18]

We estimate a random effects probit model to investigate the impact of heterogeneity in games t and $t+1$ on a team's decision to field a *yellowplayer*:

[15] Because every fifth yellow cards triggers a ban in the *upcoming* match, the self-restriction decision for round t must be made at *t-1*.

[16] The data come from the website www.betexplorer.com.

[17] Values can range between 0 (both teams have equal chances to win) and 1 (one of the teams wins with certainty). Betting odds as a measure of heterogeneity between sport contestants has been repeatedly used in the past, such as Sunde (2009), Deutscher et al. (2013), and Berger and Nieken (2016).

[18] The importance of the goalkeeper position typically leads to few or no strategic variations, and there are only 42 observations of goalkeepers who are in danger of receiving a fifth yellow card within the sample.

$$starting11_{i,t} = \beta_0 + \beta_1 yellowplayer_{i,t} + \beta_2 Het_{i,t} + \beta_3 Het_{i,t+1} + \beta_4 yellowplayer_{i,t} \cdot Het_{i,t} + \beta_5 yellowplayer_{i,t} \cdot Het_{i,t+1} + \gamma' \mathbf{X} + \alpha_i + \varepsilon_{i,t}. \quad (5)$$

The dependent variable, $starting11_{i,t}$, is a binary outcome measure that takes a value of 1 if a kicker i starts in match t and 0 otherwise.[19] Our main variable of interest is a binary variable, such that *yellowplayer=1* indicates a player with a critical number of yellow cards (4, 9, or 14) prior to match t, and *yellowplayer=0* otherwise.

To isolate the effects of heterogeneity on the decision to choose a *yellowplayer* for the starting eleven (captured in β_4 and β_5), our model must control for other factors that could affect the dependent variable. Therefore, $\mathbf{X}$ is a vector of player- and game-specific control variables, including the percentage of minutes a player is on the field in matches prior to a match t, which provides a proxy of the player's importance to the team; team size; game attendance; match day; and dummies controlling for home / away matches and regional rivalry games ('derbies'). Moreover, α_i controls for unobserved player-specific effects. Finally, $\varepsilon_{i,t}$ is the error term that captures all other unobserved factors that influence the *starting11*.

Table 1 contains the descriptive statistics for the main variables. Note that *yellowplayer=1* applies to a total of 3,277 observations for 377 different players.

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
starting11	39604	0.634	0.482	0	1
% of minutes played	39604	0.548	0.315	0	1
Het_t	39604	0.284	0.214	0	0.864
Het_{t+1}	39604	0.287	0.215	0	0.864
yellowplayer	39604	0.083	0.275	0	1

Table 1: Descriptive Statistics - Starting11.

Table 2 includes the estimation results from six variations on our proposed model.[20] These results indicate three main conclusions. First, the percentage of minutes played prior to t increases the probability of being fielded in t. This finding is not surprising; the variable *% of minutes played* was designed explicitly to proxy for player's importance to the team.

Second, the estimated coefficient of the first interaction term (fourth row) $\hat{\beta}_4$ differs significantly from 0, so the tendency to choose a *yellowplayer* as a starter decreases with the heterogeneity of the current competition. That is, the player is protected from the risk of suspension when the match appears to be decided in advance. This result is robust to all specifications in Table 2, and it provides a confirmation of our Hypothesis 1. As a complementary effect, players not at risk of a yellow-card suspension are more likely to start. These findings can easily be linked with the discouragement effect, as we discuss in Section 5.

[19] Potentially, a player might be completely excluded from the roster. However, we decided to restrict the decision to field a player in the starting lineup or leave him on the bench, because the exclusion of a player from a match tends to be due to other reasons, such as injuries or disciplinary issues. We also excluded all observations for which a player served a ban, so banned players are not coded with a 0 for the binary variable indicating the starting eleven.

[20] Robustness checks with logit models or LPM are available in Table C.1 in Appendix C. The main results are preserved.

Third, we do not find similar effects for the next round, and the heterogeneity of the upcoming game $t+1$ does not affect a team's decision (rows five and six). We refrain from taking this result as a disproof of Hypothesis 2 though, because assuming that teams put a player in the starting lineup in t just to provoke a ban in *t+1* would probably undervalue the importance of match t. Consequently, we proceed with a more subtle approach to study anticipating behavior and self-restriction in round $t+1$ in the next section.

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
% of minutes played	2.5327***	2.5328***	2.5328***	2.3850***	2.3839***	2.3839***
	(0.0305)	(0.0305)	(0.0305)	(0.0328)	(0.0328)	(0.0328)
yellowplayer	0.1812***	0.1813***	0.1912***	0.1406***	0.1410***	0.1539**
	(0.0475)	(0.0475)	(0.0602)	(0.0509)	(0.0509)	(0.0644)
Het_t	0.0643***	0.0654***	0.0651***	0.0427	0.0380	0.0379
	(0.0215)	(0.0216)	(0.0217)	(0.0429)	(0.0430)	(0.0430)
yellowplayer * Het_t	-0.3283**	-0.3293**	-0.3280**	-0.2852*	-0.2882**	-0.2875**
	(0.1416)	(0.1417)	(0.1417)	(0.1465)	(0.1466)	(0.1466)
Het_{t+1}		-0.0155	-0.0130		-0.0612	-0.0581
		(0.0149)	(0.0168)		(0.0381)	(0.0393)
yellowplayer * Het_{t+1}			-0.0393			-0.0462
			(0.138)			(0.144)
Constant	-1.0931***	-1.0891***	-1.0896***	-0.9726***	-0.8986***	-0.8995***
	(0.0340)	(0.0342)	(0.0342)	(0.0919)	(0.0336)	(0.0337)
Controls included	yes	yes	yes	yes	yes	yes
RE-Model				yes	yes	yes
Obs.	39604	39604	39604	39604	39604	39604
Pseudo-R^2	0.2373	0.2373	0.2373			
McKelvey & Zavoina's R^2				0.3575	0.3586	0.3586

- Notes: The dependent variable is *starting11*.
- This table includes player-match-level data (all players).
- Robust standard errors clustered on the match level in parentheses, * p<0.1, ** p<0.05, *** p<0.01.

Table 2: Self-restriction and the starting line-up decision.

4.3 The decision to manipulate the card system

In this section, we seek to provide evidence of strategic self-restriction in a setting that demands anticipating behavior. In detail, we examine whether heterogeneity in the current and upcoming rounds influences the probability of receiving a critical yellow card followed by a suspension. Because every fifth yellow card triggers a ban in just the subsequent match, the self-restriction decision in *t+1* must be made in t. That is, Hypothesis 1 indicates that a team self-restricts in terms of a player seeking a deliberate fifth yellow card in t if the difference in abilities in *t+1* is sufficiently strong. Hypothesis 2 suggests that self-restriction does *not* occur if the heterogeneity in *t+2* is sufficiently strong.

Only fielded players are at risk of being cautioned, so it is necessary to restrict our sample. After excluding players without any playing time, the sample covers 32,003 observations.

The dependent variable (*yellow*) is a dummy that indicates whether or not a player i has received a yellow card in match t. We suggest the following empirical model:

$$\begin{aligned} yellow_{i,t} =& \beta_0 + \beta_1 yellowplayer_{i,t} + \beta_2 Het_{i,t} + \beta_3 Het_{i,t+1} + \beta_4 Het_{i,t+2} \\ &+ \beta_5 yellowplayer_{i,t} \cdot Het_{i,t+1} + \beta_6 yellowplayer_{i,t} \cdot Het_{i,t+2} + \gamma' \mathbf{X} + \alpha_i + \varepsilon_{i,t}, \end{aligned} \quad (6)$$

which closely resembles Equation (5) in Section 4.2. However, in addition to the introduction of round *t+2*, the number of accumulated yellow cards per season and a player's position also are included as game- and player-specific controls in **X**. Table 3 presents the descriptive statistics.

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
yellow	32003	0.142	0.350	0	1
Het_t	32003	0.285	0.214	0	0.864
Het_{t+1}	32003	0.287	0.215	0	0.864
Het_{t+2}	30852	0.286	0.214	0	0.864
yellowplayer	32003	0.093	0.291	0	1
% of minutes played	32003	0.609	0.295	0	1
sumYellow (season)	32003	3.581	1.849	0	11
away	32003	0.501	0.500	0	1
attendance	32003	43.486	16.899	13.5	81.359
derby	32003	0.043	0.203	0	1
matchday	32003	19.519	8.093	6	33
centre back	32003	0.205	0.403	0	1
left/right defender	32003	0.173	0.378	0	1
left/right/central midfield	32003	0.126	0.332	0	1
offensive midfield	32003	0.064	0.245	0	1
left/right wing player	32003	0.126	0.332	0	1
forward	32003	0.190	0.392	0	1

Table 3: Descriptive Statistics - Yellow Card

The results from the probit regressions are in Table 4.[21] Our main finding is that the estimated β_5 differs significantly from 0, so the probability of reaching the five yellow card limit increases with the heterogeneity of the next round's competition (fifth column), in accordance with our proposition 1. In addition, the robust result clearly points to anticipating behavior among competitors in a multi-stage tournament setting.[22]

Despite our set of controls, these findings do not establish direct proof of strategic behavior. Thus we note that if players provoke yellow card suspensions on purpose, we would expect them to do so at the end of a game, to minimize the risk of a second yellow card within the same game. We observe exactly this trend, such that when we split the sample according to playing time remaining (models 2 to 4 and 6 to 8), it becomes apparent that the estimated coefficients of the interaction term (fifth row) grow stronger toward the end of a 90 minute game.[23] We take this finding as evidence of deliberate and strategic behavior.

Moreover, we find an effect of the heterogeneity of the match after the next one on a *yellowplayer*'s probability to get booked. This effect emerges in our preferred model 5, though the somewhat inconsistent results across estimations suggest some need for caution, rather than about interpreting it as clear evidence in support of Hypothesis 2.

Finally, the theoretical model emphasizes the importance of the difference in the weighted ability of the opponent in match $t + 2$ on the decision to self-restrict, so we ran additional estimations (beyond the baseline estimation of $\delta = 1$) and assign different weights (ranging from 1 to 5) to the winning probability of the competitor in match $t + 2$ when constructing our

[21]Because Het_{t+2} is not available for the final two games of each season, the regressions are based on 30,852 observations only.

[22]Robustness checks with logic or LPM estimates can be found in Tables C.2 and C.3 in Appendix C. The main results hold.

[23]We used a pairwise *Chow test* to ensure that the coefficient's variation is systematic.

	Model 1 all	**Model 2** > 45 min	**Model 3** > 60 min	**Model 4** > 75 min	**Model 5** all	**Model 6** > 45 min	**Model 7** > 60 min	**Model 8** > 75 min
yellowplayer	-0.0979	-0.1317*	-0.1531*	-0.1834*	-0.1795***	-0.1847**	-0.1902**	-0.2030*
	(0.0672)	(0.0753)	(0.0852)	(0.1086)	(0.0680)	(0.0792)	(0.0865)	(0.1048)
Het_t	-0.2522***	-0.1987***	-0.2645***	-0.3293***	-0.2371***	-0.1785***	-0.2514***	-0.3297***
	(0.0504)	(0.0581)	(0.0627)	(0.0733)	(0.0515)	(0.0584)	(0.0637)	(0.0757)
Het_{t+1}	-0.1353***	-0.1465***	-0.1034*	-0.0690	-0.1048**	-0.1184**	-0.0912	-0.0693
	(0.0400)	(0.0476)	(0.0529)	(0.0610)	(0.0472)	(0.0533)	(0.0573)	(0.0670)
Het_{t+2}	0.0558	0.0410	0.0609	0.0642	0.0872*	0.0635	0.0734	0.0662
	(0.0433)	(0.0517)	(0.0582)	(0.0668)	(0.0477)	(0.0539)	(0.0584)	(0.0689)
$yellowplayer * Het_{t+1}$	0.4427***	0.4543**	0.5031**	0.5588*	0.4198***	0.4237**	0.4892***	0.5495**
	(0.1699)	(0.2027)	(0.2257)	(0.2896)	(0.1514)	(0.1740)	(0.1882)	(0.2233)
$yellowplayer * Het_{t+2}$	-0.3082*	-0.2656	-0.2050	-0.1945	-0.3493**	-0.3122*	-0.2368	-0.2232
	(0.1589)	(0.1804)	(0.1929)	(0.2287)	(0.1591)	(0.1857)	(0.2015)	(0.2442)
Constant	-1.2576***	-1.3199***	-1.4314***	-1.6474***	-1.2083***	-1.1402***	-1.2430***	-1.4660***
	(0.0889)	(0.1064)	(0.1184)	(0.1402)	(0.1133)	(0.0685)	(0.0720)	(0.0823)
Controls included	yes	yes	yes	yes	yes	yes	yes	yes
RE-Model					yes	yes	yes	yes
Obs.	30852	29260	28559	27686	30852	29260	28559	27686
Pseudo-R^2	0.0180	0.0141	0.0127	0.0122				
McKelvey & Zavoina's R^2					0.0403	0.0293	0.0248	0.0214

- Notes: The dependent variable is: *yellow*
- Player-match-level data (all fielded players).
- Since Models 2-4 and 6-8 refer to the last 45/30/15 minutes of a match, observations of bookings in prior minutes are excluded.
- Robust standard errors are clustered on match level in parentheses, * p<0.1, ** p<0.05, *** p<0.01.

Table 4: The effect of heterogeneity on yellow card suspensions (Probit estimates).

Variable	Model 1	Model 5
$\hat{\beta}_6(\delta = 1)$	-0.308 (0.159)	-0.349 (0.159)
$\hat{\beta}_6(\delta = 2)$	-0.197 (0.079)	-0.198 (0.081)
$\hat{\beta}_6(\delta = 3)$	-0.117 (0.053)	-0.112 (0.053)
$\hat{\beta}_6(\delta = 4)$	-0.088 (0.040)	-0.082 (0.040)
$\hat{\beta}_6(\delta = 5)$	-0.072 (0.032)	-0.067 (0.033)

Notes: Standard errors in parentheses.

Table 5: Estimated β_6 with different weights.

heterogeneity measure Het_{t+2}. Thus we can check its impact on the decision to self-restrict by receiving a critical fifth yellow card in t. Table 5 shows how $\hat{\beta}_6$ varies when δ increases (models 1 and 5). Although the heterogeneity of the game in $t + 2$ has a negative and statistically significant impact across all weights, the effect size decreases.

5 Conclusion and Implications

This article provides evidence of strategic investment decisions in anticipation of the future need for resources. This empirical analysis of German Bundesliga soccer supports the implications of the theoretical model, indicating that heterogeneity in contests decreases the willingness to invest resources. Players therefore self-restrict in current and future competition, as long as the heterogeneity between players is sufficiently strong. When players are at risk of a ban, they tend to be excluded from the starting eleven when the contest is unbalanced. Second, players are more likely to receive a crucial yellow card that triggers a ban if the subsequent contest (which they will miss due to the ban) is sufficiently unbalanced. Third but not as pronounced, players also tend to receive a crucial yellow card if the second to next game (for which they return from their ban) is rather homogeneous.

These findings reflect a discouragement effect, such that effort provision in contests falls short of what might be expected, considering the prize at stake, due to lopsided competition. This effect prompts reduced effort in anticipation of the outcome of the current contest, but our results further indicate that competitors plan ahead, resulting in strategically lowered investments with a view to future contests. According to our findings, competitors not only adapt their resource investment to the strength of their current opponent, but also anticipate the heterogeneity of future contests and their respective needs for resources.

Our results have important policy implications as well. In multi-stage contests with multiple contestants, heterogeneity is detrimental to mediocre contestants. The best contestants can save resources against the weakest contestants without substantially lowering their chances of winning. Correspondingly, the weakest contestants save their resources in matches when playing the strongest opponents, because their ex ante chances of winning are very low. Abilities, and thus winning probabilities, are more balanced against mediocre teams, such that either the strongest nor the weakest contestants conserve resources and instead meet them at full strength. Mediocre contestants cannot afford to self-restrict in any contest without risking a negative impact on the outcome. Thus, the structure of the contest and the possibility to self-restrict for the next competition creates a disadvantage for mediocre teams and has a stabilizing effect on the top-level hierarchy. The schedule itself could be disadvantageous to some competitors too, especially if they face an opponent that most recently played in a heterogeneous match-up (Krumer and Lechner, 2018).

Tournament organizers could counteract these effects by making self-restriction advantages less pronounced. First, in the case of soccer, they could change the rules, such that a a crucial fifth yellow card would lead to a randomly drawn ban sometimes in the next five matches. Then the strategic element of self-restriction would diminish, because contestants would not know in advance whether they would lose their valuable resources at exactly the moment they do not need it. Second, schedule imbalances can be lowered if the round-robin format repeats multiple times (as in the Bundesliga, with two rounds of round-robin). Rearranging the sequence of games after each round then would prevent teams from facing opponents that profit from self-restriction because their last match was unbalanced.

References

Altmann, S., Falk, A., and Wibral, M. (2012). Promotions and incentives: The case of multistage elimination tournaments. *Journal of Labor Economics*, 30(1):149–174.

Baik, K. H. (1994). Effort levels in contests with two asymmetric players. *Southern Economic Journal*, 61(2):367–378.

Balsdon, E., Fong, L., and Thayer, M. A. (2007). Corruption in college basketball? Evidence of tanking in postseason conference tournaments. *Journal of Sports Economics*, 8(1):19–38.

Berger, J. and Nieken, P. (2016). Heterogeneous contestants and the intensity of tournaments: An empirical investigation. *Journal of Sports Economics*, 17(7):631–660.

Brown, J. and Minor, D. B. (2014). Selecting the best? Spillover and shadows in elimination tournaments. *Management Science*, 60(12):3087–3102.

Cornes, R. and Hartley, R. (2005). Asymmetric contests with general technologies. *Economic Theory*, 26(4):923–946.

Dagaev, D. and Zubanov, A. (2017). Round-robin tournaments with limited resources. National Research University Higher School of Economics Working Paper 171/EC/2017.

Dechenaux, E., Kovenock, D., and Sheremeta, R. M. (2015). A survey of experimental research on contests, all-pay auctions and tournaments. *Experimental Economics*, 18(4):609–669.

Deutscher, C., Frick, B., Gürtler, O., and Prinz, J. (2013). Sabotage in tournaments with heterogeneous contestants: Empirical evidence from the soccer pitch. *Scandinavian Journal of Economics*, 115(4):1138–1157.

Deutscher, C., Frick, B., and Ötting, M. (2018). Betting market inefficiencies are short-lived in german professional football. *Applied Economics*, 50(30):3240–3246.

Ehrenberg, R. G. and Bognanno, M. L. (1990). The incentive effects of tournaments revisited: Evidence from the European PGA tour. *Industrial & Labor Relations Review*, 43(3):74–88.

Fédération Internationale de Football Association (FIFA) (2017). *Laws of the Game 2017/18*. Zurich, Switzerland: Fédération Internationale de Football Association, Zurich, Switzerland.

Harbaugh, R. and Klumpp, T. (2005). Early round upsets and championship blowouts. *Economic Inquiry*, 43(2):316–329.

Hart, E., Avrahami, J., Kareev, Y., and Todd, P. M. (2015). Investing even in uneven contests: Effects of asymmetry on investment in contests. *Journal of Behavioral Decision Making*, 28(4):395–409.

Hill, B. (2018). Shadow and spillover effects of competition in NBA playoffs. *Journal of Sports Economics*, 19(8):1067–1092.

Konrad, K. A. (2009). *Strategy and dynamics in contests.* Oxford University Press. Oxford, UK.

Krumer, A. and Lechner, M. (2018). Midweek effect on soccer performance: Evidence from the german bundesliga. *Economic Inquiry*, 56(1):193–207.

Krumer, A., Megidish, R., and Sela, A. (2017a). First-mover advantage in round-robin tournaments. *Social Choice and Welfare*, 48(3):633–658.

Krumer, A., Megidish, R., and Sela, A. (2017b). The optimal design of round-robin tournaments with three players. Economics Working Paper Series No 1713, University of St. Gallen, School of Economics and Political Science.

Lackner, M., Stracke, R., Sunde, U., and Winter-Ebmer, R. (2015). Are competitors forward looking in strategic interactions? Evidence from the field. IZA Discussion Paper No. 9564.

Laica, C., Lauber, A., and Sahm, M. (2017). Sequential round-robin tournaments with multiple prizes. CESifo Working Paper 6685.

Leininger, W. (1993). More efficient rent-seeking – a Münchhausen solution. *Public Choice*, 75(1):43–62.

Malkiel, B. G. and Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25(2):383–417.

March, C. and Sahm, M. (2017). Asymmetric discouragement in asymmetric contests. *Economics Letters*, 151:23 – 27.

March, C. and Sahm, M. (2018). Contests as selection mechanisms: The impact of risk aversion. *Journal of Economic Behavior & Organization*, 150:114 – 131.

Raya, J. M. (2015). The effect of strategic resting in professional cycling: Evidence from the Tour de France and the Vuelta a España. *European Sport Management Quarterly*, 15(3):323–342.

Rosen, S. (1986). Prizes and incentives in elimination tournaments. *American Economic Review*, 76(4):701–715.

Ryvkin, D. (2013). Heterogeneity of players and aggregate effort in contests. *Journal of Economics & Management Strategy*, 22(4):728–743.

Sahm, M. (2019). Are sequential round-robin tournaments discriminatory? *Journal of Public Economic Theory*, 21:44–61.

Schneemann, S. and Deutscher, C. (2017). Intermediate information, loss aversion, and effort: Empirical evidence. *Economic Inquiry*, 55(4):1759–1770.

Spann, M. and Skiera, B. (2009). Sports forecasting: A comparison of the forecast accuracy of prediction markets, betting odds and tipsters. *Journal of Forecasting*, 28(1):55–72.

Stein, W. E. (2002). Asymmetric rent-seeking with more than two contestants. *Public Choice*, 113(3):325–336.

Sunde, U. (2009). Heterogeneity and performance in tournaments: A test for incentive effects using professional tennis data. *Applied Economics*, 41(25):3199–3208.

Szymanski, S. (2003). The economic design of sporting contests. *Journal of Economic Literature*, 41(4):1137–1187.

VanDerwerken, D. N., Rothert, J., and Nguelifack, B. M. (2018). Does the threat of suspension curb dangerous behavior in soccer? A case study from the Premier League. *Journal of Sports Economics*, 19(6):3240–3246.

Appendix

A Backward Induction of Game Γ

This appendix illustrates how to solve game Γ (see Figure 1) by backward induction, considering the parameters $(\theta_1, \theta_2, r, q) = (1, 1, 1/2, 3/4)$.

Case $(\mathbf{R}, \mathbf{R})$

Suppose that both player 1 and player 2 have chosen R in stage 1 and thus restrict themselves in the first match.

4th stage: player 2 vs player 3

In his second match, player 2 is unrestricted, and thus $\theta_2^k = \theta_3^k = 1$ in all nodes $k \in \{A, B, C, C'\}$.

In node A, player 2 wins the first match, and player 3 wins the second match. Thus $w_2^A = w_3^A = 1$, $\ell_2^A = \ell_3^A = 1/2$, which yields

$$x_2^A = x_3^A = \frac{(1-\frac{1}{2})^2 \cdot (1-\frac{1}{2})}{[(1-\frac{1}{2}) + (1-\frac{1}{2})]^2} = \frac{1}{8},$$

$p_2^A = p_3^A = \frac{1}{2}$, and $E_2^A = E_3^A = \frac{1}{2} + \frac{1}{8} = \frac{5}{8}$ by equations (2) – (4).

In node B, player 2 wins the first match, and player 1 wins the second match. Thus $w_2^B = 1$, $w_3^B = 1/2$, $\ell_2^B = 1/2$ and $\ell_3^B = 0$, which yields $x_2^B = x_3^B = 1/8$, $p_2^B = p_3^B = 1/2$, $E_2^B = 5/8$, and $E_3^B = 1/8$.

In node C, player 1 wins the first match, and player 3 wins the second match. Thus $w_2^C = 1/2$, $w_3^C = 1$, $\ell_2^C = 0$, and $\ell_3^C = 1/2$, which yields $x_2^C = x_3^C = 1/8$, $p_2^C = p_3^C = 1/2$, $E_2^C = 1/8$, and $E_3^C = 5/8$.

In node C', player 1 wins the first match and the second match. Thus $w_2^{C'} = w_3^{C'} = 1/2$ and $\ell_2^{C'} = \ell_3^{C'} = 0$, which yields $x_2^{C'} = x_3^{C'} = 1/8$, $p_2^{C'} = p_3^{C'} = 1/2$, and $E_2^{C'} = E_3^{C'} = 1/8$.

3rd stage: player 1 vs player 3

In his second match, player 1 is unrestricted, and thus $\theta_1^k = \theta_3^k = 1$ in both nodes $k \in \{D, E\}$.

In node D, player 2 wins the first match. Thus, $w_1^D = \frac{1}{2}p_2^B + \frac{1}{2}p_3^B = \frac{1}{2}$, $w_3^D = E_3^A = \frac{5}{8}$, $\ell_1^D = 0$, and $\ell_3^D = E_3^B = \frac{1}{8}$, which yields $x_1^D = x_3^D = 1/8$, $p_1^D = p_3^D = 1/2$, $E_1^D = 1/8$, and $E_3^D = 1/4$.

In node E, player 1 wins the first match. Thus, $w_1^E = 1$, $w_3^E = E_3^C = \frac{5}{8}$, $\ell_1^E = \frac{1}{2}p_2^C + \frac{1}{2}p_3^C = \frac{1}{2}$, and $\ell_3^E = E_3^{C'} = \frac{1}{8}$, which yields $x_1^E = x_3^E = 1/8$, $p_1^E = p_3^E = 1/2$, $E_1^E = 5/8$, and $E_3^E = 1/4$.

2nd stage: player 1 vs player 2

In their first match, players 1 and 2 are restricted, and thus $\theta_1^F = \theta_3^F = \frac{1}{2}$ in node F. Moreover $w_1^F = E_1^E = \frac{5}{8}$, $w_2^F = p_1^D E_2^B + p_3^D E_2^A = \frac{5}{8}$, $\ell_1^F = E_1^D = \frac{1}{8}$, and $\ell_2^F = p_1^E E_2^{C'} + p_3^E E_2^C = \frac{1}{8}$, which yields $x_1^F = x_2^F = 1/8$, $p_1^F = p_2^F = 1/2$, $E_1^F = 1/4$, and $E_2^F = 1/4$. Moreover, player 3's expected payoff equals $E_3^F = p_1^F E_3^E + p_2^F E_3^D = 1/4$.

Cases $(\mathbf{NR}, \mathbf{R})$, $(\mathbf{R}, \mathbf{NR})$, and $(\mathbf{NR}, \mathbf{NR})$

Using analogous procedures, we calculate the players' expected payoffs in cases in which only one or none of them self-restricts in the first match.

1st stage: decision on self-restriction

The results for the expected payoffs of players 1 and 2 are in Table A.1 (rounded to two decimal places). For the parameters under consideration, NR is a dominant strategy for both players, and thus, (NR, NR) is the only Nash-equilibrium in the first stage of game Γ.

		Player 2	
		R	NR
Player 1	R	0.25 0.25	0.18 0.31
	NR	0.31 0.18	0.22 0.22

Table A.1: Game Matrix of Stage 1

B Additional Figures

Figure B.1: Relative abilities θ_1 and θ_2 compared with $\theta_3 = 1$

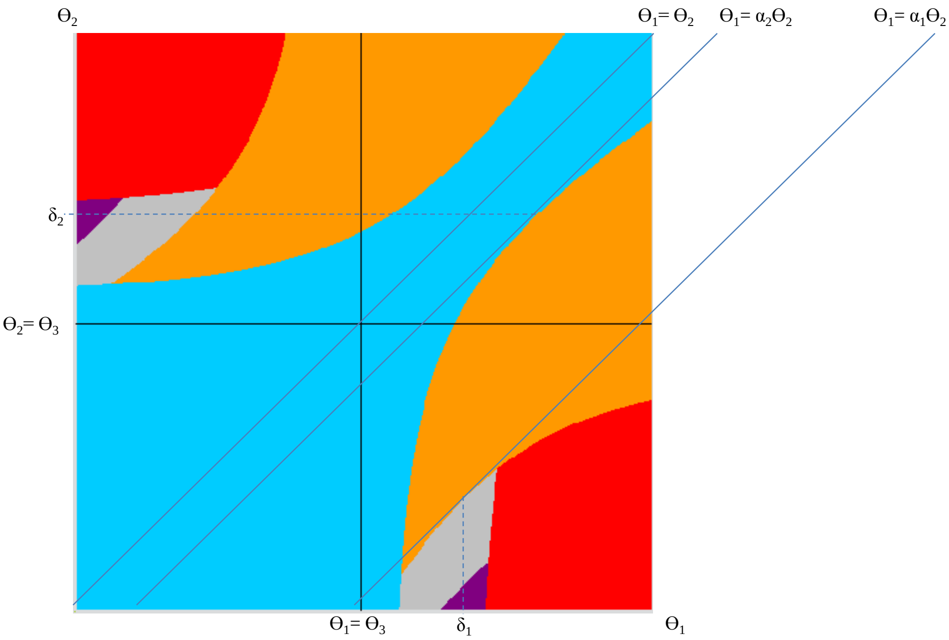
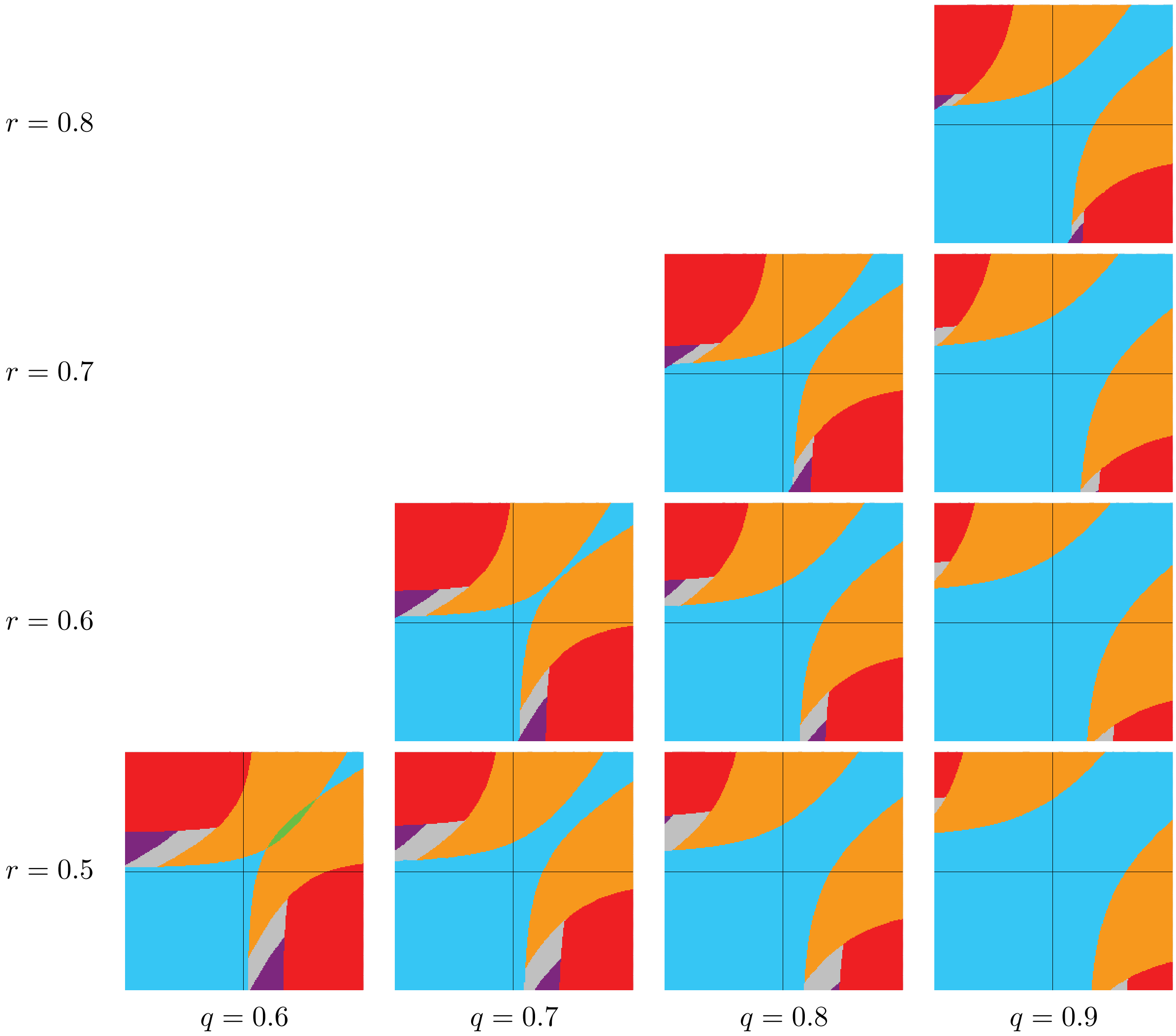


Figure B.2: Comparative Statics w.r.t. r and q



Note: This figure illustrates the effects of variations in r and q. On the horizontal (vertical) axis of each diagram θ_1 (θ_2) increases from 0.1 to 10 in steps of 2% for normalized $\theta_3 = 1$. The vertical (horizontal) bold line represents $\theta_1 = 1$ ($\theta_2 = 1$). Each color represents a certain type of equilibrium behavior of players 1 and 2 in the first stage:

- blue $\hat{=}$ no player chooses R,
- orange $\hat{=}$ only the weaker player chooses R,
- red $\hat{=}$ both players choose R,
- violet $\hat{=}$ only the stronger player chooses R,
- grey $\hat{=}$ both players choose R with positive probability (equilibrium in mixed strategies),
- green $\hat{=}$ either only the weaker player chooses R or only the stronger player chooses R or both players choose R with positive probability (multiple equilibria).

C Additional Tables

	Logit	Logit	LPM	LPM
% of minutes played	4.2331***	4.0560***	0.8137***	0.7518***
	(0.0544)	(0.0585)	(0.0069)	(0.0080)
yellowplayer	0.3057***	0.2505**	0.0480***	0.0331**
	(0.1057)	(0.1142)	(0.0142)	(0.0157)
$absoluteHet_t\,(odds)$	0.1263***	0.0684	0.0226***	0.0132
	(0.0371)	(0.0741)	(0.0063)	(0.0114)
$yellowplayer * absHet_t\,(odds)$	-0.5455**	-0.4981*	-0.0723**	-0.0622*
	(0.2479)	(0.2592)	(0.0350)	(0.0366)
$absoluteHet_{t+1}\,(odds)$	-0.0085	-0.1005	0.0020	-0.0115
	(0.0291)	(0.0678)	(0.0048)	(0.0104)
$yellowplayer * absHet_{t+1}\,(odds)$	-0.0518	-0.0538	-0.0018	-0.0017
	(0.2448)	(0.2558)	(0.0333)	(0.0358)
Constant	-1.8280***	-1.5173***	0.1334***	0.2079***
	(0.0591)	(0.0584)	(0.0095)	(0.0084)
Controls included	yes	yes	yes	yes
RE-Model		yes		yes
Obs.	39604	39604	39604	39604
R^2			0.2858	0.2855
Pseudo-R^2	0.2363			
McKelvey & Zavoina's R^2		0.3325		

- Notes: The dependent variable is *starting11*.
- Player-match-level data (all players).
- Robust standard errors are in parentheses, * p<0.1, ** p<0.05, *** p<0.01.

Table C.1: Heterogeneity and Self-Restriction (Logit and LPM estimates)

	Model 1 all	**Model 2** > 45 min	**Model 3** > 60 min	**Model 4** > 75 min	**Model 5** all	**Model 6** > 45 min	**Model 7** > 60 min	**Model 8** > 75 min
yellowplayer	-0.1748	-0.2480*	-0.3003*	-0.3948	-0.3207***	-0.3497**	-0.3737**	-0.4359*
	(0.1236)	(0.1476)	(0.1752)	(0.2469)	(0.1243)	(0.1539)	(0.1756)	(0.2316)
Het_t	-0.4633***	-0.3865***	-0.5374***	-0.7209***	-0.4354***	-0.3481***	-0.5143***	-0.7217***
	(0.0934)	(0.1143)	(0.1284)	(0.1628)	(0.0947)	(0.1139)	(0.1295)	(0.1665)
Het_{t+1}	-0.2450***	-0.2831***	-0.2073*	-0.1559	-0.1966**	-0.2367**	-0.1856	-0.1549
	(0.0741)	(0.0938)	(0.1089)	(0.1351)	(0.0868)	(0.1039)	(0.1164)	(0.1468)
Het_{t+2}	0.1067	0.0821	0.1334	0.1528	0.1616*	0.1279	0.1607	0.1605
	(0.0799)	(0.1021)	(0.1197)	(0.1491)	(0.0874)	(0.1050)	(0.1186)	(0.1511)
yellowplayer * Het_{t+1}	0.8066**	0.8793**	1.0096**	1.2551*	0.7758***	0.8331**	0.9836***	1.2256**
	(0.3154)	(0.4069)	(0.4710)	(0.6660)	(0.2760)	(0.3391)	(0.3814)	(0.4911)
yellowplayer * Het_{t+2}	-0.5735**	-0.5387	-0.4500	-0.5008	-0.6592**	-0.6288*	-0.5108	-0.5570
	(0.2914)	(0.3528)	(0.3907)	(0.4995)	(0.2926)	(0.3627)	(0.4082)	(0.5371)
Constant	-2.1434***	-2.2678***	-2.4979***	-2.9461***	-2.0670***	-1.9454***	-2.1441***	-2.5843***
	(0.1626)	(0.2062)	(0.2390)	(0.3027)	(0.2079)	(0.1326)	(0.1450)	(0.1787)
Controls included	yes	yes	yes	yes	yes	yes	yes	yes
RE-Model					yes	yes	yes	yes
Obs.	30852	29260	28559	27686	30852	29260	28559	27686
Pseudo-R^2	0.0178	0.0139	0.0126	0.0120				
McKelvey & Zavoina's R^2					0.0424	0.0347	0.0322	0.0334

- Notes: The dependent variable is *yellow*.
- Player-match-level data (all fielded players).
- Since Models 2-4 and 6-8 refer to the last 45/30/15 minutes of a match, observations of bookings in prior minutes are excluded.
- Robust standard errors clustered on match level in parentheses, * p<0.1, ** p<0.05, *** p<0.01.

Table C.2: The effect of heterogeneity on yellow card suspensions (Logit estimates).

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**	**Model 7**	**Model 8**
	all	> 45 min	> 60 min	> 75 min	all	> 45 min	> 60 min	> 75 min
yellowplayer	-0.0211	-0.0213*	-0.0201*	-0.0160*	-0.0211	-0.0213*	-0.0239**	-0.0160*
	(0.0151)	(0.0124)	(0.0115)	(0.0095)	(0.0148)	(0.0129)	(0.0117)	(0.0093)
Het_t	-0.0544***	-0.0321***	-0.0350***	-0.0288***	-0.0544***	-0.0321***	-0.0335***	-0.0288***
	(0.0109)	(0.0095)	(0.0084)	(0.0065)	(0.0108)	(0.0094)	(0.0085)	(0.0068)
Het_{t+1}	-0.0290***	-0.0240***	-0.0139*	-0.0066	-0.0290***	-0.0240***	-0.0124	-0.0066
	(0.0087)	(0.0079)	(0.0072)	(0.0057)	(0.0098)	(0.0085)	(0.0077)	(0.0061)
Het_{t+2}	0.0143	0.0082	0.0102	0.0073	0.0143	0.0082	0.0118	0.0073
	(0.0096)	(0.0088)	(0.0082)	(0.0064)	(0.0102)	(0.0088)	(0.0080)	(0.0064)
yellowplayer * Het_{t+1}	0.1003**	0.0765**	0.0703**	0.0537*	0.1003***	0.0765***	0.0690***	0.0537**
	(0.0409)	(0.0370)	(0.0348)	(0.0311)	(0.0339)	(0.0296)	(0.0267)	(0.0215)
yellowplayer * Het_{t+2}	-0.0720**	-0.0466	-0.0318	-0.0213	-0.0720**	-0.0466	-0.0342	-0.0213
	(0.0342)	(0.0283)	(0.0252)	(0.0191)	(0.0344)	(0.0299)	(0.0269)	(0.0216)
Constant	0.1094***	0.0991***	0.0802***	0.0514***	0.1094***	0.0991***	0.0840***	0.0514***
	(0.0199)	(0.0182)	(0.0168)	(0.0135)	(0.0215)	(0.0187)	(0.0176)	(0.0136)
Controls included	yes	yes	yes	yes	yes	yes	yes	yes
RE-Model					yes	yes	yes	yes
Obs.	30852	29260	28559	27686	30852	29260	28559	27686
R^2	0.015	0.009	0.007	0.004	0.0146	0.0088	0.0066	0.0044

- Notes: The dependent variable is *yellow*.
- Player-match-level data (all fielded players).
- Since Models 2-4 and 6-8 refer to the last 45/30/15 minutes of a match, observations of bookings in prior minutes are excluded.
- Robust standard errors clustered on match level in parentheses, * p<0.1, ** p<0.05, *** p<0.01.

Table C.3: The effect of heterogeneity on yellow card suspensions (LPM estimates).