

Are Competitors Forward Looking in Strategic Interactions? Evidence from the Field

by

Mario LACKNER
Rudi STRACKE
Uwe SUNDE
Rudolf WINTER-EBMER

Working Paper No. 1509

November 2015

Corresponding author mario.lackner@jku.at



JOHANNES KEPLER
UNIVERSITY LINZ

Christian Doppler Laboratory
Aging, Health and the Labor Market
cdecon.jku.at

Johannes Kepler University
Department of Economics
Altenberger Strasse 69
4040 Linz, Austria

Are Competitors Forward Looking in Strategic Interactions? Evidence from the Field*

Mario Lackner^a, Rudi Stracke^b, Uwe Sunde^b, and Rudolf Winter-Ebmer^{a, c}

^aDepartment of Economics, University of Linz (JKU)

^bDepartment of Economics, University of Munich (LMU)

^cInstitute for Advanced Studies (IHS), Vienna

November 6, 2015

Abstract

This paper investigates empirically whether decision makers are forward looking in dynamic strategic interactions. In particular, we test whether decision makers in multi-stage tournaments take heterogeneity induced changes of continuation values and the ability of their immediate opponent into account when choosing effort. Using data from professional and semi-professional basketball tournaments, we find that effort is negatively affected by the ability of the current opponent, consistent with the theoretical prediction and previous evidence. More importantly, the results indicate that the expected relative strength in future interactions does affect behavior in earlier stages, which provides support for the ‘standard’ view that decision makers are forward looking in dynamic strategic interactions.

JEL Classification: D84; D90; M51; J33

Keywords: Promotion tournament; multi-stage contest; elimination; forward-looking behavior; heterogeneity

*Corresponding author: Uwe Sunde, Seminar for Population Economics, Department of Economics, Geschwister-Scholl-Platz 1, D-80539 München; Tel.:+49-89-2180-1280; Fax:+49-89-2180 17834; Email: uwe.sunde@lmu.de. Thanks to Alex Ahammer for help with the data and Josh Angrist, Jen Brown, Florian Englmaier, Dylan Minor, Steve Stillman and seminar participants in Ohlstadt, Laax, Vienna, Istanbul and Mannheim for helpful comments.

1 Introduction

Dynamic strategic interactions characterize many situations of economic decision making. Modern macro models typically assume that decision makers are rational and forward looking, incorporating the consequences of future interactions in their current decisions. Likewise, many contexts studied in microeconomics involve strategic and dynamic components. Obviously, this is motivated by the prevalence of such decision environments in reality. One prominent example are workplace interactions, where workers interact with their employer as well as with their co-workers repeatedly, with important implications for work incentives. A prime example of how these dynamic strategic interactions work and how they are incorporated in the design of employment relations are promotion tournaments. Promotion tournaments are a common way of creating incentives for workers to exert effort on their job, but conceptually very similar tournaments are also observed in many other contexts, for instance during electoral competitions in majority voting systems, or in multi-stage procurement tournaments with shortlists.

The power and usefulness of tournament models for gaining a better understanding of behavior in such strategic interactions has been demonstrated by the seminal work of Lazear and Rosen (1981). While this benchmark model is essentially static, subsequent contributions highlight the importance of forward-looking behavior for optimal effort choices in tournaments. In particular, Waldman (1984) assumes that competing workers anticipate the signal value of a promotion for future wage negotiations, while Rosen (1986) argues that the continuation value of future promotion possibilities is an important determinant of current effort choices. Both approaches share the idea that the next promotion is not the ultimate goal, but rather a means to an end of forward-looking agents, namely either the prerequisite for future promotions to even more attractive positions within the same organization, or a signal observable by competing organizations that then allows workers to demand higher wages. Apart from the few exceptions discussed below, however, there is little to no evidence whether decision makers are indeed forward looking in dynamic strategic interactions as typically assumed in theoretical models, even though the existence and extent of forward-looking behavior is essential for many practical purposes.

This paper provides evidence on this question by investigating how current heterogeneity and the expected relative strength in future interactions affect behavior in dynamic strategic interactions. A canonical multi-stage pairwise elimination tournament model predicts the well-known adverse incentive effect of heterogeneity according to which greater heterogeneity in a given interaction reduces effort of favorites and underdogs – independent of whether decision makers are forward looking or not. The consideration of multiple stages delivers the additional prediction that the expected relative strength in future interactions has a positive effect on effort in a given (current) interaction if and only if decision makers are forward looking and take the continuation value into account. However, whether and to what extent this is the case has not been tested systematically in the existing literature. The dynamic incentive effect due to forward-looking behavior is driven by the prospects of the higher expected winning odds in the next round that a competitor tries to take advantage of by increasing current effort. Using data from professional and semi-professional basketball tournaments, we then test whether or not tournament participants are forward looking. In particular, we consider the playoffs of NBA and NCAA basketball tournaments, which provide an ideal setting for this purpose. On the one hand, these tournaments involve considerably large stakes, and, on the other hand, they provide precise information about all required elements, such as heterogeneity, effort and outcomes. Moreover, the differences between the NBA and NCAA tournament rules allow cross-validating the empirical findings.

Our findings support the view that decision makers are indeed forward looking. In particular, the results show that, everything else equal, the expectation of a weaker future opponent increases effort in the current match. The empirical results also suggest that the strategic aspect implied by the expectation of a weaker future opponent becomes more important the shorter the remaining tournament and the more pivotal the current game is, i.e. the closer the final interaction in the tournament. In addition, we find that effort is negatively affected by the ability of the current opponent, consistent with the theoretical prediction and previous evidence based on static interactions.

This paper contributes to, and extends, the existing literature on tournaments in several ways. First, we contribute to existing work that investigates the effects of heterogeneity on behavior in static tournaments, see, e.g., Baik (1994). Empirical work that

tests these predictions has typically relied on data from lab experiments (see, e.g., Bull et al., 1987, Chen et al., 2011), and only few papers have investigated the role of heterogeneity using field data from sports, see, e.g., Sunde (2009), Brown (2011), and Berger and Nieken (2014). Our work complements this literature by showing that static incentive effects of heterogeneity continue to matter in each stage of a multi-stage tournament structure. Second, this paper provides a direct test of a necessary condition of models with mechanisms that are based on dynamic implications, including models by Waldman (1984) and Rosen (1986), according to which the prospect of better outside options or of future promotions affects effort provision in earlier stages of the tournament. In particular, we show that heterogeneity induced changes of continuation values affect effort choices of tournament participants, in line with the standard assumption that decision makers are indeed forward looking in dynamic strategic interactions. Thereby, this paper complements recent work that has investigated related issues in the lab, see, e.g., Altmann, Falk, and Wibral (2012) or Stracke and Sunde (2015). Evidence from the field is scarce, however. A notable exception is work by Delfgaauw, Dur, Non, and Verbeke (2015) who implement a two-stage elimination tournament in a field experiment. They focus on the effect of variations in the structure of prizes and in the importance of noise *within* a given stage, not on the effect of continuation values *across* stages as done here. To our knowledge, the only study that indirectly accounts for dynamic incentive effect in multi-stage tournaments using match outcomes rather than effort choices is by Brown and Minor (2014). They focus on selection properties of multi-stage tournaments and investigate the influence of past effort and the strength of the expected future competitor on the probability that the stronger player wins in a given tournament interaction. While our data allow us to replicate their results, our analysis extends theirs by explicitly investigating the implications of current and future heterogeneity on effort of all tournament participants, thereby opening the black box of how observed outcomes are achieved.

This paper also relates to a very recent literature that looks at the role of expectations in terms of reference points for behavior (see, e.g., Bartling et al., 2015). While our paper shares the focus on expectations, their intention is showing how behavior changes when actual performance or outcomes falls short of previous expectations, whereas this paper is interested in the implications of expected future heterogeneity on current behavior.

The remainder of the paper is structured as follows. The next section presents a simple prototype model to derive testable hypotheses. Section 3 describes the data, the measures of heterogeneity and effort, and the empirical strategy. Section 4 presents the main results and discusses the findings from robustness checks. Section 5 concludes.

2 The Model

2.1 A Simple Tournament with Two Heterogeneous Agents

Consider a tournament with two risk neutral agents i and j who simultaneously choose effort e_i . Assume that the value of winning the tournament is R and that losing has value zero. The linear cost of effort function satisfies $c(e_i) = \frac{e_i}{a_i}$, where the parameter a_i is the ability of agent i and $a_1 \geq a_2$ holds. Under the assumption that each agent takes the effort of the competitor as given, the expected payoff function of agent i is defined as

$$\Pi_i(e_i, e_j) = p_i(e_i, e_j) \cdot R - \frac{e_i}{a_i}, j \neq i. \quad (1)$$

Assume that winning probabilities $p_i(e_i, e_j)$ are determined by a function of the ratio of efforts. In particular, winning probabilities are given by the Tullock (1980) contest technology according to which

$$p_i(e_i, e_j) = \frac{f(e_i)}{f(e_i) + f(e_j)}, \quad (2)$$

where $f(x) = x^r$. The parameter $r > 0$ measures the discriminatory power of the contest technology. Intuitively, differences in the effort choices of the agents have minor effects on winning probabilities if r is close to zero, while even minor differences in effort choices have strong effects on winning probabilities for large values of r .

Equilibrium efforts are determined by mutually best responses, and in the following we restrict attention to pure strategy equilibria for expositional purposes.¹ From the first-order optimality conditions of both agents, we obtain

$$\hat{e}_i = r \cdot a_i \cdot \frac{\theta}{(1 + \theta)^2} \cdot R \quad \text{and} \quad \hat{e}_j = r \cdot a_j \cdot \frac{\theta}{(1 + \theta)^2} \cdot R \quad (3)$$

as equilibrium effort of agents i and j , respectively, in the static tournament, where $\theta = \left(\frac{a_i}{a_j}\right)^r$ is an index of the degree of heterogeneity among agents i and j and reflects agent i 's *relative* ability vis-a-vis agent j . Equation (3) indicates that equilibrium effort for a given prize R is increasing in the discriminatory power r and in the respective agent i 's own (absolute) ability a_i , and decreasing in the degree of heterogeneity θ among the two agents. To illustrate why θ measures heterogeneity, note that θ equals one when agents are homogeneous (since the ratio $\frac{a_i}{a_j} = 1$), and increases monotonically as differences in ability between the agents increase – the ratio of abilities is thus directly related to the degree of heterogeneity. The parameter θ also depends on the discriminatory power r , however, since the outcome of the tournament does not only depend on heterogeneity, but also on the relative importance of effort and luck, which is parameterized by r . In particular, a given difference in effort levels due to heterogeneity does not matter much if the discriminatory power r is low, such that the outcome is mainly determined by luck. The same difference in effort levels has a much stronger and more direct effect on winning if the discriminatory power is high, however, since differences in efforts have a strong effect on the probability to win in this case. By accounting for the discriminatory power r , the parameter θ is thus a sufficient statistic for winning probabilities of agents i and j . Inserting equilibrium efforts in (2) immediately reveals that this is the case.

Equilibrium efforts \hat{e}_i and \hat{e}_j determine expected equilibrium payoffs. Inserting the respective expressions for equilibrium effort in (1) and defining $\Pi_i(\hat{e}_i, \hat{e}_j) \equiv \hat{\Pi}_i$ delivers

$$\hat{\Pi}_i = \frac{\theta^2 + (1 - r)\theta}{(1 + \theta)^2} \cdot R \quad \text{and} \quad \hat{\Pi}_j = \frac{1 + (1 - r)\theta}{(1 + \theta)^2} \cdot R. \quad (4)$$

¹For this contest technology and linear cost of effort, pure strategy equilibria exist as long as $r \leq 1 + \frac{a_i}{a_j}$, see Nti (1999) for details. For the subsequent analysis, this restriction is inconsequential.

An inspection of these expressions reveals that the expected payoff is increasing in θ for agent i , and decreasing in θ for agent j . Intuitively what matters for the difference in the expected payoff is relative ability, not heterogeneity. Given that we assume $a_i \geq a_j$ and define heterogeneity as $\theta = (a_i/a_j)^r$, higher values of θ imply that the relative ability advantage of agent i increases, while the same change in θ implies that the relative ability of agent j deteriorates.

2.2 The Value of Future Competition in Multi-Stage Tournaments

Consider next a multi-stage tournament where two forward looking agents with abilities $a_i \geq a_j$ compete for the prize R_{now} and the right to participate in the next stage of the tournament. Let the value of winning the next stage of the tournament be R_{fut} , and assume that both agents know the identity and thus the ability a_{fut} of their future opponent. In such a setting, current effort of forward-looking agents not only depends on the own ability and heterogeneity as in equation (3), but also on the ability of the (expected) future opponent. Intuitively, participating in future interactions of the tournament becomes more attractive the weaker the future opponent, since this implies that the chances to win against this opponent and thus to receive the prize R_{fut} are high. At the same time, it is comparably unattractive to participate in future interactions of the tournament if the future opponent is strong. The value of participation in the next stage of the tournament, the continuation value, is defined by the expected payoff for agent i (or analogously j) when competing against a future opponent with ability a_{fut} . If the identity of the future opponent is uncertain, since the parallel interaction between the other agents k and l on the same stage of the tournament is not yet decided, we define the ability of the future opponent a_{fut} as follows:

$$a_{\text{fut}} = p_k \cdot a_k + (1 - p_k) \cdot a_l , \quad (5)$$

where p_k is the probability that agent k wins against agent l in the parallel interaction. Moreover, we define $p_k = \frac{a_k}{a_k + a_l}$, since winning probabilities are determined by the degree of heterogeneity of competing agents in the parallel match.²

When defining the relative ability of agent i in the competition with the (expected) future opponent as $\kappa_i = \left(\frac{a_i}{a_{\text{fut}}}\right)^r$, it immediately follows from equation (4) that the continuation value CV_i of agent i is formally defined as

$$\text{CV}_i^*(\kappa_i) = \frac{[\kappa_i]^2 + (1-r)\kappa_i}{(1+\kappa_i)^2} \cdot R_{\text{fut}} \quad (6)$$

In line with intuition, $\text{CV}_i^*(\kappa_i)$ is strictly increasing in relative future ability κ_i , since chances to win against the future opponent and thus to receive the prize are strictly increasing in relative ability. Equilibrium effort by agent i in any non-final stage of a multi-stage tournament is thus formally defined as follows:

$$e_i^* = r \cdot a_i \cdot \frac{\theta}{(1+\theta)^2} \cdot [R_{\text{now}} + \text{CV}_i^*(\kappa_i)] \quad (7)$$

where θ measures the degree of heterogeneity between agents i and j in the current interaction, while κ_i accounts for the relative ability of agent i in the next stage of the tournament.³

2.3 Theoretical Predictions

According to condition (7), effort on any non-final stage of a multi-stage pairwise elimination tournament depends on the discriminatory power, on the agent's own ability, on within-interaction heterogeneity, and, if the agent is forward looking, on the agent's relative ability in future interactions. Note that this insight does not depend on particular simplifying assumptions of the model considered so far.⁴ In particular, the predictions are qualitatively the same in a Lazear and Rosen (1981) tournament with additive noise and

²This assumption avoids feedback effects of heterogeneity across interactions and simplifies the subsequent analysis. As we show in Appendix A.2, this simplification is not relevant for the results.

³For simplicity, we abstract from the impact of differences between $\text{CV}_i^*(\kappa_1)$ and $\text{CV}_j^*(\kappa_2)$ on heterogeneity in the current interaction. When accounting for the fact that continuation values differ across agents i and j , the actual degree of heterogeneity is given by $\theta = \left(\frac{a_i(R_{\text{now}} + \text{CV}_i^*)}{a_j(R_{\text{now}} + \text{CV}_j^*)}\right)^r$. As we show in Appendix A.2, this simplifying assumption is uncritical for the results.

⁴Details are provided in Appendix A.2.

quadratic effort cost.⁵ The only difference in this case is that measures of heterogeneity and relative abilities depend on the difference rather than the ratio of abilities.

The comparative static predictions of the model are of particular interest for the purpose of this paper. Consider first the effect of current heterogeneity on effort provision in the current stage. It turns out that heterogeneity reduces effort provision of competing agents in pairwise interactions, and that the corresponding negative effect is stronger for the stronger agent. In what follows, we use the term ‘favorite’ ($i = \text{F}$) for the stronger and the term ‘underdog’ ($i = \text{U}$) for the weaker agent:

Result 1 (Current Heterogeneity). *A higher degree of heterogeneity $\theta = \left(\frac{a_F}{a_U}\right)^r$ between a favorite and an underdog in any stage of a multi-stage pairwise elimination tournament implies*

- (a) *a lower level of absolute effort of favorites and underdogs on the respective stage of the tournament;*
- (b) *a more pronounced negative effect on effort of favorites than of underdogs.*

Proof. See Appendix [A.1](#) □

Part (a) of this result is well known from previous work – see, e.g., Baik (1994). Intuitively, the fact that (marginal) effort cost are higher for underdogs than for favorites – independent of the opponent’s ability – implies that underdogs provide less effort than favorites. Given that underdogs face a stronger opponent as heterogeneity goes up, their reaction to an increase in heterogeneity is to reduce their effort below the effort they would provide in a homogeneous tournament. Favorites, on the other hand, also provide less effort in more heterogeneous interactions since they anticipate the lower effort by underdogs – they slack off to avoid costly effort, as their chances to win are higher than in a homogeneous interaction even when providing less effort. Part (b) of [Result 1](#) follows from two observations. First, equation [\(7\)](#) shows that the (relative) reduction of effort in response to heterogeneity is the same for favorites and underdogs.⁶ Second, the marginal cost for effort, which is inversely related to ability, is lower for favorites than

⁵Details are provided in Appendix [A.3](#).

⁶The same is true for the simple tournament, see equation [\(3\)](#).

for underdogs. This implies, in turn, that favorites provide more effort than underdogs – ability has a direct positive and linear effect on effort, and an indirect positive effect through continuation values – such that the negative effect of heterogeneity is also more pronounced for favorites than for underdogs in absolute terms.

The second effect of interest concerns the question how the ability of future opponents affects effort choices in earlier stages of a multi-stage tournament. More precisely, we investigate how the future relative ability of an agent affects effort choices on any non-final stage of a multi-stage pairwise elimination tournament. Our model predicts that effort of forward looking agents is increasing in future relative ability, whereas effort choices are independent of relative future ability if agents are entirely focussing on the immediate consequence of their action:

Result 2 (Future Relative Ability). *Let a_{fut} be the expected ability of the future opponent and $i = F, U$. Effort choices of the favorite and the underdog on any non-final stage are then increasing in future relative ability $\kappa_i = \left(\frac{a_i}{a_{fut}}\right)^r$ if and only if agents are forward looking.*

Proof. See Appendix [A.1](#) □

Intuitively, the immediate reward for winning a non-final stage R_{now} is independent of future relative ability, while the continuation value is determined by future relative ability. Given that forward looking agents take both the immediate reward for winning R_{now} and the continuation value of reaching the next stage into account, they are predicted to react to changes in the continuation value. Changes in the continuation value are irrelevant for agents who entirely focus on the immediate consequences of their effort choices, however, since they fail to take the continuation value into account when choosing effort.

3 Empirical Implementation

3.1 Data

For the empirical analysis we use sports data from professional basketball. Sports data have several advantages. In particular, there is a well-defined tournament situation, and

ability as well as outcomes are readily observable. In addition, data from professional basketball have features that are particularly useful for the purpose of this paper. First, the basketball data can be used to construct a measure for effort, which is of crucial importance because it allows us to open the black box of how outcomes are determined. Second, we have access to competition performance data for each team from the regular season preceding the playoffs, i.e., the elimination tournament. This allows us to incorporate a rich set of controls that varies for each team and season. Finally, we have access to complementary data from another tournament structure, which allows us to check the validity of the results using two independent and, due to their different structure, complementary data sets.

The main part of the empirical analysis is based on data from championship tournaments in the National Basketball Association (NBA)⁷. During the regular season, a round-robin tournament is conducted in four separate leagues. After the end of the regular season, the best teams participate in a pairwise elimination tournament (the “playoffs”). Every game of basketball in the NBA is a tournament covering 48 minutes of net playing time, split up in 4 quarters. In the regular season, as well as during the playoff tournament every single game must have a winner. In case there is a tie at the end of regular time, a potentially infinite number of overtime periods of five minutes follow until a winner is determined. The empirical analysis is based on information about the full time outcome of a game. A dummy variable is used to control for games that are decided in overtime.

For organizational reasons, two separate elimination tournaments take place (the Eastern and Western conference), and the winners of each tournament only meet in a best-of-7 series of games, the “Finals”, to determine the champion.⁸ The empirical analysis focuses on the pairwise elimination tournaments that take place at the level of the conferences. Each tournament is structured in four stages; on each stage, two teams compete in best-of-5 (in stage 1 before the 2003/2004 NBA season) and otherwise best-of-7 match-ups, i.e., the winner is determined as the team that wins three or four games against the respective opponent team. The data contain information for 2,199 individual playoff games from the 1983/84 through 2013/14 seasons. While we restrict attention to

⁷All data were collected online from www.basketball-reference.com using historical boxscores and team statistics.

⁸The structure is illustrated by the playoffs of the 2013 season, see Figure 8 in Appendix B.

data from the playoff phase of the season to capture the tournament structure. Performance data from the regular season is used as background information regarding ability and other team-specific characteristics.

As a complementary data set, we use information from the National Collegiate Athletic Association (NCAA). The NCAA playoff tournaments consist of five stages leading up to the final. Each stage involves only one game between two teams (i.e., a best-of-1 winning rule), which takes place on neutral ground. We have access to data for 10 seasons from 2003 through 2013, covering a total of 682 games.

3.2 Measuring Ability

The data cover detailed information on final outcomes of games, final scores, and various statistics of team performance. One of the key variables for the purpose of this paper is a measure of ability for all teams that allows computing heterogeneity on the current stage as well as constructing a measure of relative ability expected on the next stage. In particular, ability does not only directly influence the level of effort – see equation (7) for details – but is also necessary to determine the present degree of heterogeneity (denoted θ in the theoretical model) as well as the future relative ability measure (represented by κ). A naive statistic of the number of scores or the share of games won in the regular season is readily available, but might be misleading. Due to regional separation of the leagues into four conferences, different teams face different schedules and pools of competitors, raising problems of comparability of scores in the regular season. We therefore employ the Simple Rating System (SRS) as calculated by the web-site www.basketball-reference.com.⁹ This rating system is based on the regular season point differential for each team, weighted by the team’s strength of schedule.¹⁰ This delivers a comparable measure of ability arising from the performance during the regular season that leads up to the elimination tournament of the playoffs.

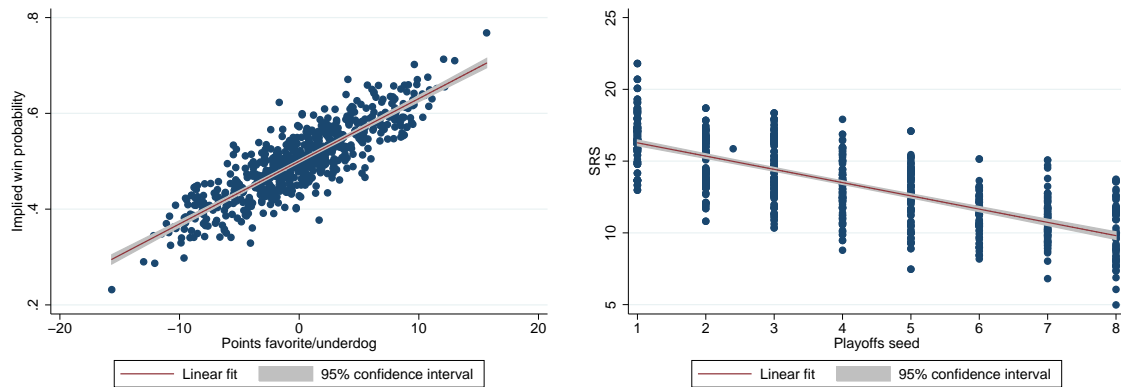
To validate the SRS-based ability measure, we compare it to betting odds and the seeds in the playoffs. Reassuringly, the winning probabilities calculated from using this

⁹This web-page is providing free sports data and calculates advanced statistics for multiple professional sports leagues. The site is run by Sports Reference LLC, <http://www.sports-reference.com>.

¹⁰As the SRS is negative for some team-years, we re-scale it as $SRS_{rescaled} = SRS + 10$ to get only positive values and thus allow for a calculation of Tullock win probabilities.

ability measure are highly correlated with betting odds as is illustrated in the left panel of Figure 1.¹¹ The SRS-based ability measure is also highly correlated with the ability-based tournament seeds in the playoffs as shown in the right panel of Figure 1. In the empirical analysis we use as a baseline a constant ability indicator, which is exclusively based on information from the regular season preceding the respective elimination tournament, and which is not influenced by the performance during the playoffs.

Figure 1: Validation of the Measure of Ability Using the SRS



Left panel: Implied win probabilities (calculated using the SRS in a Tullock-type probability function) against mean point differentials per matchup, prediction by betting markets for rounds 1-3 in NBA seasons 1991 through 2013. $N = 644$

Right panel: SRS against tournament seed for rounds 1-3 in NBA seasons 1984 through 2014. $N = 868$

To compute an empirical measure of heterogeneity (labelled $heterogeneity_t$ in the empirical analysis) on a given stage of the tournament that corresponds to θ_t in the theoretical model, let a_{it} be the SRS score of a favorite and an underdog team ($i = \{F, U\}$), respectively, who compete on stage t . Current heterogeneity on stage t is then computed as

$$heterogeneity_t = \frac{a_{Ft}}{a_{Ut}}, \quad (8)$$

where $heterogeneity_t \geq 1$ holds due to the definition of favorite and underdog teams (since $a_F \geq a_U$). The advantage of computing the measure of heterogeneity for favorites and underdogs is that the effects of changes in $heterogeneity_t$ work in the same direction for both teams. Using instead a symmetric measure of heterogeneity in terms of relative

¹¹We obtained betting odds data from www.covers.com.

ability – relating own ability and opponent ability as implied by θ_t – would imply that changes in relative ability would have opposite effects on the effort provision of favorites and underdogs. Intuitively, heterogeneity increases if the relative ability of the favorite increases, while heterogeneity decreases if the relative ability of the underdog improves. This would complicate the empirical analysis and its interpretation unnecessarily. The empirical measure of heterogeneity, $heterogeneity_t$, is thus directly linked to the measure θ_t used in the theoretical analysis. The only modification is that this measure implicitly assumes a discriminatory power of $r = 1$, which corresponds to the classic specification of a Tullock contest technology. As a robustness analysis discussed below, we relax this assumption and estimate r from the data to construct an alternative measure of heterogeneity along the lines of the theoretical model.

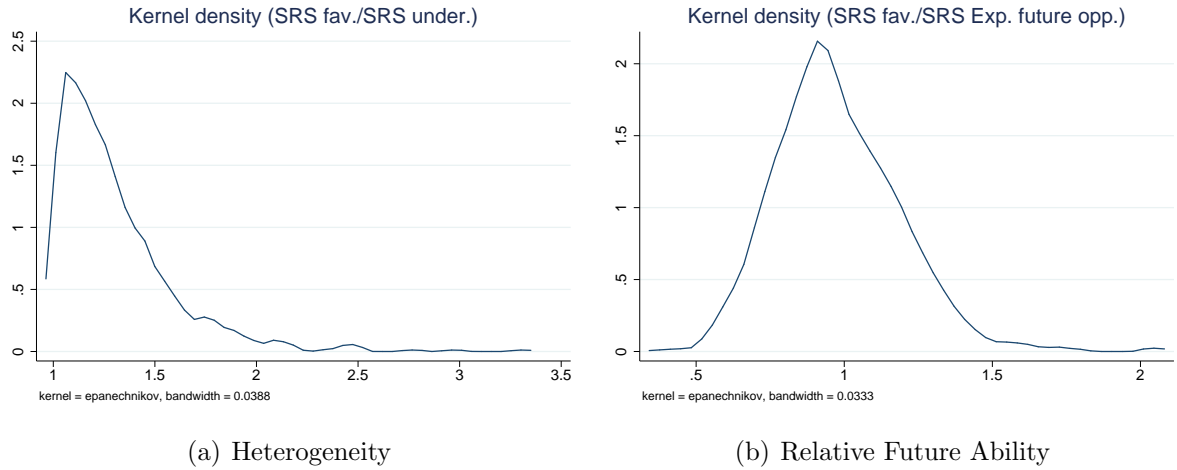
Panel (a) of Figure 2 plots the empirical kernel density of the respective heterogeneity measure and shows that games with comparatively low degrees of heterogeneity are most frequently observed in the data. High degrees of heterogeneity where $\theta_t > 1.5$ are rather the exception than the rule.

The empirical measure for relative ability of team $i \in \{\mathbf{F}, \mathbf{U}\}$ in stage $t + 1$ of the tournament against the opponent team with ability $a_{\mathbf{fut}, t+1}$ – which is labelled $E_t[rel. ability_{t+1}]$ in the following and corresponds to $\kappa_{i, t+1}$ in the theoretical analysis – is constructed in analogy to the index of heterogeneity on the current stage t . In particular, the measure of expected relative ability in future interactions, based on information available at stage t , is constructed as

$$E_t [rel. ability_{t+1}] = \frac{a_i}{E_t [a_{\mathbf{fut}, t+1}]} . \quad (9)$$

Importantly, this measure accounts for future relative ability rather than future heterogeneity, since the effect of future heterogeneity on the continuation value is different for teams that are the favorite or the underdog, respectively, in the future interaction. The expected ability of the opponent on stage $t + 1$, $a_{\mathbf{fut}, t+1}$ can be constructed in different ways, depending on the assumptions made on the expectation formation process. For simplicity and transparency, in the construction of the empirical measures we restrict attention to competitors on the immediately next stage only. As baseline measure, we consider a convex combination of the ability levels of the two teams in the parallel match-up, the winner

Figure 2: Distribution of Heterogeneity θ and Relative Future Ability κ



Notes: The figures plot kernel densities of the heterogeneity measure and the measure of expected relative ability in the next round, respectively, using the raw data used in the estimation exercises. $N = 434$

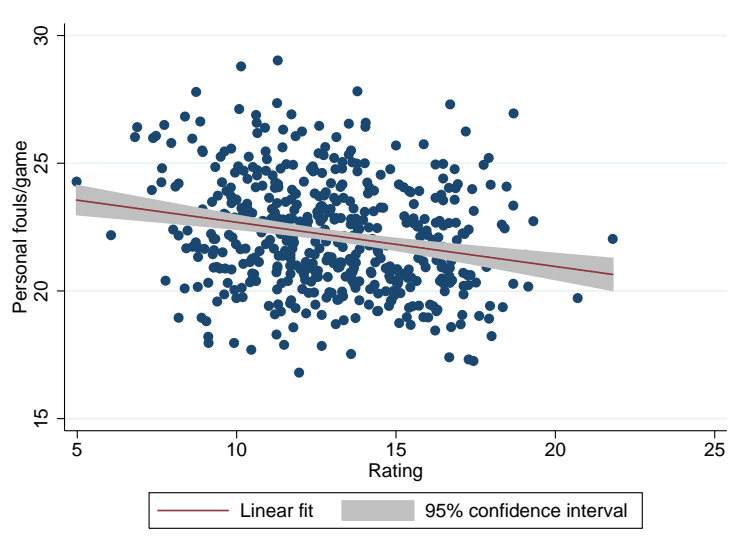
of whom will be the competitor on stage $t + 1$ of the team under consideration at the current stage t . The ability levels are weighted by the respective winning probabilities. The reason is that, in many instances, the identity of the future opponent team is not known yet when decisions are made on the current stage. The structure of the competition implies, however, that teams competing in stage t know that they will compete against the winner of the parallel stage- t match on the next stage. Given that abilities of those teams competing in this parallel match are observable, teams can form expectations about the probabilities with which they face any one of the two potential opponent teams in the next stage. In the empirical analysis, we then compute the expected ability of the opponent i in stage $t + 1$ as in condition (5). In the robustness section, we also present the results under different assumptions regarding the expected relative ability on the next stage. In particular, we consider measures that are based exclusively on the relative ability of the favorite in the parallel match-up, or adjusted measures that incorporate for each game if the competitor is already known at the time of the game, and replace the convex combination by the ability of the actual next competitor. Panel (b) of Figure 2 plots the empirical kernel density of future relative ability for current favorites.

3.3 Measuring Effort

Whereas one can think of many indicators for final outcomes or performance, constructing a measure of effort of a team or of total effort per game by two teams is not entirely straightforward. In the empirical analysis of this paper, we use the number of personal fouls that a team is called for as a proxy for effort of a team. A personal foul in basketball is defined as a breach of the rules that regulate the legal or illegal form of personal contact between players. This mostly involves attempts to prevent the opposing team from scoring. These fouls are thus called defensive fouls. Much less frequent are offensive fouls that occur during an offensive phase when an illegal scoring attempt is observed. Hence, both types of fouls measure an attempt to change the course of a game in order to win. Consequently, the number of fouls in a game is a direct indicator for the intensity of the competition and, thus, for the effort of the respective team. In principle, fouls measure how intense the defender attacks his opponent, how physically close he is in coverage, which may sometimes result in a personal foul. Notice that for this to hold it is not necessary to assume that teams explicitly decide to foul their opponent. Rather, it is presumably more likely that players try to avoid fouls in most instances, but are still more likely to foul the opponent when defending intensively. In that sense, personal fouls are an almost natural outcome of an intense game with close physical contact. The higher the intensity, the higher the probability that a foul is inadvertently committed and called. The intensity of play by a particular team, hence, should be highly correlated with the effort provided.

This measure corresponds closely to a measure of team effort since personal fouls can be committed by all players of a team. The number of personal fouls is mostly influenced by a team's own effort, and a good proxy for how much effort is put on defending the opponent. More personal fouls correspond to a more physical, thus more tiring, style of play. Finally, one might be worried that there are different reasons to commit a foul than an intensive defense, such as intentional fouls for tactical reasons. One situation where intentional fouling might be an optimal strategy for a team is when the members of the team get tired and are not able anymore to defend without committing a foul. However, even if fouls are committed by mistake, or tactically, because players are worn out, this provides a useful indicator of the physical intensity of the match for a given team as

Figure 3: Personal Fouls per Game in Regular Season vs. Ability Measure (SRS)

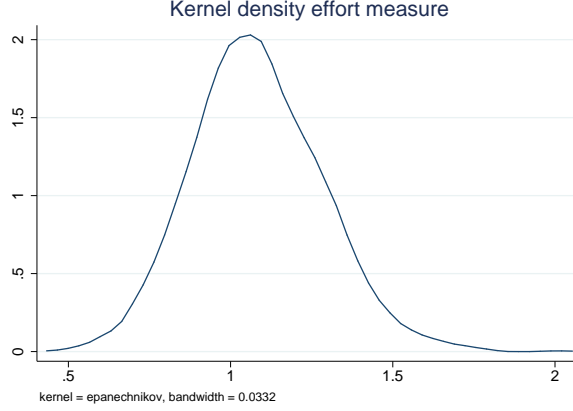


Notes: Average number of personal fouls per game in regular season against SRS. $N = 496$

fatigue is a clear indicator that the intensity – and therefore effort – has been on a high level during the game. Fouling may also be an optimal strategy to stop the clock at the end of a very close game when a team is trailing. While it is neither direct offensive or defensive effort, it is still an attempt to exert every available strategy to win the game, i.e., effort. Consequently, personal fouls appear as the most suitable of the available effort measures for the present paper.

A potential problem with this measure is that the number of personal fouls a team is called for might be affected by the ability of a team in a counterintuitive manner when thinking about effort. In particular, the theoretical model predicts that effort is increasing in ability, while it seems more reasonable to expect that more able teams are also more able to avoid fouls when defending intensively. This is also what we observe in the long-term regular season data, the pairwise correlation coefficient associated with the numbers shown in Figure (3) is -0.22 and highly significant. To control for a team's ability to avoid fouls as well as for the style of play of a particular team – the defense of some teams might explicitly decide to foul more often independent of the opponent team – our preferred proxy for team effort is the number of fouls per game relative to the average

Figure 4: Distribution of Effort Measure



Notes: The figures plot the kernel density of the effort measure, using the raw data used in the estimation exercises. $N = 434$

per-game number of fouls the team has committed in the regular season preceding the playoffs.¹² This measure is formally defined as

$$effort_{i,k,t} = \frac{\text{number of fouls in playoff game}_{i,k,t}}{\left(\frac{\sum \text{fouls regular season}_{i,k}}{\text{number of games}_k} \right)} \quad (10)$$

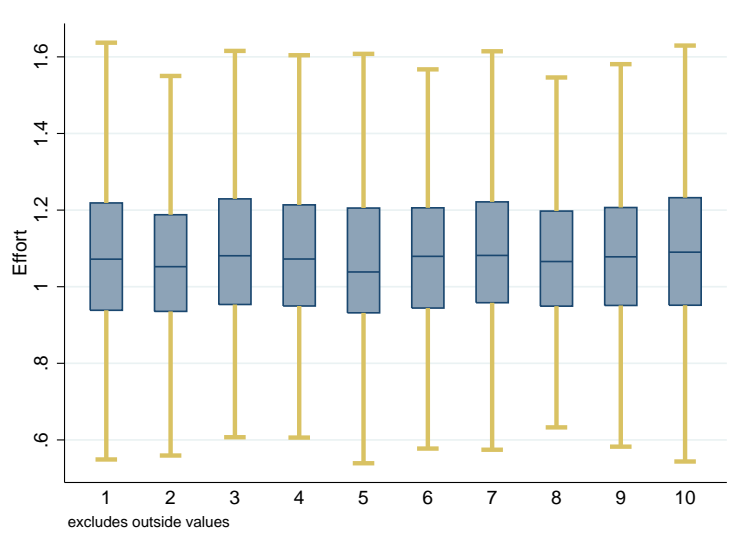
for team i in year k in stage t of the tournament. Figure 4 plots the kernel density of this effort measure and shows that teams tend to foul more often in the playoffs than in the regular season on average.

When considering the effort measure in terms of personal fouls conditional on the ability during the playoffs, however, the data reveal no clear systematic relationship. Figure 5 presents a box plot of the effort distribution during the playoffs by ability decile. There is no systematic relationship, in contrast to what one would expect if effort were misspecified in terms of capturing merely some aspect of ability (or incompetence), where one would expect a systematic relationship between effort and ability throughout the tournament.

Teams are also likely to adjust to the style of play of their opponents, as the system of the opponent's coach might differ across teams. It is therefore necessary to control for

¹²Relating the number of fouls per game relative to the average number of fouls during the regular season plays a similar role as season-team fixed effects by accounting for different styles of play, team compositions, coaching styles, etc., in a given season.

Figure 5: Effort Measure vs. Deciles of Ability Rating (SRS)



Notes: Effort, as defined in (10) for deciles of ability measured by the SRS for rounds 1–3 for NBA seasons 1984–2014. Outliers are excluded. $N = 4398$

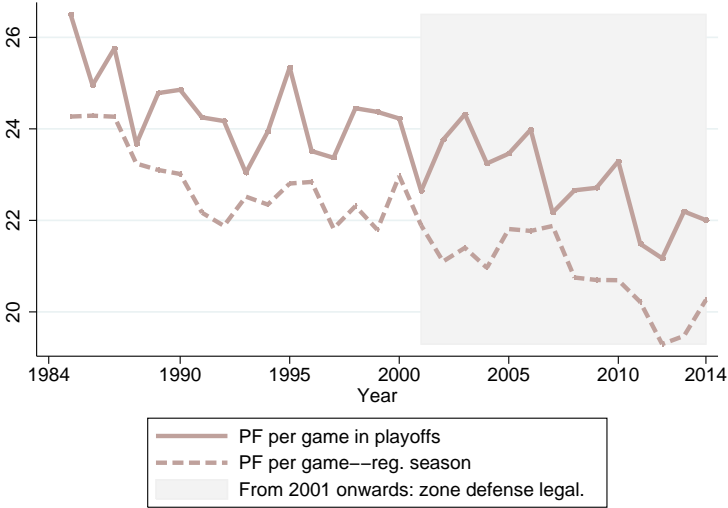
various indicators describing how opponents usually play the game. The nature of the data allows us to use regular season statistics in order to control for team-specific playing styles. One crucial measure for the likelihood that a foul is called is the style of offense the opponent plays in terms of the distance from which they make their shot attempts. An increased number of shot attempts from behind the three-point line by the opponent could therefore reduce the number of fouls, independent of effort provided.¹³ In order to account for this, we control for the opponent team’s number of three-point attempts in the regular season. Finally, we also control for the speed of the opponent’s play, as it seems quite possible that a team will commit fewer fouls if the opposing team is slowing down the game in order to reduce certain disadvantages. A good proxy for how fast an opponent plays is the absolute number of field goal attempts per minute, which is used in the form of the regular season average as a control variable.

Another concern for the effort measure could be the presence of a so-called zone defense. A zone-defense is a style of defense which is less physical and relies more on optimal positioning in space and, thus, might produce fewer fouls, independent of effort provided.

¹³The three-point line is a mark on the floor which separates the area where a successful basket counts two points from the rest of a field where it is worth 3 points. The distance between the three-point line and the basket has been the subject of multiple changes since the founding of the NBA. See www.nba.com/analysis/rules_history.html.

One practical way of operating against an opposing zone defense is to concentrate more on distance shooting. The number of attempted three-point shots should increase in the presence of a zone defense, as these defenses can be better attacked from the distance and are known for leaving three-point shots open. Figure 6 plots the evolution of fouls over time. The change in background shade indicates the rule change in the 2000/01 NBA season, which made zone defense illegal.¹⁴ The figure suggests that the rule change had no effect on the average number of fouls, and the long time trend remained unaffected. In the empirical analysis we will control for season fixed effects, which should also take care of the time trend in the average number of personal fouls.

Figure 6: Personal Fouls During the Regular Season and During Playoff Tournaments



A similar approach to use data on fouls as proxy for effort has been used before by Berger and Nieken (2014) as well as by Deutscher, Frick, Gürtler, and Prinz (2013) in the context of handball and soccer, respectively. Both studies rely on regular season data, however, which implies that they cannot control for the usual style of play of a team, as is done in the construction of the effort measure by relating effort in a given game to average effort in the regular season. Another difference to their approach is that we do not distinguish between destructive and productive effort, as both should have an

¹⁴A zone defense is a form of defense where a player is defending a certain area rather than defending against an opposing player. A detailed overview on the evolution of rules regarding illegal defense is to be found at www.nba.com/analysis/rules_history.html.

identical positive influence on the probability to win.¹⁵ In light of this, and given the intensity and strictness of rules in basketball, we view fouls as an equally or even more appropriate measure in the context of this paper.

The discussion so far has provided several arguments for the suitability of the number of personal fouls as an effort indicator in the context of this paper. In what follows, we provide more direct evidence on the arguments made so far. Recall that we use fouls as a measure of effort, since fouls measure how intense the defender attacks his opponent and how physically close he is in coverage. This type of defensive effort may then sometimes – but not always – result in a personal foul. We would thus expect that higher defensive effort increases turnovers of the opponent team (i.e. stealing the ball) and thereby helps to prevent the opponent from scoring. Since this implies that intensively defending teams a higher share of ball possession, and since ball possession is necessary to score, we would also expect that intensively defending teams score more often.

Table 1 presents the corresponding regression results that document that higher effort, indeed, increases turnovers of the opponent team, both for underdogs and favorites. Moreover, effort reduces the points scored by the opponent team out of the field (net of free throws), suggesting that one additional foul (relative to the regular season average) reduces points of the opponent team by 0.50. The effect is slightly larger when considering only the sample of underdogs, and slightly lower when considering favorites, but the differences are not significant. Finally, more fouls are apparently successful in increasing the number of own points scored, maybe because higher turnover rates make quick advances on offense more profitable. Again, the effect appears to be slightly larger for underdogs than for favorites. The downside of more effort in terms of personal fouls, however, is a higher number of free throws for the opponent team. According to the point estimates, one additional foul leads to one additional point from free throws for the opponent team.

¹⁵Lazear (1989) analyses tournament situations where high incentives increase the effort of tournament participants, which is increasing the surplus of the tournament. On the other side, high incentives may also lead to more sabotage behavior, which will, in turn, reduce the surplus. In contrast to such a situation in a firm, defensive actions of a sports team should not be interpreted as sabotage behavior, because they do not reduce the surplus of the match. While, in principle, a surplus in a sports competition is difficult to define, it should be related to the suspense or thrill of the game for the spectators, which will in any case be positively related to the intensity of the game. In a similar vein, Bartling, Brandes, and Schunk (2015) find that teams react by the number of rule breaches, measured by the assignment of cards in soccer.

Table 1: Validation: The Effect of Effort on Outcomes

<i>Sample:</i>	Opponent's			Own
	Points ^a	Turnovers	Free throws ^b	Points
<i>Pooled</i>				
<i>effort</i>	-0.505*** (0.042)	0.053*** (0.015)	1.068*** (0.019)	0.349*** (0.042)
<i>N</i>	4398	4266	4398	4398
<i>R</i> ²	0.259	0.109	0.532	0.350
<i>Underdogs</i>				
<i>effort</i>	-0.571*** (0.060)	0.055** (0.021)	1.057*** (0.025)	0.385*** (0.057)
<i>N</i>	2199	2133	2199	2199
<i>R</i> ²	0.270	0.137	0.549	0.351
<i>Favorites</i>				
<i>effort</i>	-0.436*** (0.056)	0.047** (0.021)	1.103*** (0.028)	0.326*** (0.064)
<i>N</i>	2199	2133	2199	2199
<i>R</i> ²	0.262	0.114	0.533	0.346

Robust standard errors (clustered for individual playoff-series) in round parentheses. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively.

^a Total number of opponent's points scored from the field (without points from free throws).

^b Total number of opponent's free throws taken.

Since many statistics in basketball are influenced by ability and effort choices, one might consider various other statistics as proxies for effort provision. Goldman and Rao (2012) analyze effort in regular season NBA games in the context of pressure in home games, identifying defensive rebounds as defensive and offensive rebounds as offensive effort. Any rebound in basketball, however, is strongly affected by the efforts of both teams, rendering it problematic as measure of a given team's effort. Moreover, rebound rates by definition add up to one, since the more rebounds one team achieves the lower the rebound rate of the opponent. Being a share between zero and one, rebound rates cannot

be used as an effort measure, where both teams' effort is supposed to react.¹⁶ Another indicator is the number of blocked shots, which has similar deficiencies for the purposes of this paper. An alternative statistic that should be highly correlated with defensive effort, and therefore more closely corresponds to the theoretical conceptualization of effort is the number of steals. Steals are to a lesser degree influenced by both team's effort, but they are highly influenced by abilities. In sum, the number of personal fouls appears as the measure of effort that is suited best for the purposes of this paper.

3.4 Empirical Framework

In order to test the theoretical predictions about the implications of heterogeneity on the current stage, and of expected relative ability on the next stage, we use a simple reduced form linear estimation framework that allows to test the comparative static predictions of condition (7). In particular, we estimate the empirical model

$$effort_t = \beta_0 + \beta_1 heterogeneity_t + \beta_2 E_t [rel. ability_{t+1}] + \Omega' X + \epsilon, \quad (11)$$

where $heterogeneity_t$ is the empirical counterpart to θ_t and $rel. ability_{t+1}$ is the empirical counterpart to κ_{t+1} .¹⁷ The coefficient β_0 incorporates, among other things, the discriminatory power r that is constant across agents and stages, while β_1 measures the (average) effect of absolute ability on effort. Results 1 and 2 provide directed hypotheses about the expected signs of coefficient estimates β_2 and β_3 : First, β_2 – the effect of current heterogeneity θ_t on stage-t effort – is predicted to be negative. In addition, β_2 should be higher in absolute terms in the sub-sample of favorites than in the sub-sample of underdogs. Second, β_3 – the effect of future relative ability κ_{t+1} on stage-t effort – is predicted to be

¹⁶Goldman and Rao (2012) are interested in the effect of pressure in a home game on effort. In this case, rebound rates are appropriate, but the differential effect of crowd cheering on the home vs. the away team cannot be identified with this measure.

¹⁷Estimates for more flexible specification that derives directly from a log-linearized version of (7) with $R_{now} = 0$ deliver qualitatively very similar results. In particular, substituting condition (6) into (7), taking logs and simplifying under the assumption that $r = 1$ delivers

$$effort_t = \beta_0 + \beta_1 heterogeneity_t + \beta_2 E_t [rel. ability_{t+1}] + \gamma_1 (1 + heterogeneity_t) + \gamma_2 E_t (1 + rel. ability_{t+1}) + \Omega' X + \epsilon, \quad (12)$$

but due to the collinearity of the respective terms that include $heterogeneity$ and $rel. ability$, this specification delivers estimates that are less reliable and more difficult to interpret.

positive if agents are forward looking, and zero if agents entirely focus on the immediate consequences of their actions. When applying this to the data using the measure of effort defined in equation (10), there are two minor issues that need to be noticed. First, the estimation framework does not control for the absolute ability of a team, since the effort measure is normalized with respect to the regular season average, which already accounts for ability.¹⁸ Second, the empirical specification includes additional controls for the team’s and the opponent’s styles of play. These modifications deliver the following estimation equation. Specifically, the vector X controls for team-specific playing styles such as the opponent’s free throw percentage in the regular season, the opponent’s three-point percentage, the absolute number of own shot attempts, the opponent’s absolute number of shot attempts, the number of three-point shot attempts allowed and the percentage of successful three-point shots allowed – everything measured for the regular season preceding the playoffs. Moreover, X includes a variable counting the number of previous meetings in the preceding regular season, a dummy variable equal to 1 if the team i plays at home, a dummy equal to 1 if the series are decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies¹⁹ and overtime dummies.²⁰

Note that the unit of observation in the empirical analysis is a single game. The main hypothesis underlying the analysis concerns the impact of the expected strength of the future opponent on current effort: variations in the strength of the future opponent change the incentives of a team from the outset of the game, but not within the game. Hence, we are not interested in (minute-by-minute) dynamics within a game. The dynamics within a game depends on the effort – and success – of the two teams as the game proceeds and can be considered as noise with respect to our main hypothesis.

¹⁸Unreported results show that adding ability as an additional control variable delivers coefficient estimates for this variable that are always insignificant, as one would expect if ability is already controlled for by the normalization of the effort measure. Details available from the authors upon request.

¹⁹All standings are defined from the perspective of the observed team. A standing of, e.g., ‘1-0’ indicates that the current observation is in the second game with the observed team leading the playoff series by one game.

²⁰We also estimate a specification including the number of rest days and the travel distances between team locations. The results do change neither qualitatively nor quantitatively.

4 Results

4.1 Main Results

Table 2 presents the main results from estimating model (11). The results reveal that teams indeed adjust their effort in response to the current opponent’s ability: if the ratio of a team’s own rating over the current opponent’s rating increases, teams reduce their effort. Given that favorites and underdogs in a game may have different incentives and possibilities to react to variations in heterogeneity, we also present results for splitting the sample between favorites and underdogs. Regardless of the specification, the negative effect of heterogeneity is more pronounced for favorites than underdogs. These patterns are consistent with the predictions of Result 1. Regarding the key question of this paper about the implications of variations in the expected relative ability on the next stage of the tournament, the results reveal a positive effect on effort in the current round. Thus, teams exert *more* effort in response to standing a better chance of prevailing on the *next* stage of the tournament, consistent with the predictions of Result 2. This effect is significant and visible in all specifications. In particular, there are only minor, insignificant differences between favorites and underdogs.

The log-log specification implies a straightforward quantitative interpretation: if current relative strength is higher by one percent, current effort increases by 3 percent. The shadow of the future is even larger: if the expected future opponent’s relative strength is higher by one percent, current effort is reduced by 5 percent. In terms of relative strength of the effects, the effort effect of future relative ability is typically even larger than the effect of current heterogeneity.

4.2 Robustness and Additional Results

In this subsection, we investigate the robustness of the results as well as additional implications regarding the strength of the effects in different contexts. Table 3 presents the results from estimating model (11), using three alternative specifications of the measures of heterogeneity and relative ability on the next stage. Specification (1) uses logged effort as dependent variable, regressed on heterogeneity measures in terms of logged ratios, as

Table 2: Effect of Current and Future Heterogeneity on Effort

	<i>Sample split</i>		
	Pooled sample	Favorites	Underdogs
$heterogeneity_t$	-0.026 (0.017)	-0.076*** (0.029)	0.015 (0.028)
$E_t [rel. ability_{t+1}]$	0.028* (0.016)	0.081*** (0.030)	0.056* (0.031)
Observations	4398	2199	2199
R^2	0.125	0.147	0.142

Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses. Dependent variable is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$.

stipulated by a log-linearized approximation to the theoretical condition. Alternatively, Column (2) reports results for a specification where both on the left and the right-hand-side of the equation linear (instead of logged) measures are used. Specification (3) uses a linear specification of heterogeneity in terms of differences rather than ratios. All three specifications deliver a pattern of results that is consistent with the baseline results of Table 2. If anything, the effects are even quantitatively somewhat larger.

Next, we consider round effects and the importance of the current game for reaching the next round of the tournament. Intuitively, heterogeneity in the early rounds of the tournament might have different effects, and take different forms, than in later stages, such as the semi-final, when the possibility to win the championship is very salient. Expecting to meet a weak opponent in the next round(s) increases the chances to win the championship, but more so the fewer rounds are still to go and the less uncertainty is involved regarding relative future ability. This strategic aspect may be less important when the finals are still far away. Therefore, in a first set of additional results we estimate the effects separately at the different stages of the NBA playoff tournament.

Table 3: Effect of Current and Future Heterogeneity on Effort: Alternative Specifications

	<i>Pooled sample</i>			<i>Split: favorites and underdogs</i>					
	Alt. spec. I ^a	Alt. spec. II ^b	Alt. spec. III ^c	Alt. spec. I		Alt. spec. II		Alt. spec. III	
				fav. ^a	under. ^a	fav. ^b	under. ^b	fav. ^c	under. ^c
<i>heterogeneity_t</i>	-0.021* (0.012)	-0.039 (0.032)	-0.017 (0.011)	-0.048** (0.021)	0.001 (0.019)	-0.138** (0.056)	0.029 (0.048)	-0.044** (0.020)	0.006 (0.018)
$E_t [rel. ability_{t+1}]$	0.031* (0.017)	0.049* (0.028)	0.028* (0.016)	0.072** (0.029)	0.063 (0.039)	0.133*** (0.049)	0.102* (0.053)	0.075** (0.030)	0.053* (0.031)
Observations		4398		2199		2199		2199	
R^2	0.135	0.132	0.124	0.157	0.156	0.153	0.151	0.146	0.142

Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

^a Dependent variable is defined as $\frac{\text{fouls in playoff game}}{\text{avg. fouls regular season}}$. Heterogeneity_t is defined as $\frac{SRS_{Fav}}{SRS_{Und}}$, rel. ability_{t+1} is defined as $\frac{SRS}{EXP. OPP. SRS}$.

^b Dependent variable is defined as fouls in playoff game – avg. fouls regular season. Heterogeneity_t is defined as $SRS_{Fav} - SRS_{Und}$, rel. ability_{t+1} is defined as $SRS - EXP. OPP. SRS$.

^c Dependent variable is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$. Heterogeneity_t is defined as $\frac{SRS_{Fav}}{SRS_{Und}}$, rel. ability_{t+1} is defined as $\log[\frac{SRS}{EXP. OPP. SRS}]$.

Table 4: Effect of Current and Future Heterogeneity on Effort: Stratified by Tournament Round

	<i>Favorites</i>			<i>Underdogs</i>		
	round 1	round 2	round 3	round 1	round 2	round 3
<i>heterogeneity_t</i>	-0.060* (0.034)	-0.118** (0.056)	0.131 (0.093)	0.006 (0.032)	-0.037 (0.058)	0.463*** (0.092)
$E_t [\text{rel. ability}_{t+1}]$	0.060 (0.038)	0.090* (0.054)	0.156** (0.065)	0.043 (0.039)	0.058 (0.055)	0.192*** (0.059)
<i>mean current heterogeneity</i>	0.288	0.216	0.146	288	216	146
Observations	1145	699	355	1145	699	355
R^2	0.133	0.276	0.318	0.162	0.181	0.319

The dependent variable for all sub-samples is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$. Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

Table 4 presents the results for round 1, round 2 and round 3, respectively, where round 3 corresponds to the semi-final. The results for the pooled sample showed that the negative effect of heterogeneity is present on average. However, the negative effect is weaker in later rounds where heterogeneity is less pronounced and where the heterogeneity measure that is based on regular season performance is probably less informative. In round 3 there is no negative effect of heterogeneity, but the effect is even positive for underdogs in round 3 of the playoffs, suggesting that there might be other mechanisms at play, such as emotions leading underdogs that have made it to round 3 to fight particularly hard to make it to the final (see, e.g. Kräkel (2008) for an argument along this dimension).

The positive effect of expected relative ability in future rounds is present on average, and, as revealed by Table 4, it becomes more pronounced in later rounds. This pattern that the elasticity rises up to the third – pre-final – round, is present both for favorites and underdogs. The overall result from this analysis suggests that competitors are forward looking, and that the effect is more pronounced when the stakes are higher or the uncertainty related to expected relative ability in future rounds is lower.

A second dimension along which to expect some variation in the strength of the effect, in particular of expected relative ability in future rounds, is the pivotality of a particular game, i.e., whether a game can potentially lead to a decision in the current round of the tournament regarding promotion to the next round, or not. In order to understand how strategic behavior is affected by this dimension, we split the sample by distinguishing between pivotal games and non-pivotal ones. In pivotal games, either the favorite team can decide the series in its favor, or the underdog team, or either team. For non-pivotal games there will be a further game in the same series with certainty.

Table 5 reports the results for estimating equation (11) for different sub-samples. The results document that, in line with theoretical predictions, the effect of heterogeneity is detrimental for effort provision, particularly for favorites, who appear to reduce effort in response to a greater head start advantage in terms of ability in non-pivotal games – when slacking down is not very costly in terms of potentially being eliminated from the playoffs. Moreover, the effect of expected relative ability in future interactions is stronger for favorites, in particular in pivotal games where the favorite can win the series in the next game (Column (5)). Here, the effect of forward-lookingness is particularly prevalent and rewarding. If, on the other hand, the favorite is under pressure to potentially lose the series, there is no room to act strategically: maximum effort has to be exerted right now to stay in the tournament independently of the anticipated relative ability in a future round – there is no way to either slack down if the opponent is weak or to look into the future (Column (7)). Underdogs, on the other hand, do not exhibit statistically significant effects of current heterogeneity or expected relative ability in future rounds (the effect is positive but not significantly different from zero). We interpret this pattern of results as broadly consistent with the theoretical predictions. Teams tend to accommodate their behavior to the strength of the current or future opponents as long as they can. Underdogs cannot afford to slack down anytime, and if a favorite team is in danger of losing prematurely, the current game receives all concentration. Moreover, the results suggest that teams have a rather short horizon with respect to the continuation value of the game. They react strongest if the next phase of the tournament is very close.

Table 5: Effect of Current and Future Heterogeneity on Effort: Stratified by Standings

	<i>non-pivotal games</i>				<i>pivotal games</i>			
	<i>all</i>		<i>tied</i>		<i>favorite can win</i>		<i>underdog can win</i>	
	Fav. (1)	Under. (2)	Fav. (3)	Under. (4)	Fav. (5)	Under. (6)	Fav. (7)	Under. (8)
<i>heterogeneity_t</i>	-0.076*** (0.035)	0.007 (0.033)	-0.098* (0.050)	0.010 (0.047)	-0.068 (0.061)	-0.012 (0.061)	-0.070 (0.108)	0.116 (0.118)
E_t [<i>rel. ability_{t+1}</i>]	0.068* (0.037)	0.044 (0.036)	0.108** (0.049)	0.057 (0.049)	0.140** (0.055)	0.038 (0.062)	0.027 (0.094)	0.076 (0.090)
Observations	1578		763		406		215	
R^2	0.154	0.149	0.169	0.167	0.222	0.207	0.293	0.263

The dependent variable for all sub-samples is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$. Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

4.3 Robustness: Alternative Contest Technology

The results so far are based on the assumption that the winning probability is given by a standard Tullock contest technology with discriminatory power $r = 1$. In particular, in the baseline we specified $p_k = \frac{[a_k]^r}{[a_k]^r + [a_l]^r}$ under the assumption of $r = 1$. To investigate the robustness of the results, we explore which value of r most closely predicts the outcomes we observe in the data. In order to do so, we compare outcomes of a simple probit model predicting win probabilities for respective differences in strength of the teams with simulated outcomes of the Tullock contest technology.

$$P(\text{win}_i = 1 | \text{SRS}_i, \text{SRS}_j) = \Phi(\beta_0 + \beta_1 \text{SRS}_i + \beta_2 \text{SRS}_j), \quad (13)$$

where win is a binary variable equal to 1, if team i wins the playoffs series against opponent j , Φ is the normal cdf, and SRS is the ability measure according to the Simple Rating System. The winning probabilities that are predicted from the estimation results of this probit model are plotted against the level of heterogeneity (in terms of the rating advantage of the favorite over the underdog) in Figure 9 in the Appendix. In the same figure, we plot the simulated winning probabilities using the contest technology in (2) for different levels of r (in particular, for $r = 1$, $r = 6$ and $r \rightarrow \infty$). The figure indicates that $r = 6$ provides the best fit.²¹

Table 9 in the Appendix documents that the main results are unaffected when computing the measures of heterogeneity and relative future ability not under the assumption of $r = 1$, but using the alternative computation with $r = 6$ that comes from matching the winning probabilities to the heterogeneity measure. Also the additional results for alternative specifications, and for favorites and underdogs stratified by round or standings are very similar when constructing the heterogeneity measures based on $r = 6$ instead of $r = 1$, see Tables 10, 11 and 12 in the Appendix.

4.4 Robustness: Alternative Data

In this subsection, we report results from a replication of the analysis using data from the playoffs of the annual basketball tournament of the National Collegiate Athletic Associ-

²¹Graphs with simulations for other values of r are available on request.

Table 6: NCAA Data: Effect of Current and Future Heterogeneity on Effort: Stratified by Round

	<i>Favorites</i>			<i>Underdogs</i>		
	round 1	rounds 2-3	rounds 3-5	round 1	rounds 2-3	rounds 3-5
<i>heterogeneity_t</i>	-0.147*** (0.040)	-0.241** (0.114)	-0.059 (0.168)	-0.047 (0.054)	0.099 (0.120)	0.103 (0.211)
$E_t [rel. ability_{t+1}]$	-0.081* (0.045)	0.104** (0.044)	0.109* (0.063)	0.013 (0.042)	0.109** (0.050)	0.136* (0.072)
<i>mean current heterogeneity</i>	0.377	0.163	0.136	0.377	0.163	0.136
Observations	352	330	154	352	330	154
R^2	0.256	0.193	0.269	0.126	0.157	0.240

The dependent variable for all sub-samples is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$. Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, playoff-stage dummies and season dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

ation (NCAA). This provides useful information regarding the robustness and external validity of the results, because the NCAA corresponds to a semi-professional setting and applies somewhat different rules than the NBA. From the perspective of this paper, the most important differences between the NBA and the NCAA regulation concerns the organisation of playoffs. While in the NBA each round is decided by a best-of-seven or best-of-five series, each NCAA playoff round is decided in a single game, which is held on neutral ground to avoid home bias. As each game has a pivotal character in the NCAA, forward-looking behavior in response to the expected relative ability in future rounds might be even more prevalent as the next round is always salient.

Table 6 presents the respective results for the NCAA, stratified by rounds. The results are qualitatively very similar to the results obtained with the NBA data in Table 4. In particular, favorites allow themselves some slacking off if they are paired with a weak current opponent, but this effect is not present for underdogs. Also, favorites and underdogs increase their effort if they can expect to meet a weaker opponent in the future, in particular in later rounds of the tournament.

4.5 Robustness: Alternative Construction of Relative Strength of Future Opponents

Another robustness check concerns the construction of the measure of expected future heterogeneity in terms of relative ability. Table 7 presents results for alternative specifications of the future heterogeneity. In Columns (1) and (2), we replicate the results of the baseline specification, where the heterogeneity measure is not updated in case the future opponent was already known at the time of the respective game. In Columns (3) and (4), we construct the relative future heterogeneity by discarding the underdog in the respective parallel game and only use the strength of the favorite as expected strength of the future opponent. Columns (5) and (6) represent an improved version of the baseline measure, which is updated in case the future opponent was already known at the time of a given game, and only uses the probability in case the parallel match-up is still undecided. Finally, Columns (7) and (8) again use a probability-weighted average of both potential future opponents, but now with an additional probabilistic updating of the win percentages of the future opponents.²² The results are very robust with respect to these variations. In particular, both for favorites as well as underdogs, the coefficient for future heterogeneity does not change in any of these variations.²³

²²Probabilistic updating takes into account the updated win probabilities according to the current standings in the best-of-5 or best-of-7 series using binomial updating.

²³Table 14 in the Appendix replicates the same results under the assumption of $r = 6$.

Table 7: Effect of Current and Future Heterogeneity on Effort: Construction of Relative Strength of Future Opponent

	<i>probability weighted^a</i>		<i>Future favorite only^b</i>		<i>preferred specification^c</i>		<i>preferred spec.-updated^d</i>	
	Fav. (1)	Under. (2)	Fav. (3)	Under. (4)	Fav. (5)	Under. (6)	Fav. (7)	Under. (8)
<i>heterogeneity_t</i>	-0.074** (0.030)	0.015 (0.029)	-0.091*** (0.031)	0.010 (0.027)	-0.076*** (0.039)	0.015 (0.028)	-0.091*** (0.031)	0.009 (0.027)
E_t [<i>rel. ability_{t+1}</i>]	0.079** (0.032)	0.055* (0.033)	0.091*** (0.029)	0.058** (0.030)	0.081*** (0.030)	0.056* (0.031)	0.090*** (0.028)	0.055* (0.029)
Observations	2199		2199		2199		2199	
R^2	0.147	0.142	0.149	0.143	0.147	0.142	0.149	0.142

The dependent variable for all sub-samples is defined as $\log[\text{avg. fouls regular season}] - \log[\text{fouls in playoff game}]$. Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

^a The favorite as well as the underdog in the parallel interaction is considered as the expected future opponent. The probability-weighted average of both potential future opponents is defined as the expected future opponent's ability.

^b Only the favorite in the parallel interaction is considered as the expected future opponent.

^c The preferred specification is the same as used in tables 2 through 12

^d For these specifications we use the same measure for future heterogeneity as in the preferred specification with an additional probabilistic updating of the win percentages of the potential future opponents.

4.6 Robustness: Shadow of Future Competition and Outcomes

Our data can also be used to investigate the effect of the shadow of future competition, i.e., the expected relative ability on the subsequent stage of the tournament, on outcomes in terms of winning probability or point margins, thereby replicating the analysis by Brown and Minor (2014). In particular, they estimate an empirical model

$$X_t = \beta_0 + b_1 \text{heterogeneity}_t + b_2 E_t[\text{rel. ability}_{t+1}] + b_3 \text{spillover}_{t-1} + \epsilon, \quad (14)$$

where the outcome X_t can be modeled either as whether i) the favorite wins the game or ii) the share of total points going to the favorite. In the terminology of Brown and Minor (2014), the coefficient b_1 reflects the effect of current competition (in terms of heterogeneity on the current stage), b_2 reflects the shadow of future competition (the expected relative ability on the next stage), and b_3 represents a “spillover” effect of past outcomes, measured in terms of the sum of points scored in the previous match, which could reflect exhaustion or positive effects due to a “hot hand”.

Table 8 shows the results from the perspective of favorites. We present separate results by either taking a single game or the entire series (involving maximally 5 or 7 games) as a unit of observation. The reason for considering a series is that luck is expected to play a smaller role for the ultimate result in a series than in a single game, implying that ability differences are more decisive when considering the overall outcome. With the different setup in the NCAA, where each game is decisive since there is only one game in a given round, one would expect the results should to be more comparable to Columns (1) and (2).

It turns out that the measure of heterogeneity between the favorite and the underdog on the current stage has a significant positive impact on the outcome of the game measured in the favorite’s win probability and the point margin, corroborating the results by Brown and Minor (2014) for all outcome measures and both data sets. The expected relative ability of the favorite on the subsequent stage (reflecting something like the inverse of the “shadow of future competition”) has a positive and significant effect in most

Table 8: Effect on Probability of the Favorite Winning and Favorite’s Point Margins: Heterogeneity and the Shadow of Future Competition

	NBA				NCAA	
	<i>game</i>		<i>series</i>		<i>win^a</i> (5)	<i>share of total points^b</i> (6)
	<i>win^a</i> (1)	<i>share of total points^b</i> (2)	<i>win^a</i> (3)	<i>share of total games^b</i> (4)		
<i>heterogeneity_t</i>	0.381*** (0.071)	0.029*** (0.005)	0.641*** (0.131)	0.391*** (0.074)	0.405*** (0.056)	0.081*** (0.006)
$E_t [rel. ability_{t+1}]$	0.113* (0.065)	0.011** (0.004)	0.130 (0.126)	0.157** (0.073)	0.092* (0.051)	0.000 (0.005)
<i>spillover_{t-1}</i>	-0.013 (0.013)	-0.001 (0.001)	-0.030 (0.033)	-0.018 (0.015)	0.001 (0.001)	0.000** (0.000)
N	2199	2199	434	434	682	682
R^2	0.152	0.191	0.169	0.219	0.129	0.312

All specifications in columns 1-4 (NBA) include a dummy equal to 1 if the team plays at home, a dummy equal to 1 if the series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies. specifications in columns 5 and 6 (NCAA) include round and year dummies, overtime dummies and regional dummies. Robust standard errors (clustered for individual playoff-series for columns 1 and 2 - clustered on team-year level for columns 5 and 6) in parentheses. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively.

^a Dependent variable is equal to 1 if favorite wins, 0 else.

^b Dependent variable is equal to the favorite’s share of total points (games) in game (series).

specifications.²⁴ The data do not reveal any spillover effect of the sum of total points scored in the previous playoff round, which might have to do with the fact that there are multiple days of rest between playoff series, which reduces the potential role of physical or psychological fatigue, compared to the tennis environment studied by Brown and Minor (2014). As expected, the results are quantitatively larger for the series as compared to the single NBA games or the NCAA games.

Using a direct measure of effort has the advantage that we can open the black box that links outcomes in a game to the heterogeneity of tournament participants. In particular, the empirical results suggest that the shadow of future competition effect on winning probabilities discussed by Brown and Minor (2014), indeed, is caused by the theoretical mechanism put forward in the previous section, according to which the effort reactions

²⁴The positive effect of future relative ability is also present if we use betting odds (and the respective subset of the data for which betting odds are available) rather than outcomes, in line with results reported by Brown and Minor (2014) for tennis tournaments. Details are available upon request.

to future relative ability by favorites are stronger than the corresponding reactions by underdogs. Exploring whether effort choice can increase a team's probability of winning a game is less straightforward, however. While increased effort has been shown to increase own points and turnovers and decrease opponent's points (Table 1), running a simple regression of effort choice on winning probabilities is problematic due to the presence of two obvious endogeneity problems: First, losing teams are *ceteris paribus* more likely to foul, since they are in the defense more often. Second, trailing teams have an incentive to foul in order to stop the clock towards the end of the game in order to gain possession of the ball. Even though a high-effort defense that delivers a comparatively high number of fouls is optimal and likely to improve winning probabilities in the respective situation, the fact that many of these teams still lose may bias empirical estimates of the relation between effort and a team's probability of winning a game, which essentially prevents pursuing this line.²⁵

5 Conclusion

This paper presented evidence that decision makers are forward looking in dynamic strategic interactions. The evidence is based on field data from professional and semi-professional basketball tournaments. In particular, the results suggest that tournament participants reduce their effort if current heterogeneity increases, and, more importantly, that tournament participants exert more effort in response to standing a better chance of prevailing on the next stage of the tournament. The latter effect, which is consistently

²⁵One way to overcome such an endogeneity problem could be to use expected ability of the future opponent as instrument for effort. Expected ability of the future opponent is a valid instrument because it has a proven influence on own effort, but is unrelated to situations during the game, whether a team is trailing, etc. because it is fixed at the start of the game. The corresponding results for a simple linear regression model for the win probability of the favorite for the entire sample and for a subsample of pivotal situations and first games of a series are presented in Table 15 in the Appendix. The estimates show a counterintuitive negative correlation which suggests that more effort is counterproductive such that it reduces the winning probability rather than increasing it. Using the instrumental variables strategy for the subsample of pivotal situations and first games of a series for which the impact of future competition is highest, and hence the instrument strongest with a reasonably high first stage with an F-value above 10, the coefficient becomes positive. This suggests that reverse causation may indeed be present. However, due to sample size, the positive effect is not statistically significant. Interestingly, these findings align with results reported by Bartling, Brandes, and Schunk (2015) who find that soccer teams that fall behind the expected match outcome are more likely to breach the rules and get assigned yellow or red cards, even though this leads to a worse expected match outcome.

found to be statistically and economically significant, constitutes a necessary condition for any model of rational behavior in multi-stage tournaments. The results therefore illustrate the importance of expectations about future interactions for competitive behavior and provide support for the ‘standard’ assumption in theoretical models that agents’ are forward looking. Empirical support for this important assumption from the field was lacking so far. Our findings are robust to the use of data from various tournament settings, to the construction of different heterogeneity measures. In particular, the impact of the expected relative strength in future interactions on measured effort in the current game cannot be explained by dynamic or tactical interactions during the course of the current game. Expected ability of the future opponent is fixed pre-game and does not change during a game, but it increases the incentive to work hard because the ultimate expected payoff of the tournament increases. The findings also suggest heterogeneity in the degree to which the future is incorporated in current performance. In particular, forward looking behavior seems more prevalent if the future interactions are foreseeable with greater certainty. On the other hand, future interactions affect performance less if the future is unlikely to play a role.

The results of this paper have practical implications, as they show that rational forward-looking behavior appears to be prevalent in field data, lending support to theories whose mechanisms rely on future prospects, in terms of promotions or better outside options. Moreover, the heterogeneity of the results raises the question about the size, the subjective determinants, and the salience of the continuation value that induces forward-looking behavior. A systematic analysis of these issues provides a logical next step and a promising avenue for future research.

References

- ALTMANN, S., A. FALK, AND M. WIBRAL (2012): “Promotions and Incentives: The Case of Multi-Stage Elimination Tournaments,” *Journal of Labor Economics*, 30(1), 149–174.
- BAIK, K. (1994): “Effort Levels in Contests with Two Asymmetric Players,” *Southern Economic Journal*, 61, 367–378.
- BARTLING, B., L. BRANDES, AND D. SCHUNK (2015): “Expectations as Reference Points: Field Evidence from Professional Soccer,” *Management Science*, *forthcoming*.
- BERGER, J., AND P. NIEKEN (2014): “Heterogeneous Contestants and the Intensity of Tournaments: An Empirical Investigation,” *Journal of Sports Economics*, 17, 1–30.
- BROWN, J. (2011): “Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars,” *Journal of Political Economy*, 119(5), 982–1013.
- BROWN, J., AND D. MINOR (2014): “Selecting the Best: Effort Spillover and Shadows in Elimination Tournaments,” *Management Science*, 60(12), 3087–3102.
- BULL, C., A. SCHOTTER, AND K. WEIGELT (1987): “Tournaments and Piece Rates: An Experimental Study,” *Journal of Political Economy*, 95(1), 1–33.
- CHEN, H., S. HAM, AND N. LIM (2011): “Designing Multiperson Tournaments with Asymmetric Contestants: An Experimental Study,” *Management Science*, 57, 864–883.
- DELFGAAUW, J., R. DUR, A. NON, AND W. VERBEKE (2015): “The Effects of Prize Spread and Noise in Elimination Tournaments: A Natural Field Experiment,” *Journal of Labor Economics*, 33(3), 521–569.
- DEUTSCHER, C., B. FRICK, O. GÜRTLER, AND J. PRINZ (2013): “Sabotage in Tournaments with Heterogeneous Contestants: Empirical Evidence from the Soccer Pitch,” *Scandinavian Journal of Economics*, 115, 1138–1157.
- GOLDMAN, M., AND J. M. RAO (2012): “Effort vs. Concentration: The asymmetric impact of pressure on NBA performance,” Discussion paper, UC San Diego.

- HARBRING, C., AND G. K. LÜNSER (2008): “On the Competition of Asymmetric Agents,” *German Economic Review*, 9, 373–395.
- KRÄKEL, M. (2008): “Emotions and the optimality of uneven tournaments,” *Review of Managerial Science*, 2(1), 61–79.
- LAZEAR, E. P. (1989): “Pay Equality and Industrial Politics,” *Journal of Political Economy*, 97(3), 561–80.
- LAZEAR, E. P., AND S. ROSEN (1981): “Rank-Order Tournaments as Optimal Labor Contracts,” *Journal of Political Economy*, 89(5), 841–864.
- NTI, K. O. (1999): “Rent-Seeking with Asymmetric Valuations,” *Public Choice*, 98, 415–430.
- ROSEN, S. (1986): “Prizes and Incentives in Elimination Tournaments,” *American Economic Review*, 76(4), 701–715.
- STRACKE, R., AND U. SUNDE (2015): “Dynamic Incentive Effects of Heterogeneity in Multi-Stage Promotion Contests,” *University of Munich, mimeo*.
- SUNDE, U. (2009): “Heterogeneity and Performance in Tournaments: A Test for Incentive Effects using Professional Tennis Data,” *Applied Economics*, 41, 3199–3208.
- TULLOCK, G. (1980): “Efficient Rent-Seeking,” in *Toward a Theory of the Rent-Seeking Society*. J.M. Buchanan and R.D. Tollison and G. Tullock (Eds.). Texas A&M Press, College Station, p. 97-112.
- WALDMAN, M. (1984): “Job Assignments, Signalling, and Efficiency,” *Rand Journal of Economics*, 15, 255–267.

A Theoretical Model

A.1 Proofs

Result 1 (Effect of Current Heterogeneity). *A higher degree of heterogeneity $\theta = \left(\frac{a_F}{a_U}\right)^r$ between a favorite and an underdog in any stage of a multi-stage pairwise elimination tournament implies*

- (a) *a lower level of absolute effort of favorites and underdogs on the respective stage of the tournament;*
- (b) *a more pronounced negative effect on effort of favorites than of underdogs.*

Proof. According to equation (7), equilibrium efforts of the favorite and the underdog with $a_F \geq a_U$ are defined as follows:

$$\hat{e}_F = r \cdot a_F \cdot \frac{\theta}{(1+\theta)^2} \cdot [R_{\text{now}} + \text{CV}_F^*(\kappa_F)] \quad (\text{A.1})$$

$$\hat{e}_U = r \cdot a_U \cdot \frac{\theta}{(1+\theta)^2} \cdot [R_{\text{now}} + \text{CV}_U^*(\kappa_U)] . \quad (\text{A.2})$$

To prove part (a), note that the sign of the first derivative of both expressions with respect to the degree of heterogeneity θ is determined by

$$\frac{\partial \left(\frac{\theta}{(1+\theta)^2} \right)}{\partial \theta} = \frac{1-\theta}{(1+\theta)^3} .$$

This expression is negative if $\theta = \left(\frac{a_F}{a_U}\right)^r > 1$, implying that effort decreases as heterogeneity increases. If instead $\theta = \left(\frac{a_U}{a_F}\right)^r < 1$, the expression is positive, which implies that effort increases as θ approaches 1, i.e., as heterogeneity decreases.

Regarding part (b), note that heterogeneity affects effort choices of both favorites and underdogs due to changes of the expression $\frac{\theta}{(1+\theta)^2}$. The same marginal change of heterogeneity through changes in θ is multiplied by the factor $a_F[R_{\text{now}} + \text{CV}_F^*(\kappa_F)]$ for favorites and by the factor $a_U[R_{\text{now}} + \text{CV}_U^*(\kappa_U)]$ for underdogs, however. Given that $a_F \geq a_U$ by assumption and $\text{CV}_U^*(\kappa_F) \geq \text{CV}_U^*(\kappa_U)$ due to the assumption that $a_F \geq a_U$ holds, the negative effect of heterogeneity on effort is more pronounced for favorites than for underdogs. \square

Result 2 (Effect of Future Relative Ability). *Let a_x be the expected ability of the future opponent x and $i = F, U$. Effort choices of the favorite and the underdog on any non-final stage are then increasing in future relative ability $\kappa_i = \left(\frac{a_i}{a_x}\right)^r$ if and only if agents are forward looking.*

Proof. Recall that effort choices in the multi-stage tournament are defined as follows in equation (7):

$$e_i^* = r \cdot a_i \cdot \frac{\theta}{(1 + \theta)^2} \cdot [R_{\text{now}} + \text{CV}_i^*(\kappa_i)] , \quad (\text{A.3})$$

where $\text{CV}_i^*(\kappa_i) = \frac{[\kappa_i]^2 + (1-r)\kappa_i}{(1+\kappa_i)^2}$. Agents who are not forward looking will not take the continuation value part into account, such that their effort choice is independent of κ_i . Forward looking agents, however, react to changes in their continuation value. Given that effort is strictly increasing in the continuation value according to equation (7), it is sufficient for the remainder of this proof to show that the continuation value is increasing in the own relative ability. The first derivative of the continuation value with respect to relative ability κ reads

$$\frac{\partial \text{CV}_i^*(\kappa_i)}{\partial \kappa_i} = \frac{(1+r)\kappa + 1 - r}{(1+\kappa)^3} .$$

This derivative is strictly positive in any pure strategy Nash equilibrium, given that the restriction for the existence of a pure strategy equilibrium reads $r \leq \left(\frac{a_U}{a_F}\right)^r + 1$ – see Nti (1999) for details. \square

A.2 Extensions of the Theoretical Model

The theoretical model outlined in section 2 makes several simplifying assumptions: First, the model abstracts from the effect of heterogeneous continuation values on effort choices in earlier stages, as discussed in footnote 3. Second, it is assumed that winning probabilities in the parallel match are taken as given, abstracting from the presence of feedback effects across interactions. Finally, it is assumed that the expected ability of future opponents equals a weighted average of the abilities of possible future opponents – see equation (5) – while the continuation value is usually defined as a weighted average of expected payoffs against possible future opponents. Consider a generalized version of a two-stage

pairwise elimination tournament without these simplifying assumptions. Assume without loss of generality that agents 1 and 2 meet in one of the two interactions in stage 1, while agents 3 and 4 compete in the remaining one. Current stage-1 heterogeneity in a winner-takes-all setting is then defined as $\theta_{12} = \left(\frac{a_1 CV_1^*}{a_2 CV_2^*}\right)^r$ and $\theta_{34} = \left(\frac{a_3 CV_3^*}{a_4 CV_4^*}\right)^r$, respectively, accounting for the effect of heterogeneous continuation values on effort choices in earlier stages. Moreover, we assume that the two pairwise interactions in stage 1 are simultaneous, implying that winning probabilities in the parallel math are endogenously determined. Finally, we assume that the continuation value of any agent is defined as a weighted average of expected payoffs against possible future opponents, i.e., the continuation value of agent 1 is formally defined as

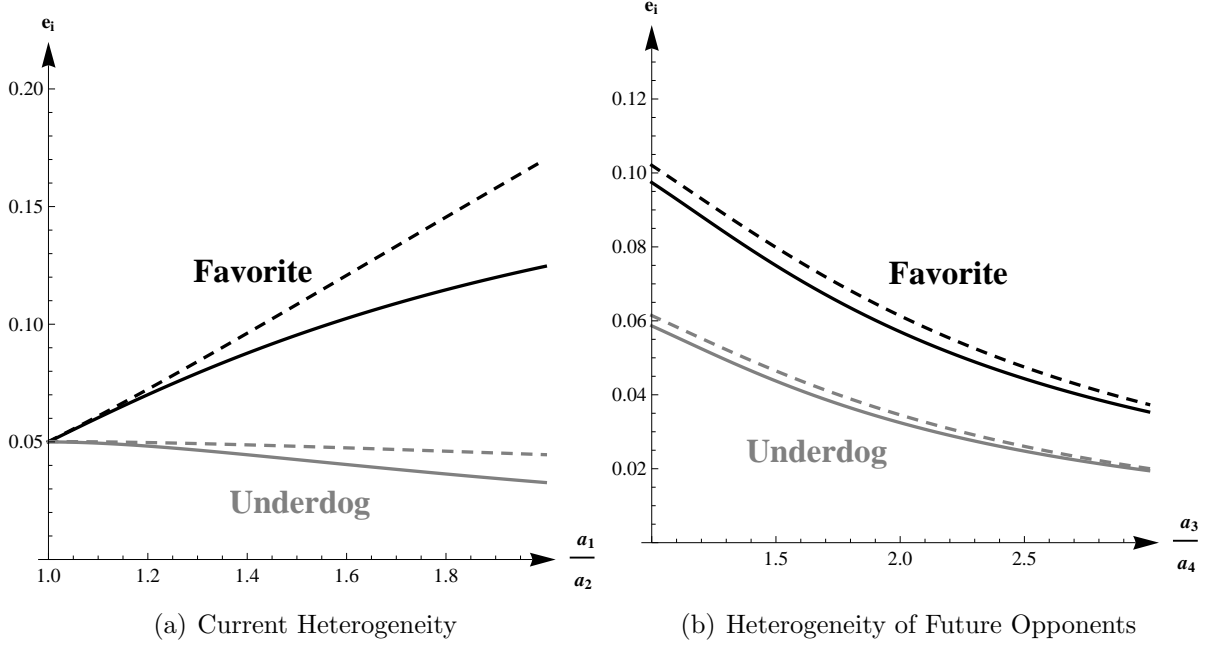
$$CV_1^* = \frac{(e_3)^r}{(e_3)^r + (e_4)^r} \Pi_{13}^* + \frac{(e_4)^r}{(e_3 + e_4)^r} \Pi_{14}^* ,$$

where Π_{1j}^* is the expected payoff of agent 1 when competing with agent j on stage 2.

In what follows, we compare equilibrium effort in the simplified model with equilibrium efforts that a generalized version of a two-stage pairwise elimination tournament predicts to investigate the importance of these assumptions.²⁶ Consider first panel (a) of Figure 7 which plots stage-1 effort of the ‘favorite’ agent 1 and the ‘underdog’ agent 2 as a function of current heterogeneity $\theta = \frac{a_1}{a_2}$, holding heterogeneity in the parallel stage-1 interaction fixed at the mean level of heterogeneity in the sample. The figure reveals that the simplified model overestimates stage-1 effort of the favorite and the underdog, given that the dotted lines that depict equilibrium predictions of the simplified model are always above the solid lines that provide the equilibrium prediction of the extended model. Given that the simplified model abstracts from heterogeneity induced differences in continuation values and therefore underestimates current heterogeneity between the favorite and the underdog, this was to be expected. Note, however, that the slope of the respective functions is very similar for low degrees of heterogeneity which constitute the majority of observations (see Figure 2). Moreover, note that the difference between solid and dotted lines is related to current heterogeneity, i.e., the higher current heterogeneity, the larger the deviation between the extended and the simplified model. Said differently, current

²⁶An analytical solution of the extended model is available from the authors upon request.

Figure 7: The Effect of Heterogeneity on Effort



Notes: The figures plot stage-1 effort of the favorite and the underdog for $r = 1$, $R_{\text{now}} = 0$, and $R_{\text{fut}} = 1$ as a function of current heterogeneity in panel (a), and as a function of heterogeneity in the parallel match in panel (b). The solid (dotted) lines depict equilibrium predictions of the extended (simplified) two-stage pairwise elimination tournament. Chosen parameters for abilities in panel (a) are $a_2 = 1$, $a_3 = 1.4$, and $a_4 = 1$, while a_1 varies between 1 and 2. Chosen parameters for abilities in panel (b) are $a_1 = 1.4$, $a_2 = 1$, and $a_3 = 1$, while a_4 varies between 1 and 3.

heterogeneity determines the importance of heterogeneity induces differences in continuation values. Given that we control for current heterogeneity in the empirical model, the estimated coefficient for the effect of current heterogeneity will account for heterogeneity due to ability differences, and for heterogeneity induced differences in continuation values.

Consider next panel (b) of Figure 7 which plots stage-1 effort of the ‘favorite’ agent 1 and the ‘underdog’ agent 2 as a function of heterogeneity between potential future opponents, i.e., of heterogeneity between agents 3 and 4 in the parallel stage-1 interaction, holding heterogeneity in the current stage-1 interaction between agents 1 and 2 fixed at the mean level of heterogeneity in the sample. The figure reveals that the simplified model still overestimates stage-1 effort of the favorite and the underdog by not accounting for heterogeneity induced differences in continuation values. Feedback effects across parallel stage-1 interactions and the method to determine continuation values do not matter much,

however, given that the slope of the respective curves is very similar across simplified and extended versions of the two-stage tournament.

The observation that neither feedback effects across parallel stage-1 interactions nor the method to determine continuation values matter much does not depend on the particular parameters chosen for the numerical exercise in Figure 7 (b).²⁷ Only the assumption to abstract from heterogeneity induced differences in continuation values in the simplified model delivers notable differences in effort if current heterogeneity is high. As discussed above, the estimated coefficient for the effect of current heterogeneity will account for heterogeneity due to ability differences and for heterogeneity induced differences in continuation values, however, implying that the simplifying assumption of the theoretical model will not affect the empirical findings.

A.3 Additive Noise Technology (Lazear and Rosen, 1981)

Consider the simplest Lazear and Rosen (1981) tournament with two agents 1 and 2 who compete for a prize R that is awarded to the agent whose performance is relatively higher. The performance y_i of agent $i \in \{1; 2\}$ reads

$$y_i(e_i) = e_i + u_i, \tag{A.4}$$

where e_i is the effort and u_i is the random performance component of agent i . Consider the optimization problem of agent i who maximizes her expected payoff

$$\Pi_i(e_i, e_j) = p(e_i, e_j) \cdot R - \frac{c(e_i)}{a_i}, \tag{A.5}$$

where $i \neq j$, $p(e_i, e_j) = \text{prob}[e_i + u_i > e_j + u_j]$ is the probability that the prize is awarded to agent i , a_i is the ability of agent i , and $c(\cdot)$ is the cost of effort function. Assume that the random performance components of both agents are independent draws from a continuous distribution with mean zero, and that the cost function is sufficiently convex

²⁷Further details available from the authors upon request.

to ensure that equilibrium efforts in the unique symmetric pure strategy Nash equilibrium are defined by first-order optimality conditions. First-order optimality conditions read:

$$\frac{\partial p(e_1, e_2)}{\partial e_1} R - \frac{c'(e_1)}{a_1} = 0 \quad (\text{A.6})$$

$$\frac{\partial p(e_2, e_1)}{\partial e_2} R - \frac{c'(e_2)}{a_2} = 0. \quad (\text{A.7})$$

Marginal winning probabilities of both agents $\frac{\partial p(e_1, e_2)}{\partial e_1}$ and $\frac{\partial p(e_2, e_1)}{\partial e_2}$ are equal for both agents, as they define the cutoff of the probability density function which is necessarily identical across agents – see Harbring and Lünser (2008) for details. This implies that the relation

$$\frac{c'(e_1)}{a_1} = \frac{c'(e_2)}{a_2} \Leftrightarrow e_1 = c' \left[\frac{a_1}{a_2} \right]^{-1} \cdot e_2 \quad (\text{A.8})$$

holds in equilibrium, where $c'[\cdot]^{-1}$ is the inverse function of marginal costs. Let $\Phi(\cdot)$ be the cumulative distribution function of the difference of random terms $u_2 - u_1$, $\phi(\cdot) = \Phi'(\cdot)$ the corresponding density function. We can then express the winning probability of agent 1 as follows:

$$p(e_1, e_2) = \text{prob}[e_1 + u_1 > e_2 + u_2] = \text{prob}[e_1 - e_2 > u_2 - u_1] = \Phi(e_1 - e_2).$$

Consequently, we obtain

$$\frac{\partial p(e_1, e_2)}{\partial e_1} = \frac{\partial p(e_2, e_1)}{\partial e_2} = \phi(e_1 - e_2). \quad (\text{A.9})$$

The combination of (A.9), (A.8), and the first-order condition of agent i delivers equilibrium efforts

$$e_1^* = c' \left[a_1 \cdot \phi \left(c' \left[\frac{a_1 - a_2}{a_2} \right]^{-1} \right) \cdot R \right] \quad (\text{A.10})$$

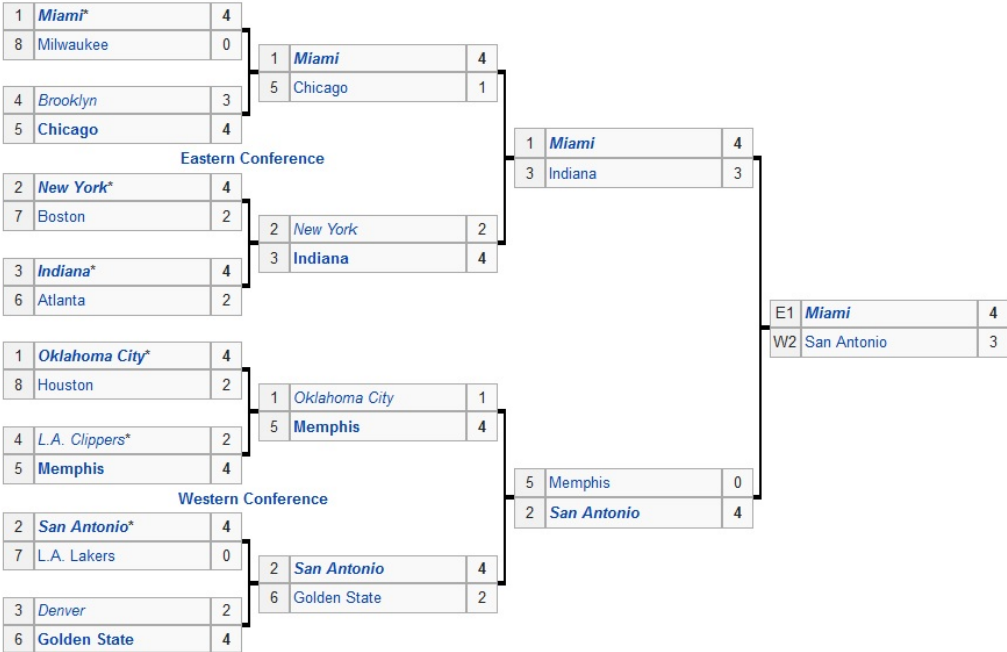
$$e_2^* = c' \left[a_2 \cdot \phi \left(c' \left[\frac{a_1 - a_2}{a_2} \right]^{-1} \right) \cdot R \right]. \quad (\text{A.11})$$

These expression for equilibrium effort show that effort is a function of own ability, relative ability defined by the difference of abilities, and of the prize at stake. Just as in the

Tullock setting discussed in the main part of this paper, the continuation value of future interactions is increasing in the own relative ability.

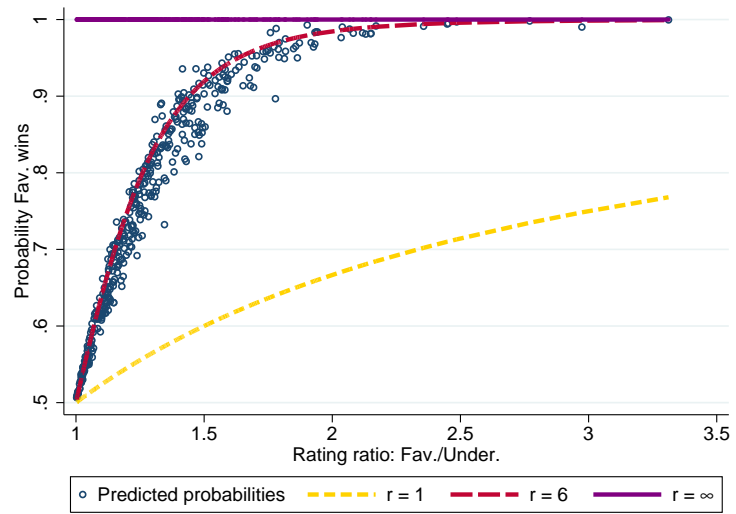
B Additional Figures

Figure 8: Illustration of the NBA Tournament Structure - Example: 2013 NBA Playoffs



Source: Wikipedia.

Figure 9: Probit model: Estimated win probabilities for different r 's



Notes: Estimated win probabilities for playoff series round 1–3, derived from Probit model explaining series win==1 by own rating and opponent rating vs. simulated Tullock probabilities for different values of r . $N = 434$

C Additional Tables

Table 9: Effect of current and future heterogeneity on Effort, $r = 6$

	<i>Pooled sample</i>	<i>Sample split</i>	
		Favorites	Underdogs
<i>heterogeneity_t</i>	-0.030* (0.017)	-0.087*** (0.031)	0.009 (0.027)
$E_t [\textit{rel. ability}_{t+1}]$	0.033** (0.016)	0.086*** (0.028)	0.056* (0.029)
Observations	4398	2199	2199
R^2	0.125	0.148	0.143

Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses. Dependent variable is defined as $\log[\textit{fouls in playoff game}] - \log[\textit{avg. fouls regular season}]$.

Table 10: Effect of current and future heterogeneity on Effort: different functional forms, r=6

	<i>Pooled sample</i>			<i>Split: favorites and underdogs</i>					
	Alt. spec. I ^a	Alt. spec. II ^b	Alt. spec. III ^c	Alt. spec. I		Alt. spec. II		Alt. spec. III	
				fav. ^a	under. ^a	fav. ^b	under. ^b	fav. ^c	under. ^c
<i>heterogeneity_t</i>	-0.024** (0.012)	-0.048 (0.032)	-0.019* (0.011)	-0.055** (0.022)	-0.002 (0.018)	-0.160*** (0.058)	0.017 (0.046)	-0.051** (0.021)	0.003 (0.017)
$E_t [\textit{rel. ability}_{t+1}]$	0.037** (0.017)	0.059** (0.026)	0.033** (0.016)	0.080*** (0.028)	0.067* (0.038)	0.139*** (0.045)	0.105** (0.049)	0.080*** (0.028)	0.054* (0.029)
Observations		4398		2199		2199		2199	
R^2	0.136	0.132	0.125	0.157	0.156	0.154	0.152	0.147	0.142

Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

^a Dependent variable is defined as $\frac{\text{fouls in playoff game}}{\text{avg. fouls regular season}}$. Heterogeneity_t is defined as $\frac{SRS_{Fav}}{SRS_{Und}}$, rel. ability_{t+1} is defined as $\frac{SRS}{EXP. OPP. SRS}$.

^b Dependent variable is defined as fouls in playoff game – avg. fouls regular season. Heterogeneity_t is defined as $SRS_{Fav} - SRS_{Und}$, rel. ability_{t+1} is defined as $SRS - EXP. OPP. SRS$.

^c Dependent variable is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$. Heterogeneity_t is defined as $\frac{SRS_{Fav}}{SRS_{Und}}$, rel. ability_{t+1} is defined as $\log[\frac{SRS}{EXP. OPP. SRS}]$.

Table 11: Effect of current and future heterogeneity on effort: stratified by round, $r = 6$

	<i>Favorites</i>			<i>Underdogs</i>		
	round 1	round 2	round 3	round 1	round 2	round 3
<i>heterogeneity_t</i>	-0.067* (0.035)	-0.141** (0.057)	0.095 (0.096)	0.001 (0.031)	-0.048 (0.055)	0.398*** (0.089)
E_t [<i>rel. ability_{t+1}</i>]	0.061* (0.035)	0.118** (0.049)	0.165** (0.065)	0.042 (0.037)	0.047 (0.053)	0.199*** (0.058)
Observations	1145	699	355	1145	699	355
R^2	0.133	0.280	0.319	0.162	0.180	0.320

The dependent variable for all sub-samples is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$. Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

Table 12: Effect of current and future heterogeneity on Effort: stratified by standings, $r = 6$

	<i>non-pivotal games</i>				<i>pivotal games</i>			
	<i>all</i>		<i>tied</i>		<i>favorite can win</i>		<i>underdog can win</i>	
	Fav.	Under.	Fav.	Under.	Fav.	Under.	Fav.	Under.
<i>teterogeneity_t</i>	-0.093*** (0.036)	0.004 (0.031)	-0.114** (0.053)	0.004 (0.044)	-0.075 (0.064)	-0.015 (0.060)	-0.066 (0.111)	0.115 (0.116)
E_t [<i>rel. ability_{t+1}</i>]	0.082** (0.034)	0.050 (0.033)	0.116** (0.045)	0.058 (0.046)	0.134** (0.055)	0.038 (0.062)	0.021 (0.090)	0.082 (0.089)
Observations	1578		763		406		215	
R^2	0.156	0.149	0.171	0.167	0.222	0.207	0.293	0.264

The dependent variable for all sub-samples is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$. Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

Table 13: NCAA data: effect of current and future heterogeneity on Effort: stratified by round, r=4

	<i>Favorites</i>			<i>Underdogs</i>		
	round 1	rounds 2-3	rounds 3-5	round 1	rounds 2-3	rounds 3-5
<i>heterogeneity_t</i>	-0.146*** (0.040)	-0.241** (0.114)	-0.061 (0.168)	-0.047 (0.051)	0.095 (0.119)	0.096 (0.208)
$E_t [rel. ability_{t+1}]$	-0.069* (0.039)	0.099** (0.043)	0.103* (0.061)	0.014 (0.037)	0.105** (0.047)	0.133* (0.069)
Observations	352	330	154	352	330	154
R^2	0.255	0.192	0.268	0.126	0.157	0.240

The dependent variable for all sub-samples is defined as $\log[\text{fouls in playoff game}] - \log[\text{avg. fouls regular season}]$. Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, playoff-stage dummies and season dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

Table 14: Effect of current and future heterogeneity on Effort - robustness check, $r = 6$

	<i>Future favorite only^a</i>		<i>probability weighted^b</i>		<i>preferred specification^c</i>		<i>preferred spec.-updated^d</i>	
	Fav.	Under.	Fav.	Under.	Fav.	Under.	Fav.	Under.
$heterogeneity_t$	-0.091*** (0.031)	0.010 (0.027)	-0.088*** (0.031)	0.009 (0.027)	-0.087*** (0.031)	0.009 (0.027)	-0.091*** (0.031)	0.009 (0.027)
$E_t [rel. ability_{t+1}]$	0.091*** (0.029)	0.058** (0.030)	0.086*** (0.029)	0.057** (0.029)	0.086*** (0.028)	0.056* (0.029)	0.090*** (0.028)	0.055* (0.029)
Observations	2199		2199		2199		2199	
R^2	0.149	0.143	0.148	0.143	0.148	0.143	0.149	0.142

The dependent variable for all sub-samples is defined as $\log[\text{avg. fouls regular season}] - \log[\text{fouls in playoff game}]$. Coefficients for additional variables controlling for team specific characteristics are not reported due to space limitations. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, standings dummies and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively. Robust standard errors (clustered for individual playoff-series) in round parentheses.

^a Only the favorite in the parallel interaction is considered as the expected future opponent.

^b The favorite as well as the underdog in the parallel interaction is considered as the expected future opponent. The probability-weighted average of both potential future opponents is defined as the expected future opponent's ability.

^c The preferred specification is the same as used in tables 2 through 12

^d For these specifications we use the same measure for future heterogeneity as in the preferred specification with an additional probabilistic updating of the win percentages of the potential future opponents.

Table 15: Effort and the Probability of Winning for Favorites: Using Ability of Future Opponent as Instrumental Variable

	Win	Win small ^a	IV: Win ^b
<i>effort</i>	-0.017*** (0.002)	-0.015*** (0.004)	0.036 (0.039)
<i>Observations</i>	2199	840	840
<i>R</i> ²	0.168	0.174	0.007
<i>First-stage</i> <i>F-stat.</i>			10.272

The dependent variable is a binary variable equal to 1 if the team wins the game. Robust standard errors (clustered for individual playoff-series) in round parentheses. All specifications include a dummy equal to 1 if team plays at home, a dummy equal to 1 if series is decided in best-of-7 mode with best-of-5 as the base category, playoff-stage dummies, and overtime dummies. *, ** and *** indicate statistical significance at the 10-percent level, 5-percent level, and 1-percent level, respectively.

^a Sample is restricted to all first games of a series and all pivotal games without games where both teams can win the playoff series.

^b 2SLS estimation using the expected future opponent's rating (with r=6) as an instrument.