Selecting the Best? Spillover and Shadows in Elimination Tournaments*

Jennifer Brown
Northwestern University
NBER

Dylan B. Minor
Northwestern University

May 2014

Abstract

We consider how past, current, and future competition within an elimination tournament affect the probability that the stronger player wins. We present a two-stage model that yields the following main results: (1) a shadow effect—the stronger the expected future competitor, the lower the probability that the stronger player wins in the current stage and (2) an effort spillover effect—previous effort reduces the probability that the stronger player wins in the current stage. We test our theory predictions using data from high-stakes tournaments. Empirical results suggest that shadow and spillover effects influence match outcomes and have been already been priced into betting markets.

Keywords: *Elimination tournament, dynamic contest, contest design, effort choice, betting markets.*

*Brown: jen-brown@kellogg.northwestern.edu; Minor: d-minor@kellogg.northwestern.edu. We gratefully acknowledge the helpful comments of Craig Garthwaite, Jin Li, Mike Mazzeo, Chad Syverson, and seminar participants at Stanford GSB, UC-Berkeley (Haas), University of Rochester (Simon), Columbia GSB, Northwestern (Kellogg), University of Cologne, University of Heidelberg, Northwestern-Toulouse Joint Workshop on Industrial Organization, IIOC, LMU's Workshop on Natural Experiments and Controlled Field Studies, Organizational Economics workshop at Queen's School of Business, and the Tournaments, Contests and Relative Performance Evaluation Conference.

Competition for employment and education, innovation funding, and design opportunities can all be framed as multi-stage elimination tournaments in which players are knocked out over successive stages of the event. These contests are often designed to increase player effort—indeed, much of the theoretical and empirical literature focuses on contests as incentive mechanisms. Yet, tournaments may also serve as selection mechanisms, identifying the "best" candidates as overall winners. In labor tournaments where employees' latent talents are not directly observable, firms may organize contests to reveal their workers' relative abilities.[1]

In this paper, we study how the strategies of heterogeneous players in match-pair elimination tournaments are shaped by past, current, and future competition. More specifically, we examine how these intertemporal effects influence a tournament's ability to reveal the strongest player as the winner. Negative spillover from past stages may make current effort more costly and may depress performance, and the shadow of tough future competition decreases a player's expected future payoffs and may also lead to lower current effort. The differential impact of past and future competition across players in a given match changes the effectiveness of tournaments as a selection mechanism. We find that both negative spillover and tough future competition increase the probability that a weak candidate wins.

Our results have practical implications; whether the contest's objective is to encourage effort, select a strong winner, or both, we find evidence suggesting that firms, educators, and other contest designers may need to consider the role of past and future competition in structuring incentives.

In personnel tournaments, workers risk elimination as they advance through corporate management levels. In most contexts, retention of the highest quality worker is most desirable. For example, GE's former CEO, Jack Welch, designed an explicit elimination tournament to select his successor (Konrad 2009). Competition between firms may also be knockout events. In 2010, GE announced a three-stage elimination tournament, the Ecomagination Challenge, to award $200 million to the firm that developed the best smart grid technologies. More commonly, architectural firms may compete for large contracts and investment banks may compete for new clients over several elimination stages. Political races also may involve elimination—a candidate must win his party's primary election to compete in the general election to hold office. Many sporting events are also structured as elimination tournaments.

In each of these examples, effort is clearly important; firms want to hire managers, designers, bankers and innovators who will invest heavily in the activity at hand, voters

[1] In contrast, Lazear (1986) discusses how performance pay may attract higher quality workers *into the firm* when the firm cannot readily observe innate worker ability.

want their representatives to work hard on their behalf, and spectators enjoy high action games. However, selection may also be a prime objective of the contest organizer—a client may desire the most creative design firm, voters may value the most skilled politician, and a board may want the smartest executive to lead the company.

We explore elimination tournaments as selection mechanisms with a two-stage match-pair model. One particular strength of our model is that its predictions are framed in terms of outcomes. As a result, they are testable—in contrast with effort that is notoriously difficult to measure in the field, tournament wins and losses are readily observable. We test our theoretical predictions using the outcomes of high-stakes matches; we exploit the random assignment of players in professional tennis tournament draws. Examining the effect of changes in the skill of the expected competitor in the next round, we find strong evidence of a shadow effect. In addition, spillover in tennis tournaments appears to have a particularly negative impact on the stronger player. We also examine tennis betting markets and find that bookmakers' prices reflect both spillover from past competition and the shadow of future opponents.

The literature on the type of tournament that we study—sometimes called "knock-out" tournaments—begins with Rosen's (1986) model of a multi-stage contest where players have Tullock-style (1980) contest success functions. One significant difference between Rosen (1986) and our current paper is that we use the contest success function that appeared earlier in Lazear and Rosen (1981). More importantly, Rosen (1986) is not focused on shadow and spillover effects; instead, his main result explains the skewed compensation distributions found in many firms. Harbaugh and Klumpp (2005) study a special case of Rosen's model with a single prize and introduce a version of spillover. In contrast to our result that effort in the first stage has a relatively larger impact on the stronger player's probability of success, they model a contest in which low-skill players are disadvantaged in the final round. This result is generated by a set of assumptions that differs from those in our model; in particular, Harbaugh and Klumpp assume that effort is costless and therefore completely exhausted in the final stage, and the total supply of effort is fixed and equal for all players. In their set-up, the stronger player conserves his effort in anticipation of stiff competition in a final stage match against an equally skilled opponent, whereas the weaker player always exerts more effort than the stronger player in the first stage. The weaker player's first-stage exertion "spills over" into the next stage and further reduces his chance of winning the event.

Our effort spillover prediction also relates to previous work on fatigue in dynamic competition. Ryvkin (2011) presents a winner-take-all model where homogeneous players face a binary effort decision and effort has no explicit cost—these features are in stark contrast to our model where players are heterogeneous and effort is a continuous and costly choice

variable in a multi-prize tournament. In his work, fatigue accumulates across stages and players have no opportunity to refresh their effort resources. Among other results, he finds that equilibrium effort is decreasing in fatigue—similar to our notion of negative spillover between tournament stages. Schmitt et al. (2004) study the opposite phenomenon—positive spillover—in rent-seeking contests. They find both theoretical and experimental evidence that positive spillover leads to more first-period expenditure. Our contribution complements and extends these theoretical and experimental results to the field. Moreover, we consider the shadow of expected future competition in addition to the impact of spillover from the past.

Related to our interest in the effect of future competitor, Ryvkin (2009) considers the elasticities of a player's equilibrium effort with respect to the abilities of his opponents across several tournament formats. In elimination tournaments with weakly heterogeneous players, he finds that the abilities of opponents in the more distant future have a lower impact on a player's equilibrium effort than does the ability of the current opponent. While Ryvkin (2009) focuses on players' effort, we are particularly concerned with tournament outcomes.

Several papers have explored the use of tournaments as a selection mechanism. Searls (1963) compares the statistical properties of single- and double-elimination contests and predicts that single-elimination events—the type of tournament that we consider in this paper—are most likely to select the highest ability player as the winner. While our model allows players to make strategic effort decisions in response to past and future competition, Ryvkin and Ortmann (2008) and Ryvkin (2010) compare the selection efficiency of three tournament formats when contestants do not choose effort. In these models, as in the one that we present in the text below, a player's success is probabilistic. In contrast, Groh et al. (2012) model an environment in which heterogenous players choose their level of effort but participate in a perfectly discriminating contest. In an all-pay auction, they explicitly consider various contest designer's objectives, including selection. They find that common seeding rules that match weakest to strongest players in the semifinals maximize the probability that the strongest player wins overall. Clark and Riis (2001, 2007) also examine one-stage, perfectly discriminating contests and explore how various prize rules can improve selection. Modeling a different type of strategic choice, Hvide (2002), Cabral (2003), Hvide and Kristiansen (2003) study outcomes in contests in which competitors choose their degree of risk taking.

Of course, a tournament's ability to select the best is only important when contest designers face a field of heterogeneous competitors. Empirically, Sunde (2009) tests the incentive effect of player heterogeneity in professional tennis tournaments. He finds that heterogeneity impacts the effort choice of the stronger player more than it changes the effort of the weaker

player in a match; for an equal change in rank disparity, the increase in the number of games won by the stronger player is smaller than the decrease in the number of games lost by the weaker player. In addition to the concurrent heterogeneity studied in Sunde's work, we also examine heterogeneity across multiple stages of an event, exploring the incentive impact of ability differences with past, current, and (expected) future opponents. The effects of player heterogeneity on effort in one-shot tournaments has been studied both theoretically (e.g. Baik 1994; Moldovanu and Sela 2001; Szymanski and Valletti 2005; Minor 2011) and empirically (e.g. Knoeber and Thurman 1994; Brown 2011).

The paper is organized as follows: Section 1 presents a two-stage model of an elimination tournament. We derive several propositions and outline the testable hypotheses. In Section 2, we describe our data and empirical strategy for testing these predictions. Section 3 describes the results from past tournaments and discusses the spillover and shadow effects in the context of betting markets. We conclude in Section 4 and discuss the implications of our findings for contest designers.

1 Theory

We study a theory of knockout tournaments in which matches within a given stage are staggered over time. Players in later matches learn the identity of their potential future opponent from outcomes of earlier matches. However, players in these earlier events can only form expectations about their future opponent. This formulation captures both sequential and simultaneous features of competition—just as players in simultaneous matches must form expectations about the outcomes of parallel games, so too must players in early matches of sequential tournaments form expectations about later matches.

This tournament format is often found in practice. For example, in firms, simultaneous promotions to division vice-president may be rare. Instead, the identity of the new appointee is known to other workers still competing for a parallel executive spot—the hopeful workers now know their future opponent for further advancement. This structure is in contrast with other models of elimination tournaments where all matches in a given stage occur simultaneously rather than sequentially (for example, see Stracke (2011)) or all participants compete against each other simultaneously in pools, rather than as pairs (Fu and Lu 2012). We use an additive noise model, as in Lazear and Rosen's (1981) foundational work on one-shot labor tournaments, to focus on the dynamics of a multi-stage elimination tournament.[2]

In the following section, we explore the role of past and future competition on tournament

[2] Additive noise models have been used in much of the labor tournaments literature since Lazear and Rosen (1981); for examples, see Konrad (2009).

outcomes. We present a model that is simple enough to clearly inform our empirical tests, yet rich enough to capture common features of high-stakes, multi-stage tournaments. Specifically, we model an elimination tournament with heterogeneously skilled players competing in sustained competition—one could imagine professionals of varying abilities competing over months or years for a prized position within the firm.[3] For expositional ease, we first present the spillover and shadow effects separately in Section 1.2 and 1.3. Then, in Section 1.4, we present an analysis of the effects operating simultaneously and also enrich the notion of cross-round spillover. Combining the effects does not change the general predictions of the model.

Our theory results describe the probability that the stronger player wins in different stages of the elimination event. These predictions speak directly to our broader research question of "selecting the best." That is, our comparative statics results provide predictions about when the strongest player is most likely to advance to future rounds of competition and, ultimately, win the tournament.

1.1 Model Set-Up

Consider a two-stage elimination tournament with four risk-neutral players, where the players who win in the first stage advance to the final stage. The overall tournament winner receives a prize of V_W, while the second-place competitor receives a prize V_L. Let $V_W > V_L > 0$ and define the prize spread $\Delta V = V_W - V_L$. For simplicity, we assume no discounting across stages. Let player i's total cost be a quadratic function of his effort x_i and his cost type c_i.[4] The convexity assumption on the cost function is common in the literature on tournaments and captures the notion that additional units of effort are increasingly costly for competitors.[5]

For simplicity, we denote a player's cost function as $\frac{1}{2}x^2$ and a player i's total cost as $c_i \frac{1}{2} x_i^2$. We assume that cost types, c_i, vary across all players and are commonly known amongst competitors.[6] We will describe players with relatively low costs as being "stronger" than players with relatively high costs.

[3] In a sports context, our model better reflects the dynamics of an endurance event (e.g. tennis) than competition requiring short bursts of effort (e.g. powerlifting).

[4] One could define a mapping $E : \mathbb{R}_+^N \to \mathbb{R}_+^1$ that collapses levels of N effort-generating activities to the real line. The overall cost of effort is then strictly increasing in the resultant scalar x_i.

[5] Our results hold for more general convex cost functions, $\gamma(x_i)$, provided that both players' effort choices are sufficiently sensitive to a change in marginal benefit. In particular, we require that $\frac{\gamma''(x_1^*)}{\gamma''(x_2^*)} < \frac{c_2}{c_1}$. Versions of the model with linear and other quasi-convex costs also produce similar results.

[6] For ease of exposition, we model heterogeneity through players' cost types. However, several alternative models produce identical results: for example, we can also capture heterogeneity across valuations by defining a player's valuation of the prize as $\frac{V}{c_i}$ or allow the impact of an additional unit of effort on the probability of winning to vary across competitors. It can also be shown that capturing heterogeneity by varying cost function convexity leads to similar results.

Recall that matches in the first-stage are sequential. Assume that players 3 and 4 compete first. Then, player 1 faces player 2 knowing the outcome of the previous match. Players 3 and 4 will form an expectation of the strength of future competition, knowing only the identity of two potential opponents. Without loss of generality, we assume that player 3 won his match against player 4.

1.1.1 Final Stage of the Tournament

Assume that player 1 won his first-stage match. To find the equilibrium of the multi-stage game, we begin by analyzing the strategies of player 1 and his opponent player 3 in the *final* stage. Define player 1's expected payoff function as

$$\pi_{1,final} = P_1(x_1, x_3)\,\Delta V - \frac{1}{2}c_1 x_1^2 + V_L \tag{1}$$

where his probability of winning takes the following form:

$$P_1(x_1, x_3) = \begin{cases} 1 \text{ if } x_1 + \varepsilon_1 > x_3 + \varepsilon_3 \\ \frac{1}{2} \text{ if } x_1 + \varepsilon_1 = x_3 + \varepsilon_3 \\ 0 \text{ otherwise} \end{cases} \tag{2}$$

where $x_i + \varepsilon_i$ is player i's level of output and output is a function of both effort x_i and a random noise term ε_i. In definition (2), the probability that player 1 wins is increasing in his own effort and decreasing in the effort of his opponent.

Define $\boldsymbol{\varepsilon} = \varepsilon_3 - \varepsilon_1$ and let $\boldsymbol{\varepsilon}$ be distributed according to some distribution G such that probability (2) can be written as

$$P_1(x_1, x_3) = \Pr(x_1 - x_3 > \boldsymbol{\varepsilon}) = G(x_1 - x_3) \tag{3}$$

Now, player 1's payoff function (1) can be written as

$$\pi_{1,final} = G(x_1 - x_3)\,\Delta V - \frac{1}{2}c_1 x_1^2 + V_L \tag{4}$$

and his first order condition is

$$\frac{\partial \pi_{1,final}}{\partial x_1} = G'(x_1 - x_3)\,\Delta V - c_1 x_1 = 0 \tag{5}$$

Following Konrad (2009) and Ederer (2010), we assume that G is distributed uniformly

with the following support[7]

$$G \sim U\left[-\frac{1}{2}a, \frac{1}{2}a\right]$$

and, therefore,

$$G' = \frac{1}{a}$$

The assumption that G is uniformly distributed removes the strategic interdependence of players' *current* period effort choices (Konrad 2009).[8] This allows us to isolate the consequences of past effort choices and potential future competition on current-stage effort. In a firm context, this would assume that a worker's optimal effort choice is independent of the identity of his current opponent; of course, in earlier stages, his optimal effort depends on his expectations about future opponents' identities. The results hold broadly if we relax this assumption of same-stage independence and allow players' optimal effort choices to depend on both their current and future opponents.[9]

Rewriting the first order condition (5) yields:

$$\frac{\partial \pi_{1,final}}{\partial x_1} = \frac{\Delta V}{a} - c_1 x_1 = 0$$

which we can rearrange as the following expression:

$$x_i^* = \frac{\Delta V}{a c_i} \text{ for } i = 1, 3 \tag{6}$$

Assume for the remainder of the analysis that player 1 is the stronger player $(c_1 < c_3)$. Then, expression (6) implies that player 1 exerts more effort in the final stage $(x_1^* > x_3^*)$. Therefore, the stronger player is more likely to win in the final stage, relatively to his weaker opponent—that is, the better player is more likely to be "selected" as the overall tournament winner.

In the final round, since both players are guaranteed at least the second prize V_L, only the difference between first and second prize matters to competitors. As expected, a larger

[7] To ensure that probabilities are well-defined, we require that

$$0 < \frac{\frac{\Delta V}{a c_1} - \frac{\Delta V}{a c_3} + \frac{a}{2}}{a} < 1$$

This condition ensures that $G(\cdot) \in (0, 1)$.

[8] Results from Ryvkin (2009) support this simplifying assumption. He shows that, when players' relative abilities are uniformly distributed, a "balanced" seeding can eliminate the dependence of a player's equilibrium effort on his opponent's ability.

[9] A version of the model with more general distributions that allow for same-stage interdependence, including the normal distribution, is available in an online appendix.

prize spread leads to more effort from both players, though the stronger player increases his effort more than the weaker player. Also, increasing the noise around effort (i.e., increasing a, the width of the support of G) reduces effort, particularly for the stronger player.

1.1.2 First Stage of the Tournament

Define z_1 and z_2 as the efforts of players 1 and 2 in the first stage. Player 1's expected payoff function in the first stage is

$$\pi_{1,first} = P_1(z_1, z_2)\tilde{V}_1 - \frac{1}{2}c_1 z_1^2 \tag{7}$$

where $\tilde{V}_1$ is his continuation value (i.e., his payoff in the final stage):

$$\tilde{V}_1(x_1, x_3) \equiv \pi_{1,final} = G(x_1 - x_3)\Delta V - \frac{1}{2}c_1 x_1^2 + V_L$$

Equation (7) yields the first order condition

$$\frac{\partial \pi_{1,first}}{\partial z_1} = \frac{\tilde{V}_1}{a} - c_1 z_1 = 0$$

which we can rearrange, for either player, as the following expression:

$$z_i^* = \frac{\tilde{V}_i}{ac_i} \text{ for } i = 1, 2 \tag{8}$$

Fixing a player's continuation value, his effort z_i^* is decreasing in c_i. Since the continuation value itself is also decreasing in c_i, first-stage effort z_i^* is increasing in a player's ability (decreasing in c_i).[10]

Recall that, at the start of their match, players 1 and 2 already know the outcome of the other first-stage match between players 3 and 4. Of course, this means that players 3 and 4 did not know exactly the identity of *their* future opponent. Instead, we assume that they formed an expectation of their continuation value as follows:

$$\mathrm{E}\left[\tilde{V}_i\right] = p_{1|i}\tilde{V}_i(x_i^*, x_1^*) + \left(1 - p_{1|i}\right)\tilde{V}_i(x_i^*, x_2^*) \text{ for } i = 3, 4$$

[10] It can be shown, using expression (9), that

$$\frac{\partial \tilde{V}_i}{\partial c_i} = -\frac{1}{2}\frac{\Delta V^2}{a^2 c_i^2} < 0.$$

where $p_{1|i}$ is the equilibrium probability that player 1 wins knowing that he will face player i in the final stage.[11] Note that player i cannot influence this probability $p_{1|i}$ because it is a function of the *realized* outcome of the completed match between players 3 and 4. This simplifies our analysis, since player i's first-stage effort z_i does not change this probability $p_{1|i}$. Thus, for players 3 and 4, we can express their effort as

$$z_i^* = \frac{\mathrm{E}\left[\tilde{V}_i\right]}{ac_i} \text{ for } i = 3, 4$$

and the analysis described above for players 1 and 2 applies similarly.

1.2 Shadow of Future Competition

We can use the model to understand the impact of known or expected future competition on the likelihood that stronger players advance to future stages of the tournament—of course, this then influences the likelihood that a high-skill player is selected as the overall winner.

Consider an increase in the skill of the future opponent. This change has the effect of decreasing the continuation value for both players 1 and 2 in the *first* stage. In the following analysis, we show that if player 1 has a lower cost of effort than player 2, then he will decrease his first-stage effort more than player 2.

We can express player i's first-stage effort as

$$z_i^* = \frac{\tilde{V}_i}{ac_i} = \frac{\left(\frac{\frac{\Delta V}{ac_i} - \frac{\Delta V}{ac_3} + \frac{1}{2}a}{a}\Delta V - \frac{1}{2}c_i\left(\frac{\Delta V}{ac_i}\right)^2 + V_L\right)}{ac_i} \tag{9}$$

To identify the effect of a change in the effort cost of the future opponent, we take the derivative

$$\frac{\partial z_i^*}{\partial c_3} = \frac{\Delta V^2}{a^3 c_i c_3^2} > 0$$

Thus, an increase in the skill of the future opponent (i.e. a decrease in c_3) decreases a player's effort in the first stage. This is consistent with Ryvkin (2009) who finds that, in tournaments with weakly heterogeneous players, effort depends negatively on future opponents' skill levels.

Since we are additionally concerned with tournament outcomes, we next ask: Which first-stage player is more sensitive to the change in the future competition? Let $c_1 < c_2$. Then,

$$\frac{\partial z_1^*}{\partial c_3} = \frac{\Delta V^2}{a^3 c_1 c_3^2} > \frac{\Delta V^2}{a^3 c_2 c_3^2} = \frac{\partial z_2^*}{\partial c_3}$$

[11] When player 1 is stronger than player 2, $\tilde{V}_i\left(x_i^*, x_1^*\right) < \mathrm{E}\left[\tilde{V}_i\right] < \tilde{V}_i\left(x_i^*, x_2^*\right)$ for $i = 3, 4$.

This means that, for a given increase in the talent of the future competitor, player 1 decreases his effort even more than player 2. This gives us the first proposition:

Proposition 1 *As the skill of the future competitor in the final stage increases (declines), the stronger player becomes less (even more) likely to win in the first stage and thus less (even more) likely to be selected as the overall tournament winner.*

Figure 1 provides some intuition for the result. Marginal cost and benefit are presented on the vertical axis and effort is shown on the horizontal axis. By definition, the marginal cost of the weaker player lies above the marginal cost of the stronger competitor. In the model, the marginal benefit of effort is always larger for the stronger player; however, for simplicity in the figure, we make the conservative assumption that both players enjoy the same marginal benefit of effort. When the marginal benefit of effort is low, the difference between the stronger and weaker players' efforts is $EffortGap_{Low}$ and when the marginal benefit of effort is high, the difference is $EffortGap_{High}$. When the future competitor is more skilled, both of the current players experience a decrease in their marginal benefit of effort, a move from MB_{High} to MB_{Low}. Since $EffortGap_{High} > EffortGap_{Low}$, players facing a more skilled opponent provide more similar levels of effort, and this reduces the probability that the stronger player wins in the current stage. The reverse is true as the future competitor becomes less skilled; in this case, the gap between current players' efforts increases and this improves the stronger player's chance of success.

A limit argument is also illustrative. Consider a current stage with two differently-skilled players, where the winner advances to face some final typical opponent. Given his superior skills, the stronger player is more likely to win in the current stage. Now, consider what happens when the next round competitor is impossible to beat: Both players in the current stage will exert almost no effort. This hurts the chances of the stronger player, since his likelihood of winning declines towards 50%. In contrast, the weaker player in the current stage sees his probability of success improve towards 50%. The prospect of an impossibly strong future opponent turns the current match into a coin-flip and, as a result, reduces the chances that the stronger player wins. Similar intuition applies to the case where the future competitor changes from unbeatable to typical.

1.3 Effort Spillover

We can also examine effort spillover between stages of the tournament. Spillover can take either a positive or negative form. Positive spillover might reflect learning-by-doing, skill building or momentum within a firm. For example, an innovation team whose proposal

advances to a second stage of funding might benefit from its first-stage experiences, both technical and relational. With positive spillover, second-stage effort is less costly than first stage effort. In contrast, negative spillover might reflect fatigue or reduced resources in later stages. For example, architects competing in design competitions might exhaust their creative resources in early stages and have only limited energy for second-stage proposals. In this case, second-stage effort is more costly than first-stage effort.[12]

Consider a scenario where effort expended by a player in the first stage influences his marginal cost of effort in the final stage.[13] We can rewrite player 1's final-stage payoff as

$$\pi_{1,final} = G\left(x_1 - x_3\right)\Delta V - \frac{1}{2}kc_1x_1^2 + V_L$$

where k reflects the change in total cost induced by first stage effort.

To study a negative spillover effect, we let a player's marginal cost of effort in the final stage increase ($k > 1$) and final stage effort is strictly decreasing in the degree of negative spillover. With positive spillover, a player's marginal cost of effort in the final stage decreases ($k < 1$) and final stage effort is increasing in positive spillover.

Now, equilibrium effort is

$$x_1^* = \tfrac{\Delta V}{kac_1}$$

Straightforward calculations show that negative (positive) spillover reduces (increases) a player's final-stage payoff. Consequently, first-stage effort decreases (increases) with negative (positive) spillover.

Thus, negative spillover implies a lower probability of success in the final stage, holding the opponent's effort and skill constant. Of course, the opposite is true for positive spillover. Since

$$\frac{\partial x_1^*}{\partial k} = -\frac{\Delta V}{k^2ac_1} < -\frac{\Delta V}{k^2ac_3} = \frac{\partial x_3^*}{\partial k}$$

when both players in a match suffer similar negative spillover, the stronger player is more adversely affected. As a result, he is relatively less likely to win. In the limit, $G\left(x_1^* - x_3^*\right) \to 0.5$ as the degree of negative spillover $k \to \infty$.

We summarize this finding in the second proposition:

Proposition 2 *A common proportional increase in effective cost type decreases the probability that the stronger player wins.*

[12] Different notions of spillover have been explored in the literature in settings where players with exogenous, fixed resources make effort allocation decisions over multiple periods of play. For recent examples, see Harbaugh and Klumpp (2005) and Sela and Erez (2013).

[13] If previous effort appears only as a fixed cost in the final stage, we would expect no change in final-stage effort.

Figure 2 illustrates the spillover effect. For both players, negative spillover increases the marginal cost of effort and reduces the levels of effort exerted in the competition. However, a common proportional increase in marginal cost leads to a larger change in effort for the stronger player, relative to the weaker player. Since $EffortGap_{WithoutSpillover} > EffortGap_{WithNegativeSpillover}$, players experiencing negative spillover provide more similar levels of effort, and this reduces the probability that the stronger player wins in the current stage.

Again, a limit argument provides further intuition. Consider a current stage with two players who experience typical levels of negative spillover. Given his superior skills, the stronger player is more likely to win in the current stage. Now, consider what happens when spillover increases dramatically. Facing very high costs, both players will exert similarly low levels of effort. This hurts the chances of the stronger player, since his likelihood of winning declines towards 50%. In contrast, the weaker player in the current stage see his probability of success improve towards 50%. Thus, negative spillover evens the playing field.

Proposition (2) suggests that, with negative spillover, weaker players might support costlier competitive conditions—for example, a weaker player might advocate for more stringent common standards or more difficult tasks. Overall, however, the direction and impact of spillover depends on the context and, thus, is an empirical question.

Our result that negative spillover levels the playing field in both stages is in contrast to Harbaugh and Klumpp's (2005) finding that intertemporal tradeoffs level the playing field for the first stage, but do the opposite in the final stage. Their result is sensitive to the assumptions that effort is costless and that players' total efforts are equally constrained.

Spillover need not be modeled as a common proportional increase in marginal cost; in the next section, we allow the degree of spillover to be a function of first stage effort and achieve similar results.

1.4 Combined Shadow and Spillover Effects

In the text above, we separately present the models of effort spillover and the shadow of future competition; now, we consider these effects simultaneously and allow spillover to be an increasing function of first-stage effort. Combining the effects does not change the general predictions of the previous analysis—the prospect of a stronger future competitor and the presence of negative spillover continue to even the playing field.

1.4.1 Spillover and Shadow - Final Stage

Under this formulation, our first order condition for the final stage yields equilibrium effort choice

$$x_i^* = \frac{\Delta V}{k(z_i)ac_i}$$

where $k(\cdot)$ reflects the degree of spillover from the previous stage and is a strictly increasing function of first stage effort z_i. As expected, greater first-stage effort results in lower effort in the final stage. Further, this effect is amplified for the stronger type since $c_1 < c_2$. The final stage spillover effect is

$$\frac{\partial x_i^*}{\partial k(z_i)} = \frac{-\Delta V}{k(z_i)^2 ac_i} < 0$$

Since $\frac{\partial x_1^*}{\partial k(z_1)} < \frac{\partial x_2^*}{\partial k(z_2)} < 0$, a common level of spillover reduces the disparity between participants' efforts in the final stage, since $\frac{\partial x_1^*}{\partial k(z_1)} < \frac{\partial x_2^*}{\partial k(z_2)} < 0$. As a result, the stronger player is less likely to win in the final stage.

1.4.2 Spillover and Shadow - First Stage

Next, we consider effort decisions in the first stage. We write $k(z_i)$ as k_i to simplify the notation in this section and express player i's payoff as

$$\pi_i = G_{first}(\cdot)\left(\left(\frac{\frac{\Delta V}{k_i ac_i} - \frac{\Delta V}{k_j ac_j} + \frac{a}{2}}{a}\right)\Delta V - \frac{1}{2}k_i c_i\left(\frac{\Delta V}{k_i ac_i}\right)^2 + V_L\right) - \frac{1}{2}c_i z_i^2$$

The first order condition for the first stage is

$$\frac{\partial \pi_i}{\partial z_i} = \frac{\left(\frac{\frac{\Delta V}{k_i ac_i} - \frac{\Delta V}{k_j ac_j} + \frac{a}{2}}{a}\right)\Delta V - \frac{1}{2}k_i c_i\left(\frac{\Delta V}{k_i ac_i}\right)^2 + V_L}{a} + G_{first}(\cdot)\left(-\frac{\Delta V^2}{2k_i^2 a^2 c_i}\right)\frac{\partial k_i}{\partial z_i} - c_i z_i = 0$$

which then gives us the following expression for first-stage equilibrium effort:

$$z_i^* = \underbrace{\frac{\left(\frac{\frac{\Delta V}{k_i ac_i} - \frac{\Delta V}{k_j ac_j} + \frac{a}{2}}{a}\right)\Delta V - \frac{1}{2}k_i c_i\left(\frac{\Delta V}{k_i ac_i}\right)^2 + V_L}{ac_i}}_{\text{shadow effect}} + \underbrace{\frac{G_{first}(\cdot)\left(-\frac{\Delta V^2}{2k_i^2 a^2 c_i}\right)\frac{\partial k_i}{\partial z_i}}{c_i}}_{\text{spillover effect}} \tag{10}$$

With no spillover, the left term is the shadow effect that we described in Section 1.2. Again, the stronger player becomes less likely to win in the first stage as the skill of the

future competitor increases.

The right term reflects spillover. Consider the effect of introducing spillover. Since $G_{first}(\cdot) \geq \frac{1}{2}$ and $c_1 < c_2$, the negative spillover effect is greater in magnitude for the stronger player. This increases the chances that the weaker player wins; thus, spillover has the effect of evening the playing field. That is, *ceteris paribus*, spillover increases the chance of an upset.

1.5 Model Predictions

The theory model outlined above provides the following main predictions:

1. **Shadow of Future Competitors:** The stronger the expected competitor in the next stage, the lower the probability that the stronger player wins in the current stage. Empirically, for a given pair of competitors, we expect that the stronger player is less likely to win when the winner of the current match will face a stronger future opponent.

2. **Effort Spillover between Stages:** Increased negative (positive) spillover decreases (increases) the probability that the stronger player wins in the final stage. Empirically, for a given pair of competitors experiencing negative spillover, we expect that similar levels of past exertion will make it less likely that the stronger player wins.

Although not the focus of the current paper, other predictions follow immediately from our analysis: (a) a steeper prize structure improves the stronger player's probability of success in all stages; (b) the noisier the effort-to-output relationship, the lower the probability that the stronger player wins in either stage; and (c) fixing the competitors' abilities and given a sufficiently large (small) second-place prize, the weaker player's probability of winning is greater (smaller) in the final stage, relative to the first stage. Proofs for these additional results are available from the authors by request.

Note that the model's main implications are framed in terms of outcomes, allowing us to readily test these predictions by observing tournament winners. In the following sections, we describe our data and empirical analysis.

2 Data

Professional tennis offers an ideal environment in which to test the empirical implications of the theory.[14] Tennis events are single-elimination tournaments—only winning players

[14] While tennis tournament organizers may have various objectives beyond selection, it is the *structure* of these tournaments that lends itself to our empirical tests. One would expect tournament competitors to respond to the structure and incentives, not the reason for that contest design.

advance to successive stages until two players meet in the final stage to determine the overall winner. Prizes increase across stages with the largest prize going to the overall winner, and the distribution of prizes is known in advance for all tournaments. The financial stakes are substantial and vary across events—for example, the total purse for the 2009 US Open singles competition was $16 million with a $1.7 million prize for first place, while the total purse for the 2009 SAP Open was $531,000 and the winner received $90,925.

Our empirical analysis exploits the random nature of the initial tournament draw. By ATP rules, the top 20 to 25% of players in an event (the "seeds") are distributed across the draw: the top two seeds are placed on opposite ends of the draw; the next two seeds are randomly assigned to interior slots on the draw; the next four seeds are randomly assigned to other slots; etc. After the seeded players have been assigned, the remaining players are then randomly placed in matches prior to the start of the event.[15] This variation provides the identification for our empirical approach—we can observe the same skilled player compete against a variety of randomly-assigned opponents. For example, in our data, we can observe the fourth best player in the world play against competitors ranked 50th, 100th, and 250th in the first round of the same tournament over different years.[16]

The structure of tennis tournaments is particularly conducive to studying the shadow of future competition—both players (and the econometrician) know the competitors in the parallel match. In some cases, players know exactly who they would face in the next round; in other cases, they can make reasonable predictions about upcoming opponents. Measures of players' abilities are also observable to competitors and researchers—past performance data, as well as world rankings statistics, are widely available.

Data from professional tennis has been used in other research: Walker and Wooders (2001) used video footage and data from the finals of 10 Grand Slam events to identify mixed strategies. Malueg and Yates (2010) study best-of-three contests using four years of data from professional tennis matches with evenly-skilled opponents. They find that the winner of the first set of a match tends to exert more effort in the second set than does the loser and, in the event of a third set, players exert equal effort. Forrest and McHale (2007) use professional tennis bookmaking data and find a modest long-shot bias. Gonzalez-Diaz et al. (2012) use data from US Open tournaments to assess individual players' abilities to adjust their performance depending on the importance of the competitive situation. They find that heterogeneity in this ability drives differences in players' long-term success. Using detailed

[15] Note that the seeding is done according to rank within a tournament; the top seed in one event may have a different ATP ranking than the top seed in another tournament.

[16] Using simulations, we examined the potential for a mechanical relationship between match outcomes and our variables of interest due to tournament seeding rules. Results of the simulations suggest that these rules are not driving our empirical results. Details of the simulation are presented in an online appendix.

data from the men's and women's professional tennis circuits, Gilsdorf and Sukhatme (2008a and 2008b) find that larger marginal prizes increase the probability that the stronger player wins.

2.1 Professional Tennis Match Data

To test the predictions outlined in the theory, we examine the behavior of professional tennis players in 615 international tournaments on the ATP World Tour between January 2001 and June 2010. The data, available at http://www.tennis-data.co.uk, include game-level scores and player ranks for men's singles matches. The four "Grand Slam" events—the Australian, French, and US Opens, and Wimbledon—are included in the data. All of the tournaments are multi-round, single-elimination events played over several days.

Tournament draws may include 28, 32, 48, 56, 96 or 128 players. Of the 615 events in the data, 432 tournaments consist of five rounds of play—rounds 1 and 2, quarterfinals, semifinals, and the final. Six rounds are played in 129 events. Fifty-four tournaments, including the Grand Slam events, consist of seven rounds of play. Most ATP events are best-of-three sets, while the Grand Slam events are best-of-five sets.[17] Depending on the number of competitors, first-round byes may be awarded to the top-ranked players.[18]

World rankings (officially called the South African Airways ATP Rankings) are based on points that players accumulate over the previous 12 months. ATP points directly reflect the pyramid structure of tournaments; more points are awarded to players who advance in top tournaments. For example, a Grand Slam winner earns the maximum points awarded for a single event.[19] ATP rankings are simply a rank-order of all players by their accumulated points. In our analysis, we use the ATP rankings as a measure of players' skill levels.

Table 1 presents summary statistics from 25,758 men's professional tennis matches, reported separately for five-, six- and seven-round events. The stronger player wins approximately 65% of the matches; on average, betting markets predict this outcome in approximately the same proportion. On average, matches are decided after approximately 17 games in five- and six-round tournaments and 28 games in seven-round tournaments, many of which are decided by best-of-five sets.

[17] To win a set, a player must win at least six games and at least two games more than his opponent. A game is won by the player who wins at least four points and at least two more than his opponent. Set tie-break rules vary by tournament.

[18] Byes automatically advance a player to the next round.

[19] For details of the world ranking system, see the ATP World Tour Rulebook, available online at www.atpworldtour.com.

3 Results

In this section, we present empirical tests of the theory predictions. We first examine performance data from professional tennis matches, reporting empirical evidence of both spillover and shadow effects. Next, we ask whether shadow and spillover effects have been priced into betting markets. Although this additional analysis is not a direct test of the theory, it does provide further support for the importance of understanding these phenomena.

3.1 Spillover and Shadow Effects in Match Outcomes

Proposition 1 states that tougher future competition will decrease the stronger competitor's probability of success in the current stage. This prediction follows from the observation that while stronger future competition will cause both players to decrease their effort in the current period, the current effort of the better-ranked player decreases more than the current effort of his worse-ranked opponent. Proposition 2 considers the role of spillover in effort choice and predicts that negative spillover favors the weaker player. The direction of the spillover effect is often an empirical question; however, one might expect negative spillover in events that require intense effort exertion over a short period of time. In professional tennis, players may face a higher cost of effort if exertion in previous matches induced lasting fatigue.

The following specification allows us to study the effects of shadow and spillover simultaneously:

$$\begin{aligned} Strongwins_{mrs} &= \beta_0 + \beta_1 Future_{mrs} + \beta_2 StrongPastGames_{mrs} \\ &\quad + \beta_3 WeakPastGames_{mrs} + \beta_4 Current_{mrs} + \alpha X_r + \gamma Z_s + \varepsilon_{mrs} \end{aligned} \tag{11}$$

where $Strongwins_{mrs}$ is a binary indicator of whether the better-ranked player in match m won in round r of tournament s; $Future_{mrs}$ represents the expected ability of the opponent in the next round; $Current_{mrs}$ represents the degree of heterogeneity in players' skills in the current match; $StrongPastGames_{mrs}$ is the number of games played in all previous rounds of the tournament by the better-ranked player; $WeakPastGames_{mrs}$ is the number of games played in all previous rounds of the tournament by the worse-ranked player; X_r is a matrix of round fixed effects (e.g. first, second, quarterfinal); Z_s is a matrix of tournament-year fixed effects (e.g. 2008 U.S. Open) that capture average event-specific differences (e.g. temperature, purse and media attention), and ε_{mrs} is the error term.

We estimate all equations using a linear probability model (OLS) with a robust variance estimator that is clustered at the tournament-year level to account for correlation in players' performances across matches in a given tournament in a given year. Results are very similar

for a probit specification and are available from the authors by request.

$Future_{mrs}$ is the rank of the stronger of the possible opponents in the next round (i.e. the stronger player in the parallel match).[20] Results are qualitatively similar if we instead use an average of potential future opponents' ranks or a transformation suggested by Klaassen and Magnus (2003).[21] $Current_{mrs}$ is the ratio of the ranks of the worse player and the better player.

We report results for regression (11) by tournament size, separating five-, six- and seven-round events. This accounts for differences in tournament structures—for example, the quarterfinal competitor casts a shadow on the second round in a five-round tournament and the fourth round of a seven-round event; and accumulated spillover in a quarterfinal match in a five-round event may have a considerably different effect than in a seven-round event.

3.1.1 Results: Match Outcomes

Table 2 presents results for the main specification for five-, six- and seven-round events.[22] The coefficient on the shadow ($Future_{mrs}$) is positive and statistically significant in all three regressions ($p < 0.01$). This suggests that the stronger (i.e. better ranked) the future opponent, the lower the probability that the stronger player wins in the current round. For a one standard-deviation decrease in future opponent's rank (i.e. increase in ability), we estimate that the probability that the stronger player wins in the current round decreases by approximately 3.2 to 5.7 percentage points, depending on the tournament size. Given that the probability that the stronger player wins is approximately 65%, on average, a one-standard deviation increase in the shadow is associated with a 3 to 8% decline in the probability of winning.

Coefficient estimates for the two spillover variables take on predicted signs and are statistically significant in all cases ($p < 0.01$). More previous games for the stronger player decreases the probability he wins in the current match, while more previous games for the weaker player increases the chance that the stronger player wins. A one standard-deviation increase in the number of previous games for the stronger player is associated with a decline of approximately 7 to 13 percentage points in his probability of winning in the current match; this represents a 10 to 20% decline. A one standard-deviation increase in the number of previous games for the weaker player is associated with a decline of approximately 4 to 7 percentage points in his probability of winning in the current match; this represents a 6 to

[20] Due to data limitations, the exact sequence of matches is not broadly available.

[21] Klaassen and Magnus (2003) calculate a player's ability as $R = K + 1 - \log_2(rank)$, where K is the total number of rounds in the tournament and $rank$ is the player's tournament seed.

[22] Note that we set both players' previous games to zero for the first round because they experience no spillover, and we exclude final round observations because those players face no shadow.

11% decline in the probability that the weaker player wins.

The history of the stronger player appears to drive his current success more than the history of his opponent—coefficient estimates on the number of games played by the stronger player are larger in magnitude than the coefficient estimates for the weaker player, although the magnitudes are significantly different only for the seven-round tournament ($p < 0.01$). This finding is not surprising, in light of the theory. Exertion from past rounds increases the marginal cost of effort for both players; however, a common proportional increase in marginal cost leads to a larger change in effort for the stronger player, pushing players to provide more similar levels of effort. This reduces the probability that the stronger player wins in the current stage.

As expected, the coefficient estimates on the skill disparity measure for the current match are all positive and statistically significant ($p < 0.01$). This suggests that increased player heterogeneity increases the probability that the stronger player wins. On average, a one-standard deviation increase in the rank ratio—an increase in the disparity between players' abilities—improves the probability that the stronger player wins by approximately 4 percentage points.

3.1.2 Alternative Specifications

Tournament Stakes

Shadow and spillover results are robust to different tournament-level controls. Replacing the tournament fixed effects in our main specification with more detailed controls for the court type, surface type and the natural log of the total tournament purse (in 2010 US dollars) yields shadow and spillover effects similar to those presented in Table 2. Results are presented in Table 3.

Although not the focus of our current paper, this robustness exercise also allows us to assess whether the stronger player is more likely to win in higher stakes events—a prediction that follows from the theory, if increases in the total purse do not materially change the shape of the distribution of prizes. In Table 3, the coefficient estimates on the total prize are positive and statistically significant, consistent with the prediction that larger prizes increase the probability that the stronger player wins ($p < 0.01$).

Long Shadow

Motivated by the simple two-stage model in Section 1, the main empirical specification in Table 2 considers the impact of the expected opponent in the next round of the tournament. It is possible, however, that players respond to a "longer" shadow—in principle, players could look at the full tournament roster and adjust their effort in response to the overall presence of a very highly skilled player in the event. For example, these effort adjustments could take

the form of changes in physically and mentally costly training activities for an upcoming tournament. To consider this possibility empirically, we expand the main specification to include the ATP rank of the most able opponent in the tournament.[23]

Results are reported in Table 3. In all cases, the coefficient estimates for the rank of the most able opponent in the tournament are positive and statistically significant ($p < 0.01$). That is, the presence of a strong player in the tournament—not necessarily even in the next immediate rounds—may lower the probability that the stronger player wins in the current match. The coefficient estimates associated with the more immediate shadow effect are similar to those in the main specification and remain statistically significant ($p < 0.01$). Both the short and long shadows cast by skilled competitors affect the probability that the stronger player wins in the current round.

3.2 Shadow and Spillover Effects in Betting Markets

In this section, we explore whether markets account for the shadow and spillover dynamics of multi-stage competition. Indeed, using data from professional betting markets, we find compelling evidence that subtle spillover and shadow effects have been incorporated into prices.

The efficiency of prediction and betting markets has been studied extensively in the literature; for examples, see the survey by Vaughan Williams (1999). Prediction markets are founded on the argument that by aggregating information, competitive markets should result in prices that reflect all available information (Fama 1970). Therefore, driven by aggregated information and expectations, prediction market prices may offer good forecasts of actual outcomes (Spann and Skeira 2003).

Similarly, betting odds reflect bookmakers' predictions of future outcomes. Betting odds may change as new information becomes available to the bookmaker and with changes in the volume of bets that may be driven by individual bettors' private information. As with formal prediction markets, we might expect betting odds to provide good forecasts. Spann and Skeira (2008) compare forecasts from prediction markets and betting odds using data for German premier soccer league matches. They find that prediction markets and betting odds provide equally accurate forecasts.

To examine whether betting markets incorporate information about shadow and spillover effects, we estimate a regression similar to equation (11) . Now, instead of a binary indicator of the actual outcome, the dependent variable is the probability that the stronger player wins the match as implied by betting markets.

[23] For most players, this is the best-ranked player in the tournament; for matches involving the best-ranked player, the most able opponent is the second best-ranked competitor in the event.

Our data include closing odds from professional bookmakers for pre-match betting.[24] Woodland and Woodland (1999) note that bookmakers adjust odds based on the volume of bets, making the odds available as the betting market closes particularly rich in information. In our analysis, we use the median of the available odds data since the data from no single firm covered all matches.[25] Overall, there was little variation between odds posted by different bookmakers for the same match, perhaps because participants in tennis betting markets tend to be specialists and there is little casual betting (Forrest and McHale 2007).

Table 1 reports the implied probabilities that the stronger player wins across rounds in five-, six- and seven-round tournaments. On average, the stronger player is predicted to win; the betting market favors the stronger player approximately 63% of the time, with slightly more favorable predictions in high-stakes, seven-round tournaments. The accuracy of odds market predictions suggests that information beyond simple rankings are being priced into the market. Between 2001 and 2010, predictions from the market are correct for 69% of the 25,633 matches for which betting data are available. Given that the stronger player actually wins in 65% of the matches, one might not be surprised by this accuracy if the market always predicted that the better-ranked player wins. However, in 19% of the matches, the betting odds imply that the weaker player is expected to win. Interestingly, these market predictions are accurate nearly 63% of the time. That is, betting markets do almost as well predicting an upset as they do predicting a win by the stronger player. This is particularly notable since a naive assessment of the ATP rankings in these matches might suggest that the odds are still solidly against the weaker player; in predicted upsets, the mean rank of the weaker player is 98, roughly 1.7 times higher than his opponent's rank of 59.

3.2.1 Results: Betting Market Predictions

Table 4 reports results for regressions where the dependent variable is the probability that the stronger player wins as implied by the betting market. Overall, coefficient estimates suggest that the betting predictions incorporate information about players' past, current, and expected future competition.

Coefficient estimates for the effect of a stronger future opponent are positive and statistically significant for the three regressions ($p < 0.01$). For a one standard-deviation decrease in future opponent's rank (increase in ability), we estimate that the implied probability that the

[24] Data from 11 betting firms (Bet365, Bet&Win, Centrebet, Expekt, Ladbrokes, Gamebookers, Interwetten, Pinnacles, Sportingbet, Stan James, and Unibet) are included in our main dataset obtained from www.tennis-data.co.uk. Several betting firms also offer in-play betting, but we focus our analysis on pre-match bets only.

[25] We calculate the probability odds from the decimal odds in the original data. Probability odds are 1 / (decimal odds - 1) .

stronger player wins in the current round decreases by approximately 1.1 to 3.2 percentage points; this represents a roughly 3% decline in the implied probability of winning.

Since betting markets close at the start of the match, players' past exertion information is readily available to bookmakers. Indeed, coefficient estimates for the stronger and weaker players' number of previous games are statistically significant and take on the expected signs all cases ($p < 0.01$ in 5 of 6 cases; $p < 0.1$ in 1 of 6). More previous games played by the stronger player is associated with a decrease in the expectation that he succeeds, while more previous games played by the weaker player is associated with an increase in the expectation that the stronger player wins. The magnitudes of these effects also align with predictions from our model—stronger players are more sensitive to an additional unit of spillover, compared to the weaker players.

Greater heterogeneity in players' abilities may increase the market's expectation that the stronger player wins—the coefficient on rank ratio is positive and statistically significant in all cases ($p < 0.01$).

Overall, we find strong evidence that prices in tennis betting markets reflect the shadow and spillover effects predicted by our model.

3.2.2 Unobserved Player Heterogeneity Across Rounds

One advantage of the betting market data is that we can identify things that might otherwise be outside of the econometrician's observation. In particular, we can identify when there is a *predicted* upset—this prediction is based on observations of the bookmaker and not simply the ranks of the players. For example, if a player has a minor injury or seems to be in the midst of a short winning streak, his world rank would not reflect this transient state. However, bookmakers could integrate this information into their predictions about match outcomes.

We can identify predicted upsets by comparing the implied probability of the betting odds to the rank-based outcome prediction (i.e. the prediction that the stronger player is more likely to win). If the betting odds predict that the worse-ranked player has a better than 50% chance of winning, then there is some unobserved (to us) positive shock for him (and/or negative shock for the stronger player). Deviations from the ranked-based predictions that persist over multiple rounds suggest a state-dependent component of play. We take a conservative approach to identify this state-dependence.

There are 2085 predicted upsets in the data, representing roughly 7% of all matches. In 67 cases, a single player was predicted to cause multiple upsets in the same event. Fifty-seven of these instances involved two upsets in the same tournament; ten cases involved three predicted upsets. This means that more than 96% of predicted upsets did not persist

beyond a single round. Overall, we find little evidence that match outcomes are driven by unobserved state dependence.

4 Conclusion

In this paper, we present a two-stage, match-pair tournament model that provides two sharp results: (a) a shadow effect of future competition—the tougher the expected competitor in the final stage, the lower the probability that the stronger player is selected as the winner in the first match; (b) an effort spillover effect—negative spillover has a stronger adverse effect on the higher-skilled player, relative to its impact on the weaker player's probability of success.

We test our two main theoretical hypotheses using data from professional tennis matches. We find evidence of a substantial shadow effect and also identify a negative spillover effect in tennis tournaments. In a second analysis, we use probability odds data from bookmakers to show that betting markets recognize and price in these spillover and shadow effects.

4.1 Implications for the Firm

Our findings have implications in terms of the structure of elimination tournaments. Tournaments are often designed to identify high-ability candidates in environments where the contest organizer cannot readily observe innate talent. In a firm context, our results suggest ways by which a manager can improve the likelihood of promoting the strongest candidate.

Shrouding the skill of a strong future opponent increases players' continuation values, relative to the case where the player faces a stronger rival with certainty. This will elicit more effort, particularly from the stronger player, and improve the probability that the stronger player will win in the current match. Of course, the opposite is true if the contest designer shrouds the identity of a weaker future opponent. Overall, a shrouding policy could elicit more effort (and thus improve the likelihood of selecting a strong winner) in a setting where the future opponent is more likely to be strong, rather than weak. In promotion contests within the firm, a manager who suspects his workers will face a particularly high-skilled competitor should be discouraged from posting explicit information about the skill and identity of this future threat. In practice, a credible shrouding policy could be implemented by always delaying the announcement of winners from parallel competitions.

Limiting negative spillover by allowing competitors opportunities to refresh their resources between stages may also increase the probability that the stronger type wins. For example, in an innovation contest, firms should be given adequate time between stages to

raise additional funds and pursue more advanced technology improvements. Similarly, a firm may wish to institute a "work-life balance" program that promotes employee wellness, discourages career-related burnout, and improves the probability that the firm's labor tournament promotes the strongest workers.

In addition, firms may want to encourage positive spillover through learning. For example, managers could analyze and provide constructive feedback about workers' performances, allowing them to better accumulate skills over stages of the promotion tournament.

References

[1] Baik, K. H. 1994. Effort Levels in Contests with Two Asymmetric Players. *Southern Economic Journal*, 61(2): 3-14.

[2] Brown, J. 2011. Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars. *Journal of Political Economy*, 119(5): 982-1013.

[3] Cabral, L. 2003. R&D Competition When Firms Choose Variance. *Journal of Economics and Management Strategy*, 12: 139-150.

[4] Clark, D., and C. Riis. 2001. Rank-order tournaments and selection. *Journal of Economics*, 73: 167-191.

[5] Clark, D., and C. Riis. 2007. Contingent payments in selection contests. *Review of Economic Design*, 11(2): 125-137.

[6] Ederer, F. 2010. Feedback and Motivation in Dynamic Tournaments. *Journal of Economics and Management Strategy*, 19(3): 733-769.

[7] Fama, E. F. 1970. Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25: 383-417.

[8] Forrest, D., and I. McHale. 2007. Anyone for Tennis (Betting)? *European Journal of Finance*. 13(8): 751-768.

[9] Fu, Q., and J. Lu. 2012. The Optimal Multi-Stage Contests. *Economic Theory*, 51: 351-382

[10] Gilsdorf, K. F., and V. Sukhatme. 2008a. Testing Rosen's Sequential Elimination Tournament Model: Incentives and Player Performance in Professional. *Journal of Sports Economics*, 9(3): 287-303.

[11] Gilsdorf, K. F., and V. Sukhatme. 2008b. Tournament incentives and match outcomes in women's professional tennis. *Applied Economics*, 40(18): 2405-2412.

[12] Gonzalez-Diaz, J., O. Gossner, and B. W. Rogers. 2012. Performing Best When It Matters Most: Evidence From Professional Tennis. *Journal of Economic Behavior & Organization*, 84(3): 767-781.

[13] Groh, C., B. Moldovanu, A. Sela, and U. Sunde. 2012. Optimal seedings in elimination tournaments. *Economic Theory*, 49(1): 59-80.

[14] Harbaugh, R., and T. Klumpp. 2005. Early Round Upsets and Championship Blowouts. *Economic Inquiry*, 43(2): 316-329.

[15] Hvide, H. K. 2002. Tournament Rewards and Risk Taking. *Journal of Labor Economics*, 20: 877-898.

[16] Hvide, H. K., and E. G. Kristiansen. 2003. Risk Taking in Selection Contests. *Games and Economic Behavior*, 42: 172-179.

[17] Klaassen, F. J. G. M. and J. R. Magnus. 2003. Forecasting the winner of a tennis match. *European Journal of Operational Research*, 148(2): 257-267.

[18] Knoeber, C. R. and W. N. Thurman. 1994. Testing the Theory of Tournaments: An Empirical Analysis of Broiler Production. *Journal of Labor Economics*, 12(2): 155-179.

[19] Konrad, K. A. 2009. Strategy and Dynamics in Contests. Oxford University Press.

[20] Lazear, E. P. 1986. Salaries and Piece Rates. *Journal of Business*, 59(3): 405-431.

[21] Lazear, E. P. and S. Rosen. 1981. Rank-Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy*, 89(5): 841-864.

[22] Malueg, D. A., and A. J. Yates. 2010. Testing Contest Theory: Evidence from Best-of-Three Tennis Matches. *Review of Economics and Statistics*, 92(3): 689-692.

[23] Minor, D. B. 2011. Increasing Effort Through Softening Incentives. Working Paper.

[24] Moldovanu, B., and A. Sela. 2001. The optimal allocation of prizes in contests. *American Economic Review,* 91(3): 542-558.

[25] Rosen, S. 1986. Prizes and Incentives in Elimination Tournaments. *American Economic Review*, 76(4): 701-715.

[26] Ryvkin, D. 2009. Tournaments of Weakly Heterogeneous Players. *Journal of Public Economic Theory,* 11(5): 819-855.

[27] Ryvkin, D. 2010. The Selection Efficiency of Tournaments. *European Journal of Operational Research*, 206: 667-675.

[28] Ryvkin, D. 2011. Fatigue in dynamic tournaments. *Journal of Economics and Management Strategy*, 20(4): 1011-1041.

[29] Ryvkin, D., and A. Ortmann. 2008. The Predictive Power of Three Prominent Tournament Formats. *Management Science*, 54(3): 492-504.

[30] Schmitt, P., R. Shupp, K. Swope, and J. Cadigan. 2004. Multi-period rent-seeking contests with carryover: Theory and experimental evidence. *Economics of Governance*, 5: 187-211.

[31] Searls, D. T. 1963. On the Probability of Winning with Different Tournament Procedures. *Journal of the American Statistical Association*, 58(304): 1064-1081.

[32] Sela, A., and E. Erez. 2013. Dynamic Contests with Resource Constraints. *Social Choice and Welfare,* 41(4): 863-882.

[33] Spann, M., and B. Skiera. 2003. Internet-based virtual stock markets for business forecasting. *Management Science*, 49: 1310-1326.

[34] Spann, M., and B. Skiera. 2008. Sports Forecasting: A Comparison of the Forecast Accuracy of Prediction Markets, Betting Odds and Tipsters. *Journal of Forecasting*, 28: 55-72.

[35] Stracke, R. 2011. Multi-Stage Pairwise Elimination Contests with Heterogeneous Agents. Working Paper.

[36] Sunde, U. 2009. Heterogeneity and Performance in Tournaments: A Test for Incentive Effects Using Professional Tennis Data. *Applied Economics*, 41(25): 3199-3208.

[37] Szymanski, S., and T. M. Valletti. 2005. Incentive effects of second prizes. *European Journal of Political Economy*, 21(2): 467-481.

[38] Tullock, G. 1980. Efficient rent seeking. In: J. M. Buchanan, R. D. Tollison, and G. Tullock. (eds.) Toward a theory of the rent-seeking society. College Station: Texas A&M University Press.

[39] Vaughan Williams, L. 1999. Information Efficiency in Betting Markets: a Survey. *Bulletin of Economic Research*, 51(1) January: 1-39.

[40] Walker, M., and J. Wooders. 2001. Minimax Play at Wimbledon. *American Economic Review*, 91(5): 1521-1538.

[41] Woodland, B., and L. Woodland. 1999. Expected utility, skewness and the baseball betting market. *Applied Economics*, 31: 337-345.

Figure 1: Effect of the shadow on effort

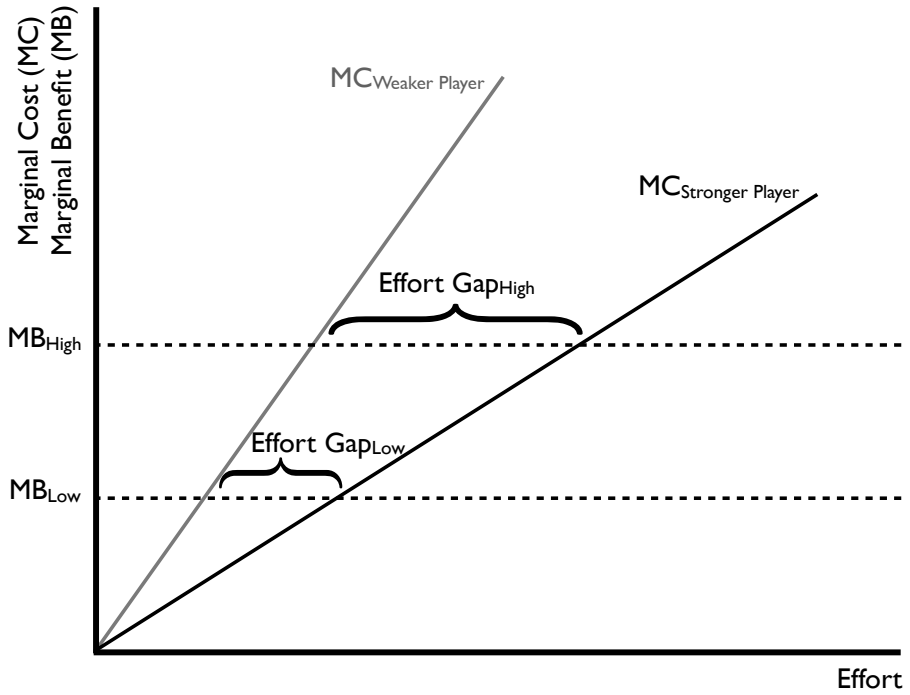

Figure 2: Effect of spillover on effort

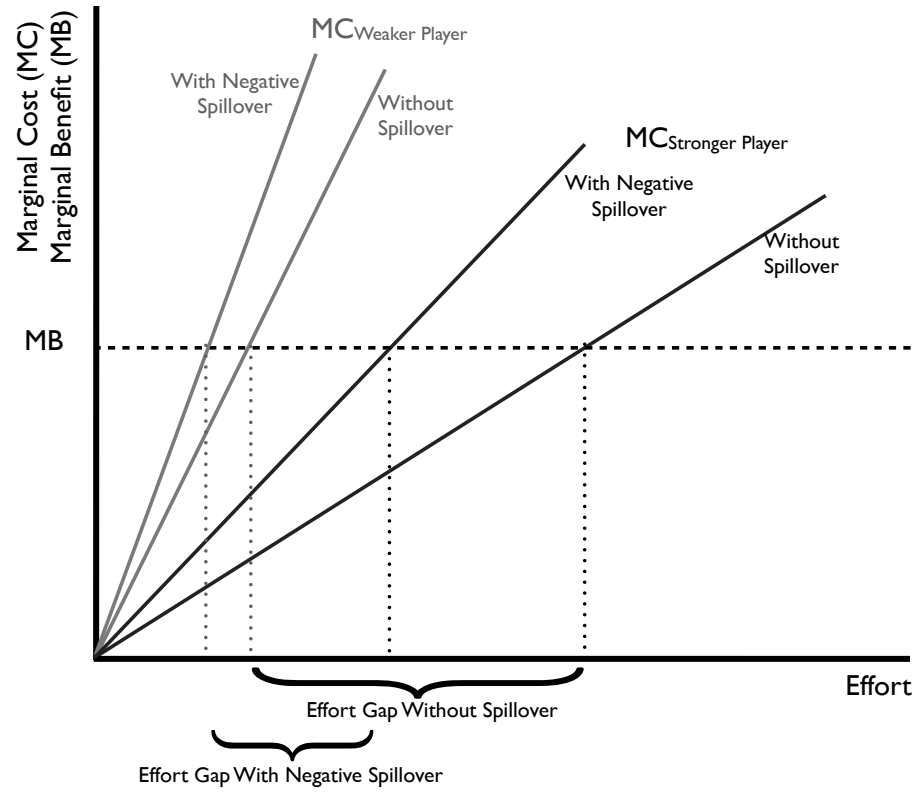

Table 1 - Summary statistics

	Five round tournaments	Six round tournaments	Seven round tournaments
# of tournaments	432	129	54
# of matches played			
Total*	12,758	6,767	6,233
1st round	6,749	2,923	2,903
2nd round	3,435	2,043	1,716
3rd round	-	1,029	858
4th round	-	-	432
Quarterfinal	1,718	516	216
Semifinal	856	256	108
Means			
% of matches in which stronger player wins	64.2% (47.9)	62.9% (48.3)	68.9% (46.3)
Betting market prediction (%) that stronger player wins	62.7% (18.6)	62.4% (18.6)	67.3% (21.1)
Expected future opponent rank	48.35 (41.7)	25.88 (26.1)	29.21 (29.9)
Stronger player's previous games	15.85 (20.9)	14.29 (21.3)	25.43 (37.2)
Weaker player's previous games	16.52 (21.4)	20.75 (24.0)	29.96 (39.4)
Current rank ratio (worse / better rank)	5.79 (15.7)	6.71 (20.7)	9.09 (29.3)

Notes: Values in parentheses are standard deviations. * Matches from the final round of all tournaments are excluded from the counts, means and standard deviation. "Expected future opponent rank" is the rank of the stronger player in the parallel match. "Stronger player's previous games" is the number of games played in all previous rounds of the tournament by the better-ranked player. "Weaker player's previous games" is the number of games played in all previous rounds of the tournament by the worse-ranked player. The number of previous games for the stronger and weaker players is set to zero for the first round of all tournaments. "Current rank ratio" is rank of the worse-ranked player divided by the rank of the better-ranked player.

Table 2 - Actual match outcomes

Dependent variable:	Stronger player wins in current match (0% or 100%)		
	Five round tournaments	Six round tournaments	Seven round tournaments
Expected future opponent rank	0.0762***	0.0844***	0.1737***
	(0.0113)	(0.0292)	(0.0224)
Stronger player's previous games	-0.3342***	-0.3641***	-0.3510***
	(0.0800)	(0.0752)	(0.0615)
Weaker player's previous games	0.1998**	0.2926***	0.1885***
	(0.0774)	(0.0790)	(0.0491)
Current rank ratio (worse / better rank)	0.2747***	0.1818***	0.1323***
	(0.0457)	(0.0650)	(0.0438)
Fixed effects			
Round	X	X	X
Tournament-Year	X	X	X
R-squared	0.05	0.05	0.06
# of observations	12,758	6,767	6,233

Notes: Values in parentheses are robust standard errors, clustered by tournament-year (e.g. 2008 U.S Open). Matches from the final round of all tournaments are excluded. "Expected future opponent rank" is the rank of the stronger player in the parallel match. "Stronger player's previous games" is the number of games played in all previous rounds of the tournament by the better-ranked player. "Weaker player's previous games" is the number of games played in all previous rounds of the tournament by the worse-ranked player. The number of previous games for the stronger and weaker players is set to zero for the first round of all tournaments. "Current rank ratio" is rank of the worse-ranked player divided by the rank of the better-ranked player.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3 - Alternative specifications

Dependent variable: Stronger player wins in current match (0% or 100%)

	Tournament stakes			Long shadow		
	Five round tournaments	Six round tournaments	Seven round tournaments	Five round tournaments	Six round tournaments	Seven round tournaments
Expected future opponent rank	0.0586***	0.0727***	0.1703***	0.0663***	0.0780***	0.1724***
	(0.0107)	(0.0265)	(0.0220)	(0.0110)	(0.0293)	(0.0222)
Stronger player's previous games	-0.3023***	-0.3776***	-0.2883***	-0.2885***	-0.3407***	-0.3177***
	(0.0742)	(0.0637)	(0.0529)	(0.0787)	(0.0750)	(0.0612)
Weaker player's previous games	0.1551**	0.2474***	0.1793***	0.2113***	0.3149***	0.1796***
	(0.0747)	(0.0764)	(0.0488)	(0.0766)	(0.0779)	(0.0486)
Current rank ratio (worse / better rank)	0.2972***	0.1857***	0.1331***	0.1720***	0.1351**	0.0790**
	(0.0438)	(0.0678)	(0.0433)	(0.0347)	(0.0582)	(0.0319)
ln(Total Purse in 2010 US	4.5795***	5.0128***	11.3372***			
	(1.3409)	(1.2389)	(1.7053)			
Rank of best player				2.4543***	3.9033***	20.2632***
				(0.2281)	(0.8130)	(1.7109)
Fixed effects						
Indoor/Outdoor controls	X	X	X			
Surface Type controls	X	X	X			
Round	X	X	X	X	X	X
Tournament-Year				X	X	X
R-squared	0.02	0.02	0.04	0.07	0.06	0.07
# of observations	12,549	6,767	6,233	12,758	6,767	6,233

Notes: Values in parentheses are robust standard errors, clustered by tournament-year (e.g. 2008 U.S Open). Matches from the final round of all tournaments are excluded. "Expected future opponent rank" is the rank of the stronger player in the parallel match. "Stronger player's previous games" is the number of games played in all previous rounds of the tournament by the better-ranked player. "Weaker player's previous games" is the number of games played in all previous rounds of the tournament by the worse-ranked player. The number of previous games for the stronger and weaker players is set to zero for the first round of all tournaments. "Current rank ratio" is rank of the worse-ranked player divided by the rank of the better-ranked player. "Rank of the best player" generally equals the rank of the best player in the tournament; for matches involving the best-ranked player, this variable equals is the rank of the second best competitor in the event.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4 - Betting market data

Dependent variable:	Implied probability that the stronger player wins in current period (%)		
	Five round tournaments	Six round tournaments	Seven round tournaments
Expected future opponent rank	0.0314***	0.0422***	0.1065***
	(0.0048)	(0.0116)	(0.0122)
Stronger player's previous games	-0.1341***	-0.2103***	-0.2224***
	(0.0256)	(0.0308)	(0.0280)
Weaker player's previous games	0.0520*	0.1196***	0.1076***
	(0.0276)	(0.0305)	(0.0244)
Current rank ratio (worse / better rank)	0.2311***	0.1568***	0.1258***
	(0.0332)	(0.0522)	(0.0402)
Fixed effects			
Round	X	X	X
Tournament-Year	X	X	X
R-squared	0.10	0.10	0.12
# of observations	12,634	6,711	6,148

Notes: Values in parentheses are robust standard errors, clustered by tournament-year (e.g. 2008 U.S Open). Matches from the final round of all tournaments are excluded. "Expected future opponent rank" is the rank of the stronger player in the parallel match. "Stronger player's previous games" is the number of games played in all previous rounds of the tournament by the better-ranked player. "Weaker player's previous games" is the number of games played in all previous rounds of the tournament by the worse-ranked player. The number of previous games for the stronger and weaker players is set to zero for the first round of all tournaments. "Current rank ratio" is rank of the worse-ranked player divided by the rank of the better-ranked player.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Online Appendix

"Selecting the Best? Spillover and Shadows in Elimination Tournaments"

May 2014

A General Distribution Case

The model in the body of the paper presents results when the noise in players' output is distributed uniformly; recall that, in Section 1 of the main paper, we define $\boldsymbol{\varepsilon} = \varepsilon_3 - \varepsilon_1$ and assume that $\boldsymbol{\varepsilon}$ is distributed according to $G \sim U\left[-\frac{1}{2}a, \frac{1}{2}a\right]$. Similar results can be derived for any unimodal and symmetric distribution $G(\cdot)$ with mean zero. Again assume that the first-stage matches are resolved sequentially; players 1 and 2 know that player 3 won his parallel match to advance to the final stage.

Player 1's payoff function for the final stage can be written as

$$\pi_{1,final} = G(x_1 - x_3)\Delta V - \frac{1}{2}c_1(x_1)^2 + V_L$$

and his first order condition is

$$\frac{\partial \pi_{1,final}}{\partial x_1} = G'(x_1 - x_3)\Delta V - c_1 x_1 = 0$$

Similarly, player 3's first order condition is

$$G'(x_3 - x_1)\Delta V - c_3 x_3 = 0$$

Since $G(\cdot)$ is symmetric about its mean, it follows that $G'(x_1 - x_3) = G'(x_3 - x_1)$. This implies the following in equilibrium:

$$\frac{x_1^*}{x_3^*} = \frac{c_3}{c_1} \tag{1}$$

Although changes in the prize spread or the noise around players' output affect equilibrium effort, the *ratio* of players' efforts is constant. It follows, for example, that an increase in the prize spread that leads to higher equilibrium effort from both competitors will necessarily increase the absolute spread between players' efforts. In turn, this increases the probability that the stronger player wins in the current stage since his probability of winning is $G(x_1 - x_3)$. In contrast, as equilibrium effort falls—for example, from the adverse effects of negative spillover—the absolute spread between players' efforts decreases. Here, the probability that the stronger player wins declines with equilibrium effort levels.

Since $G(\cdot)$ can be any unimodal symmetric distribution, the impact of changes in the variance of $G(\cdot)$ depends on the exact distribution and its parameters. The top panel of Figure A1 provides an illustration of two PDFs of $G(\cdot)$, both normal distributions centered

at zero with standard deviations of 1 and 2, respectively.

First consider region A. When the players are relatively similar in ability and thus choose similar equilibrium efforts, reducing the variance means a "thickening" of the density. This provides greater incentives for both players, as the marginal return to effort is greater. Therefore, when players are similar in ability, the probability that the stronger player wins increases as the variance decreases.

Now consider region B in which the ability difference between players is substantial and decreased variance means a "thinning" of the density. This weakens incentives for both players, as the marginal return to effort is reduced. Therefore, in this region, decreased variance reduces the probability that the stronger player wins.

Shadow Effect

To study the impact of a change in the ability of the future competitor on first-stage outcomes, we consider the case where player 3 becomes a stronger opponent (i.e., c_3 decreases). Two inequalities are sufficient for player 1 to weakly decrease his effort relative to his current opponent's effort and thus reduce his probability of winning in the first stage: (a) $G'(x_1^* - x_3^*) \geq G'(x_2^* - x_3^*)$ and (b) $G''(x_1^* - x_3^*) \geq G''(x_2^* - x_3^*)$.

When noise is distributed uniformly, both of these weak inequalities hold. For more general distributions, we must analyze these inequalities over several cases. Consider a contest where ε, the difference in players' additive noise terms, is drawn from a normal distribution, illustrated in the bottom panel of Appendix Figure A1. Since the ordering of players' efforts is critical for the analysis, we outline three cases.

Inequality (a): $G'(x_1^* - x_3^*) \geq G'(x_2^* - x_3^*)$

Ordering 1) When $x_3^* < x_2^* < x_1^*$, players 1 and 2 expect to face a future opponent who is weaker than both of them. For example, in the bottom panel of Figure A1, suppose that $x_1^* - x_3^*$ lies at E and $x_2^* - x_3^*$ lies between D and E. Here, we violate inequality (a). However, if player 1 and 2 are sufficiently similar in ability, the increase in player 1's effort can still be greater despite a smaller change in his continuation value since $c_1 < c_2$.

Ordering 2) When $x_2^* < x_1^* < x_3^*$, the future opponent is always stronger than both current players. In this case, it is unambiguous that the stronger player has a greater increase of effort. For example, when $x_1^* - x_3^*$ lies between C and D and $x_2^* - x_3^*$ is to the left of that, we find that $G'(x_1 - x_3) > G'(x_2 - x_3)$.

Ordering 3) When $x_2^* < x_3^* < x_1^*$, the future opponent is stronger than player 2, but weaker than player 1. When $x_1^* - x_3^* > |x_2^* - x_3^*|$, inequality (a) is violated. However, when $x_1^* - x_3^* \leq |x_2^* - x_3^*|$, inequality (a) is satisfied. In Figure A1, inequality (a) is satisfied when $x_1^* - x_3^*$ falls between D and E and $x_2^* - x_3^*$ falls below C.

Inequality (b): $G''(x_1^* - x_3^*) \geq G''(x_2^* - x_3^*)$

To analyze the impact of a change in the future opponent, we focus on the continuation value in the first stage.[1] Recall that

$$\pi_{i,final} = \widetilde{V}_i = G(x_i - x_3)\Delta V - \frac{1}{2}c_i(x_i)^2 + V_L$$

so that

$$\frac{\partial \widetilde{V}_i}{\partial c_3} = -\frac{\partial x_3^*}{\partial c_3} G'(x_i^* - x_3^*)\Delta V > 0$$

where, using the implicit function theorem,

$$\frac{\partial x_3}{\partial c_3}|_{FOC,i} = \frac{x_3^*}{G''(x_3^* - x_i^*)\Delta V - c_3} < 0.$$

Thus, the magnitude of $\frac{\partial x_3^*}{\partial c_3}$ depends on $G''(\cdot)$, the slope of $G'(\cdot)$. The denominator of this expression is the SOC and, therefore, is negative.

Again, we consider the three cases.

Ordering 1) When $x_3^* < x_2^* < x_1^*$, it is ambiguous whether the inequality holds. When player 3 is worse than both 1 and 2, $x_3^* - x_1^*$ and $x_3^* - x_2^*$ both lie on the left-hand-side of the PDF. Although both slopes are always positive, there are cases where $G''(x_3^* - x_2^*) < G''(x_3^* - x_1^*)$ and other cases where $G''(x_3^* - x_2^*) > G''(x_3^* - x_1^*)$.

Ordering 2) When $x_2^* < x_1^* < x_3^*$, it is likely that the inequality holds. For this case, $x_3^* - x_1^*$ and $x_3^* - x_2^*$ both lie on the right-hand side of the PDF. The inequality holds unambiguously in regions in which the PDF is concave in its argument for both $x_3^* - x_1^*$ and $x_3^* - x_2^*$. For the normal distribution in the lower panel of Figure A1, this occurs when both $x_3^* - x_1^*$ and $x_3^* - x_2^*$ lie to the right of the peak and left of the inflection point. When the differences both fall to the right of the inflection point, the inequality does not hold. When $x_3^* - x_1^*$ and $x_3^* - x_2^*$ fall on different sides of the inflection point, the result is ambiguous. In short, as long as x_3^* is sufficiently similar to x_1^*, the inequality holds.

Ordering 3) When $x_2^* < x_3^* < x_1^*$, the inequality is always satisfied. Since $x_3^* - x_1^*$ and $x_3^* - x_2^*$ are located on the left- and right-hand sides of the PDF, respectively, $G''(x_3^* - x_1^*) > 0$ and $G''(x_3^* - x_2^*) < 0$.

In summary, the intersection of both inequalities requires that the future opponent player 3 be not too much less skilled than player 1. Assuming that the spreads between players' abilities are similar across parallel matches, we will observe most often the case where the expected future opponent is more similar in ability to the stronger current player (relative to

[1] Note that $G'(z_1 - z_2) = G'(z_2 - z_1)$ in the FOCs.

the weaker current player). Therefore, the shadow effect is expected in the typical scenarios in which player 3 is either slightly better than both 1 and 2; or worse than player 1 but more similar in skill to player 1 than player 2. For cases in which player 3 is either much worse or much better than the current players, the presence of a shadow effect is possible but not assured by the theory. Empirically, the presence of such ambiguous cases should reduce our ability to identify any shadow effect in the data.

Spillover

In Section 1 of the main paper, we describe negative spillover as increasing players' effective cost types. Assume that two players experience the same level of exertion in the first stage, leading to the same proportional increase in cost types in the final stage. The ratio of their efforts remains unchanged; however, final stage efforts are lower and thus the absolute spread in efforts is smaller, and the stronger player is less likely to win the match. Therefore, as we found in the uniform case, spillover evens the playing field.

Simulation

In Section 2 of the main paper, we briefly describe the seeding rules used to structure professional tournaments. One concern might be that these seeding rules introduce a mechanical relationship in the data—while seeded and unseeded players are randomly matched in the first round, the identity of potential opponents in the future rounds can be constrained by the initial draw. To understand the roles of ATP's seeding protocol, we undertook several simulation exercises and conclude that mechanical relationships are not driving our empirical results.

Simulation Algorithm

The first step of the simulation generates an initial draw for a 32-player tournament using ATP seeding rules. Recall that a 32-player event includes eight seeded players, 24 unseeded players and five rounds of play (1st, 2nd, quarterfinal, semifinal and final rounds). In the first round, eight matches include one seeded player and one unseeded player, and eight matches include two unseeded players.

For a visual description of the seeding task, imagine the draw as a list of 16 pairs of players arranged vertically; these are the 16 matches of the first round. For this size of tournament, the top eight players are seeded according to their rank. In the draw, Seed 1 and Seed 2 are positioned in the very top and very bottom positions, respectively. This positioning ensures that they cannot face each other until the final round. Seed 3 and Seed 4 are randomly assigned to the two most-interior positions, ensuring that they too cannot face each other until the final round and cannot face Seed 1 or Seed 2 until the final round. Seeds

5, 6, 7 and 8 are then randomly assigned to other open matches, positioned in such a way that no seeded players can face another until the quarterfinal round. Finally, the remaining 24 players are randomly assigned to the open positions in the draw.

More specifically, we use the following ATP rules to restrict the (otherwise) random assignment of players. Call the ith seeded player S_i.

- In the first round, all seeded players face an unseeded player.
- Seeded players cannot face each other until the quarterfinal round.
- S_1 and S_2 cannot face each other until the final round.
- S_3 and S_4 cannot face each other until the final round.
- S_5, S_6, S_7 and S_8 cannot face each other until the semifinal round.
- Neither S_1 nor S_2 can face S_3 or S_4 before the semifinal round.
- None of S_1, S_2, S_3 and S_4 can face S_5, S_6, S_7, or S_8 until the quarterfinal.
- After the top eight players are assigned to the draw, all open positions are randomly filled by unseeded players.
- Any seeded player who wins in the first round must face an unseeded player in the second round.

We generate two versions of the 32-player draw and, for each version, simulate 1,000 tournaments to create the simulation dataset.

Simulation Version 1

In the first simulation, we use players' tournament rank (i.e. values 1 to 32) as a measure of player skill and make match outcomes probabilistic.

We use the following rule to determine the winner of each match. The stronger player wins if:

$$61.7 + 0.379 \left(\frac{Rank_{weak}}{Rank_{strong}} \right) > x \tag{2}$$

where $Rank_{strong}$ and $Rank_{weak}$ are the ranks of the stronger and weaker players in the match, respectively, and x is a random draw from a uniform distribution with support 0 to 100. Otherwise, the weaker player wins. The parameters used in expression (2) are obtained from the main ATP dataset for the first round of five-round events. We estimate

$$Strongwins_{mt} = \alpha_0 + \alpha_1 \left(\frac{Rank_{weak}}{Rank_{strong}} \right)_m + \varepsilon$$

where $Strongwins_{mt}$ is a binary indicator of whether the better-ranked player in match m won in a stated round of a tournament and $Rank_{strong}$ and $Rank_{weak}$ are the ranks of the stronger and weaker players in the match, respectively. This estimation yields two coefficients: $\alpha_0 = 61.7$ and $\alpha_1 = 0.379$. These coefficients are used in expression (2) to construct the probabilities associated with players' success in the simulated matches.

With our simulated draws and results, we can estimate the following regression using a linear probability model:

$$strongwins_m = \gamma_0 + \gamma_1 Future_m + \varepsilon$$

where $strongwins_m$ is a binary indicator of whether the better-ranked player in match m won in a stated round of a tournament and $Future_m$ represents the ability of the stronger opponent in the next round, as determined by the initial simulated draw. Note that it is not necessary to include tournament fixed effects because the tournaments are identical in all respects except draws and outcomes.

Concerned with a potential mechanical shadow effect caused by tournament seeding, we focus our attention on the magnitude and statistical significance of γ_1. We adapt the simulation procedure to generate second and quarterfinal data, advancing players according to the simulated outcomes.

Simulation Version 2

The second simulation is identical to the first with one exception: While the first version used players' ranks within the tournament, the second version uses simulated world ranks as the measure of players' skills.

We use the following algorithm to simulate players' world ranks: First, for each tournament, we randomly assign unique world ranks between 1 and 50 for eight players and between 50 and 170 for twenty-four players. We then use players' world ranks to identify the top eight players as tournament seeds. Match outcomes in this version are determined by the following rule:

$$61.7 + 0.379 \left(\frac{WorldRank_{weak}}{WorldRank_{strong}} \right) > x \tag{3}$$

where $WorldRank_{strong}$ and $WorldRank_{weak}$ are the world ranks of the stronger and weaker players in the match, respectively, and x is a random draw from a uniform distribution with support 0 to 100. Otherwise, the weaker player wins. The remainder of the simulation

proceeds as in version 1.

Simulation Results

With the simulation data, we can test whether there is a mechanical relationship between the expected ability of the opponent in the next round and the stronger player's success in the current round. Specifically, we regress an indicator of whether the stronger player wins on the skill measure of the stronger player in the parallel match, as determined by the initial simulated draw.

Results, presented in Appendix Table A1, show that when we include all first rounds matches, there is a small, positive and statistically significant relationship between the indicator that the stronger player wins and the skill of the expected future opponent. This suggests that we could be overestimating the magnitude of the shadow effect in the first round. However, this relationship does not appear in later rounds, relieving concerns of a pervasive mechanical effect. In fact, some estimates of this relationship are negative in later rounds, although not statistically different from zero.

One solution that we propose is to drop matches that include the top two seeds of the tournament, at least in the first round—recall that the top two seeds are the only players who are not randomly assigned to a position in the draw. The results from regressions using simulation data are reassuring. Without the top two seeds, we now do not observe any statistically significant relationship between the indicator that the stronger player wins and the skill of the expected future opponent in any round. That is, we do not observe any mechanical relationship that could be driving our empirical results.

ATP Data Results Excluding the Top Two Seeds

In light of the simulation results, we estimated the main specification using the ATP dataset and excluding matches with the top two seeded players; results are presented in Appendix Table A2 for the match outcomes and Appendix Table A3 for the implied probabilities in the betting market. Coefficient estimates in these specifications are very similar in magnitude to those in the main specification and are identical to the main results in terms the pattern of statistical significance. That is, the shadow effect survives dropping the top seeded players from the analysis and the observed effect does not appear to be driven by the ATP's seeding protocol.

Table A1: Simulation results

Round	Ability measure	Shadow effect	All players	Excluding top two seeds
First	Rank in tournament	Coefficient estimate	0.117	0.041
		Standard error	0.052	0.057
		p-value	0.024	0.473
		# of observations	16,000	14,000
First	Simulated world rank	Coefficient estimate	0.077	0.024
		Standard error	0.017	0.039
		p-value	0.000	0.537
		# of observations	16,000	14,000
Second	Rank in tournament	Coefficient estimate	0.002	0.039
		Standard error	0.084	0.090
		p-value	0.977	0.665
		# of observations	8,000	6,668
Second	Simulated world rank	Coefficient estimate	-0.007	-0.009
		Standard error	0.018	0.020
		p-value	0.711	0.633
		# of observations	8,000	6,654
Third	Rank in tournament	Coefficient estimate	0.070	0.062
		Standard error	0.124	0.137
		p-value	0.575	0.648
		# of observations	4,000	3,080
Third	Simulated world rank	Coefficient estimate	-0.018	-0.017
		Standard error	0.027	0.030
		p-value	0.492	0.579
		# of observations	4,000	3,045

Notes: Estimates are obtained from regressions of an indicator of whether the stronger player wins on the skill measure of the stronger player in the parallel match, as determined by an initial simulated draw. Each simulation run includes 1,000 32-player tournaments.

Table A2 - Actual match outcomes excluding top two seeds

Dependent variable:	Stronger player wins in current match (0% or 100%)		
	Five round tournaments	Six round tournaments	Seven round tournaments
Expected future opponent rank	0.0631***	0.0673**	0.1659***
	(0.0125)	(0.0297)	(0.0238)
Stronger player's previous games	-0.3863***	-0.3252***	-0.4355***
	(0.0934)	(0.0829)	(0.0728)
Weaker player's previous games	0.2054**	0.2678***	0.1886***
	(0.0947)	(0.0928)	(0.0538)
Current rank ratio (worse / better rank)	0.2425***	0.1653***	0.1281***
	(0.0419)	(0.0630)	(0.0423)
Fixed effects			
Round	X	X	X
Tournament-Year	X	X	X
R-squared	0.06	0.06	0.07
# of observations	10,339	5,781	5,706

Notes: Values in parentheses are robust standard errors, clustered by tournament-year (e.g. 2008 U.S Open). Matches from the final round of all tournaments are excluded. "Expected future opponent rank" is the rank of the stronger player in the parallel match. "Stronger player's previous games" is the number of games played in all previous rounds of the tournament by the better-ranked player. "Weaker player's previous games" is the number of games played in all previous rounds of the tournament by the worse-ranked player. The number of previous games for the stronger and weaker players is set to zero for the first round of all tournaments. "Current rank ratio" is rank of the worse-ranked player divided by the rank of the better-ranked player.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3 - Betting market data excluding top two seeds

Dependent variable:	Implied probability that the stronger player wins in current period (%)		
	Five round tournaments	Six round tournaments	Seven round tournaments
Expected future opponent rank	0.0266***	0.0331***	0.1040***
	(0.0054)	(0.0120)	(0.0122)
Stronger player's previous games	-0.1904***	-0.2219***	-0.2535***
	(0.0352)	(0.0357)	(0.0332)
Weaker player's previous games	0.0809**	0.1101***	0.1202***
	(0.0386)	(0.0372)	(0.0270)
Current rank ratio (worse / better rank)	0.2077***	0.1432***	0.1222***
	(0.0303)	(0.0492)	(0.0392)
Fixed effects			
Round	X	X	X
Tournament-Year	X	X	X
R-squared	0.10	0.11	0.12
# of observations	9,898	5,621	5,571

Notes: Values in parentheses are robust standard errors, clustered by tournament-year (e.g. 2008 U.S Open). Matches from the final round of all tournaments are excluded. "Expected future opponent rank" is the rank of the stronger player in the parallel match. "Stronger player's previous games" is the number of games played in all previous rounds of the tournament by the better-ranked player. "Weaker player's previous games" is the number of games played in all previous rounds of the tournament by the worse-ranked player. The number of previous games for the stronger and weaker players is set to zero for the first round of all tournaments. "Current rank ratio" is rank of the worse-ranked player divided by the rank of the better-ranked player.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$