HAL
open science

Heterogeneity and Performance in Tournaments: A Test for Incentive Effects using Professional Tennis Data

Uwe Sunde

To cite this version:

Uwe Sunde. Heterogeneity and Performance in Tournaments: A Test for Incentive Effects using Professional Tennis Data. Applied Economics, 2009, 41 (25), pp.3199-3208. 10.1080/00036840802243789 . hal-00582225

HAL Id: hal-00582225

https://hal.science/hal-00582225v1

Submitted on 1 Apr 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Heterogeneity and Performance in Tournaments: A Test for Incentive Effects using Professional Tennis Data

Journal:	*Applied Economics*
Manuscript ID:	APE-07-0162
Journal Selection:	Applied Economics
Date Submitted by the Author:	03-Mar-2007
Complete List of Authors:	Sunde, Uwe; IZA
JEL Code:	M52 - Compensation Methods and Their Effects (stock options, etc.) < M5 - Personnel Economics < M - Business Administration and Business Econ; Marketing; Accounting, J41 - Contracts: Human Capital, Wage Models, Internal Labor Markets < J4 - Particular Labor Markets < J - Labor and Demographic Economics, J33 - Compensation Packages|Payment Methods < J3 - Wages, Compensation, and Labor Costs < J - Labor and Demographic Economics
Keywords:	Elimination Tournaments, Incentives in Tournaments, Tennis

SCHOLARONE™
Manuscripts

For Peer Review

Heterogeneity and Performance in Tournaments: A Test for Incentive Effects using Professional Tennis Data

Uwe Sunde*
IZA, Bonn

March 3, 2007

Abstract

This paper provides an approach to test whether greater heterogeneity of contestants leads to lower effort exertion in elimination tournaments, as predicted by conventional tournament models. This prediction is difficult to test with real world data because effort is difficult to measure. Based on a simple behavioral model, testable implications are derived and an identification strategy is suggested that allows to test for an incentive effect of heterogeneity even when effort is unobservable. The application uses data from professional tennis tournaments and provides evidence that heterogeneity affects the incentives to exert effort.

JEL-classification: J4, J3, M5

Keywords: Elimination Tournaments, Uneven Tournaments, Incentives in Tournaments, Tennis

*Contact Address: IZA, Schaumburg-Lippe Str. 5, 53113 Bonn, tel: +49/228-3894-221, email: sunde@iza.org.

For Peer Review

1 Introduction

Firms frequently use promotion tournaments to provide their workers with incentives for effort exertion by posting a prize to be won by the relatively best performing contestant. Theory predicts that the total amount of effort to win the prize exerted in the course of a tournament is highest when the contestants are homogeneous (see Lazear and Rosen 1981, Rosen 1986, and Prendergast 1999). Uneven tournaments, where one contestant is *ex ante* stronger than the other(s), result in lower effort exertion of the contestants. Intuitively, the larger the initial disadvantage of the underdog, the more costly it is for him in terms of effort to compensate for the handicap. On the other hand, the favorite knows about his advantage and can reduce his efforts without endangering his prospects of success. The assumption of performance incentives through prizes has received some empirical support, see e.g. Ehrenberg and Bognanno (1990a, 1990b) for evidence from golf tournaments.[1] However, the negative incentive effect of heterogeneity of contestants has so far not been tested with non-experimental data because of the difficulties associated with measuring heterogeneity and effort.[2]

This paper is a first attempt to close this obvious gap in the literature. Testing the incentive effects of heterogeneity with real world data is important in view of the widespread use of tournaments and the fact that heterogeneity of contestants is the rule rather than the exception. I use data from the final two rounds of the most prestigious tennis tournaments for professionals from the Association of Tennis Professionals (ATP) to test the prediction that higher *ex ante* heterogeneity reduces incentives to provide effort. Tennis data are particularly well suited for this, since they perfectly replicate the simple setting of a two-person tournament, allowing for the clearest theoretical predictions. The structure of information available in tennis data closely resembles the structure and requirements of information in theoretical models of one-shot tournaments. At the same time, tennis contests resemble the environment of a corporate tournament, and involve considerable stakes for the players. Prizes are known for each tournament and at each stage. The strength of players can be inferred from their position in the world ranking prior to the respective tournament, allowing to construct a measure of the heterogeneity of the contestants that is generally not available in other data. Rankings and prizes in the data are predetermined and common knowledge before the match begins, which

[1] Several theoretical predictions concerning incentives through prizes have recently been confirmed in empirical studies using firm-level data, see Eriksson (1999) and Bognanno (2001).

[2] Abrevaya (2002) examines the determinants of winning probabilities of *ex ante* unequal players in professional bowling but does not consider effort exertion. Experimental studies dealing with uneven tournaments include Bull, Schotter, and Weigelt (1987), Schotter and Weigelt (1992), and more recently Harbring and Ruchala (2003). None of these directly tests the effect of varying heterogeneity on effort.

For Peer Review

meets the requirements of the theory and avoids problems of endogeneity. Moreover, unlike in the context of intermediate-level firm data, tennis tournaments end with the final. Thus, prize money for reaching and/or winning the final is a good proxy for the prize of winning the elimination tournament implied by matches on the two last rounds of tennis tournaments.

Similar to settings in firms, effort exerted by tennis players is not directly observable by the principal or the audience (including the econometrician), which constitutes the major empirical problem in testing for the existence of an incentive effect caused by player heterogeneity. On the basis of a simple structural model of a tennis match, this paper derives testable implications that can be tested even if effort is not directly observable. The identification strategy rests on the observation that the omitted variable bias caused by unobservable effort has different effects for favorites and underdogs in a match. The empirical results suggest that the hypothesis that heterogeneity has no effect on incentives and effort exertion can be rejected.

The paper proceeds as follows. Section 2 describes the data, section 3 presents a simple statistical model of a tennis match that can be used to derive behavioral predictions, and that serves as basis for the empirical implementation. The section ends by describing the identification strategy used in the application. Section 4 presents the empirical results, and section 5 concludes.

2 Data from Professional Tennis Tournaments

The data are provided by the ATP and comprise information for 156 tournaments of the most and second most prestigious categories, Grand Slam and Master Series, for male tennis professionals for the years 1990 until 2002.[3] In particular, the data stem from the four Grand Slam tournaments per year: Australian Open (Melbourne), Roland Garros (Paris), U.S. Open (New York), and Wimbledon (London). Additionally, the following eight Masters tournaments are included in the data: Cincinnati, Hamburg, Monte Carlo, Montreal, NASDAQ-100 (Miami), Pacific Life (Indian Wells), BNP Paribas (Paris), and Rome. For each tournament and year, the data contain detailed match information including the names of the respective players as well as their current ranking in the ATP Champion's Race, where players are ranked according to their performance in Grand Slam and Masters tournaments as well as their result in the Masters Cup and their five best results from International Series Tournaments in the year previous to the match, the numbers of games and sets won by each player in each respective match, and the identity of the winner of each match.[4] To ensure sufficient random variation in the heterogeneity

[3] Data and relevant links are available on the internet under http://www.atptennis.com.

[4] Missing rankings are replaced by the last available ranking, possibly from the previous year. Note that the lower the rank, the better a player, i.e. the player ranked 1 had the best performance of all tennis professionals

For Peer Review

Table 1: **Summary Statistics**

		Observations (N)	Mean	Std. Dev.
Individual Games won		920	14.67	5.95
Individual Games won per set	P	920	4.84	1.36
Absolute difference in Ranks	HET	460	19.30	25.03
Total Prize Money ($1,000)		156	2,646.24	1,449.48
Prize for winning current round ($1,000)		460	246.29	153.78

of opponents and avoid effects of initial seedings, only data for the last two rounds, that is for semifinals and finals, are used.[5] Eliminating matches with missing information on games won or suspended matches due to injury leaves a total of 460 matches comprising 920 observations for individual players. Table 1 provides some summary statistics.

3 Empirical Model

3.1 A Simple Structural Model of a Tennis Match

Consider the following simple tournament model: two players, a favorite f and an underdog u compete for the (stage) prize \$ that accrues to the winner of a tennis match. The two players differ with respect to their ability a_i, $i = f, u$, to play, where $a_f > a_u$. Both players can choose their effort levels, $e_f \geq 0$ and $e_u \geq 0$, respectively, in order to affect their probability to win the match. The match itself consists of a certain number of balls, and, abstracting from the precise rules for winning a tennis match, the likelihood of winning the match increases in the number of balls won.[6] Let the strength of play of a player during a given ball b be given by S_{fb} and S_{ub}, respectively. This strength of play depends on the players' respective effort levels e_{ib}, their ability a_i and a vector of other characteristics X, including factors like the stage price, the type of soil, season etc. Let us assume for simplicity that this relationship is linear, such that

$$S_{ib} = \alpha a_i + \beta e_{ib} + \gamma' X_i + \varepsilon_{ib} \quad , \quad i = f, u \,. \tag{1}$$

The last term, ε_{ib}, reflects a random component affecting the strength of play during a given ball b. This component of luck can be interpreted as a propensity to commit a mistake, or

over the year.

[5] This setting is conservative as any effects should be biased towards zero if the slacking-off due to heterogeneity is stronger in earlier rounds of the tournament.

[6] As will become clear below, abstracting from the precise technicalities involved, like the crucial role of the number of sets won in determining the match winner, simplifies the analysis substantially without affecting the results too much.

For Peer Review

equally to hit an ingenious ball. Given that there are many ways to commit a mistake and loose or win a ball, it is natural to assume that ε_{ib} is distributed according to a type-1 extreme value distribution, given that it is the first failure in one of many dimensions that leads to a mistake and the end of the ball. For convenience, we assume that ε_{fb} and ε_{ub} are distributed independently, which can be justified by the fact that players commit their mistakes individually and independently from their opponent on the other side of the net (assuming technical fouls are forbidden). The favorite wins a given ball b if

$$S_{fb} \geq S_{ub} \quad \Leftrightarrow \quad \alpha a_f + \beta e_{fb} + \gamma' X_f - \alpha a_u + \beta e_{ub} + \gamma' X_u > \varepsilon_{ub} - \varepsilon_{fb} \,, \tag{2}$$

while the underdog wins it otherwise. Given this structure, $\varepsilon_{ub} - \varepsilon_{fb}$ are independent draws from a log Weibull distribution.[7] The probability to win a given ball b is therefore given by a standard logit model, such that

$$\Pr(\text{favorite wins ball}) = \frac{\mathrm{e}^{\alpha(a_f - a_u) + \beta(e_{fb} - e_{ub}) + \gamma'(X_f - X_u)}}{1 + \mathrm{e}^{\alpha(a_f - a_u) + \beta(e_{fb} - e_{ub}) + \gamma'(X_f - X_u)}} \,, \tag{3}$$

and the associated likelihood given information on an entire match could be estimated as

$$\mathcal{L}_f = \prod_{(\#\text{ balls } f)} \frac{\mathrm{e}^{\alpha(a_f - a_u) + \beta(e_f - e_u) + \gamma'(X_f - X_u)}}{1 + \mathrm{e}^{\alpha(a_f - a_u) + \beta(e_f - e_u) + \gamma'(X_f - X_u)}} \cdot \prod_{(\#\text{ balls } u)} \frac{1}{1 + \mathrm{e}^{\alpha(a_f - a_u) + \beta(e_f - e_u) + \gamma'(X_f - X_u)}} \,. \tag{4}$$

The log odds ratio of the favorite rather than the underdog winning a given ball is consequently given by

$$\ln\left(\frac{\Pr(\text{favorite wins ball})}{\Pr(\text{underdog wins ball})}\right) = \alpha(a_f - a_u) + \beta(e_{fb} - e_{ub}) + \gamma'(X_f - X_u) \,, \tag{5}$$

which is a function of the relative strength of play of both players, $S_{fb} - S_{ub}$.

However, in the application to ATP data, we do not have information on every single ball played, or the associated effort levels, so that this model is not directly applicable. As is explained in the data section, the data are aggregated on the level of an entire match, including the total numbers of games and sets won by each player. Moreover, the odds ratio is of little use in tennis, where the rules of the game not necessarily imply that the player winning the most games wins the match. The model so far does neglect issues like the rules of winning a tennis match, in particular the endogeneity of the duration of a match. For the purpose of estimating a model of effort choice, it is important to control for the inherent endogeneity of the number of sets played, or the duration of the match in terms of the total number of games played. Hence, applying the odds ratio in this context does not seem appropriate, in particular since it is merely instrumental and does not allow to use all individual observations independently for the estimation.

[7] See Kennedy (2003), p. 275–276, for details.

For Peer Review

Consequently, in what follows we modify the framework in order to be applicable to the data at hand. Given the proportions data on the match level, we concentrate on average effort exertion during a match, rather than on the effort exerted to win a given ball. Optimal strategies of a player refer to the effort level for the entire match. This assumption neglects any dynamics of effort exertion during a match. On the other hand, it reflects the usual notion of a one-shot decision prevalent in the (theoretical and empirical) tournament literature (see e.g. Lazear and Rosen, 1981). The probability to win a match increases monotonically in the proportion of games won by a player, rather than the odds ratio of a given point. A sensible implementation has to control for the endogeneity of the duration of the match. This can be done by normalizing the number of games won by a player using information on the number of sets played, and estimating the model separately for favorites and underdogs, allowing for dependencies in the error structure. Such a specification takes into account that, in order to increase their chance to win the tennis match, players have to win as many games as possible, rather than the majority of games. With data on players $i, j = f, u$ meeting in match m in tournament k, a corresponding specification of the relative strength of play during the match that can be directly estimated by OLS is

$$\ln P_{imk} = \alpha(a_{imk} - a_{jmk}) + \beta(e_{imk} - e_{jmk}) + \gamma' X_{mk} + u_{imk} \,, \tag{6}$$

where P_{imk} is the average number of games per set (or match) that player i wins, which is directly related to the average differential in the strength of play, $S_f - S_u$ during the match under observation. This reflects a modified version of a linear probability model to deal with proportions data that uses observations on both players.[8] Potential dependencies between the opponents in a given match are taken into account by allowing the error components ϵ_{imj} to be heteroskedastic and correlated between the players of a given match m.

3.2 Derivation of Testable Hypotheses

That stage prizes provide stronger incentives for effort exertion, while effort provision decreases with the heterogeneity of players is a well-known result in the tournament literature (see e.g. Lazear and Rosen, 1981, Rosen, 1986, and Prendergast, 1999). This subsection derives these results within the framework laid out before, following the treatment of Prendergast (1999). Consider the optimal strategies of a player, given that he can only choose one effort level for the entire match, and concentrate on pure strategy equilibria.[9] For simplicity, abstract from

[8] See Greene (1997, pp. 894-896) and Kennedy (2003, pp. 264-266) for formal derivations.

[9] These assumptions constitute a substantial simplification. In fact, Klaassen and Magnus (2001) show that the probability of winning a point is not independent from who won the previous point, and that some points and

For Peer Review

Editorial Office, Dept of Economics, Warwick University, Coventry CV4 7AL, UK

5

notation for match and tournament, and let the utility of a player be additively separable in the utility derived from winning the stage price and the disutility of exerting effort,

$$u_i = w(\$) - c(e_i)\,, \tag{7}$$

where as before \$ reflects the stage prize, and $w(\cdot)$ is a monotonically increasing function. The cost of exerting effort, $c(\cdot)$, is an increasing and convex function of effort e_i.

The probability of winning the match, and thus the stage prize, is a function $\pi(\cdot)$ that corresponds to the relative strength of play of both players at the match level, analogous to the strength of play during a ball in condition (2). Then, the player's problem is to choose effort in order to maximize expected utility,

$$\max_{e_i} Eu_i = \max_{e_i} w(\$)\pi_i\left(a_i - a_j, e_i - e_j, X\right)) - c(e_i)\,, \tag{8}$$

where π_i is defined as the probability of player i winning the match analogously to (3). As before, this probability can be shown to depend on the relative strength of play during the match. Hence, $\pi(\cdot)$ is a monotone transformation of P_{imk} as expressed in condition (6).[10] Also note that, analogously to (3), the probability $\pi(\cdot)$ is characterized by a logistic distribution. The optimal effort choice e_i^* is characterized by the first order condition satisfying

$$\frac{\partial Eu_i}{e_i} \stackrel{!}{=} 0 \quad \Leftrightarrow \quad w(\$)\frac{\partial \pi(\cdot)}{\partial e_i} \stackrel{!}{=} \frac{\partial c(e_i)}{\partial e_i}\,. \tag{9}$$

Existence and uniqueness is ensured by the properties of the cost function $c(\cdot)$, and the function governing the winning probability $\pi(\cdot)$ of player i. The equilibrium effort allocation is given by the effort of player i, e_i^*, and the corresponding effort level of the competitor player, e_j^*, such that neither player has an incentive to deviate from his choice given the other player's choice and the environment, implying

$$e_i^* = \left(c'\right)^{-1}\left(w(\$)\frac{\partial \pi(\cdot)}{\partial e_i}\right)\,. \tag{10}$$

Condition (9) directly delivers several testable hypotheses concerning the optimal effort choice. Consider first the favorite's effort choice, e_f^*: Everything else equal, the equilibrium effort of the favorite player, e_f, is higher, the higher the stage prize \$; and the lower the ability difference to the underdog, $(a_f - a_u)$.[11] Concerning the underdog's effort choice, a similar set

games are more important to players than others. Similarly Walker and Wooders (2001) provide evidence from ten tennis matches that the observed serving behavior is consistent with an equilibrium in mixed strategies.

[10] In particular, the probability is strictly monotonically increasing and concave in the average number of games per set the player wins.

[11] The first result follows directly from condition (9); the second result follows from the fact that the marginal change in the probability of winning $\partial\pi(\cdot)/\partial e_i$ is equal to the density in the distribution of $\epsilon_u - \epsilon_f$ evaluated at the respective ability differential $a_f - a_u$. Since the error distribution is logistic and the density of the logistic has a mode at 0, equilibrium effort is highest for equally strong opponents with $a_f = a_u$ who then exert the same effort in equilibrium.

For Peer Review

of results can be derived: Everything else equal, the equilibrium effort of the underdog player, e_u, is higher, the higher the stage prize \$; and the lower his (relative) ability $(a_f - a_u)$.[12]

3.3 Empirical Implementation and Identification

The simple framework illustrates how the incentive effects arising from the stage prize and the player heterogeneity rest on the assumption that players can affect their winning probability, and thus their probability of winning the prize, by choosing the effort with which they play. In most previous empirical studies using e.g. data on golf tournaments, effort is observed directly in terms of individual scores, and the tournament winner is determined by who requires the smallest number of hits for all rounds. Most importantly, effort is easily observed for each player independently of other players, and a measure of ability in terms of a ranking is available. In tennis, on the other hand, as in corporate tournaments, a player's effort is hardly observable independently from the opponent, posing a particular conceptual difficulty for testing effects of prizes and heterogeneity in ability on absolute rather than relative performance.

However, it is intuitive that not only relative, but also absolute efforts of both players affect the observed outcomes in terms of games. While a player affects his probability to win the match by his (absolute) effort choice, the probability is essentially a function of relative effort, as is illustrated by the simple framework presented before.

Testing the hypothesis about the incentive effects of stage prizes is straightforward, since the only way that stage prizes can enter the optimization is through effort exertion, illustrated by condition (9). Testing the hypothesis about heterogeneity leading to lower effort provision in uneven tournaments is a little more intricate than testing for incentive effects of prizes. Note that heterogeneity in ability $a_{fmk} - a_{umk}$ can be measured by the difference of the players' ATP ranking before the match. Recalling the empirical specification of strength of play given by (6) and the behavioral model presented before, ability differences enter the individual optimization problem in two ways. The total effect of players' heterogeneity is given by

$$\frac{d \ln P_{imk}}{d(a_{fmk} - a_{umk})} = \frac{\partial \ln P_{imk}(\cdot)}{\partial(a_{fmk} - a_{umk})} + \frac{\partial \ln P_{imk}(\cdot)}{\partial e_{imk}} \frac{\partial e_{imk}}{\partial(a_{fmk} - a_{umk})} \ . \tag{11}$$

A higher superiority of the favorite over the underdog in terms of ability necessarily leads to a larger differential in the strength of play between the two, and therefore to quicker and clearer decisions, that is more games won by the favorite and fewer by the underdog. This is the

[12] Note that for the underdog, condition (9) has to be rewritten analogously, since the probability of winning the match is determined by the inverse strength of play. Again, the first result follows directly from this first order condition; the second result follows from a similar argument as before concerning the change in the winning probability, which corresponds to the logistic density function.

For Peer Review

case even if the efforts of the players are not affected by heterogeneity in their ability. This plain *capability effect* is given by $\partial \ln P_{imk}/\partial(a_{imk} - a_{jmk}) = \alpha$, the first term in condition (11). However, as was shown in the last section, both players exert less effort when the chances of winning are more unequal, as exemplified by a higher superiority of the favorite over the underdog, $(a_{fmk} - a_{umk})$. Using the number of games won per set as outcome variable, agency theory would predict that the total number of games decreases with the heterogeneity of the two players. This is the *incentive effect* of the heterogeneity as predicted by tournament theory, the second term on the right hand side of condition (11).[13] Hence, heterogeneity affects the observable outcome variable through a direct capability effect, the first term on the right hand side of equation (11), and an indirect effort effect, the second term.

Distinguishing capability and incentive effects is important. From the tournament designer's view, the incentive effect is crucial, since it can be affected by the appropriate incentive scheme, while capability is a datum. From a methodological point of view, testing tournament theory essentially means identifying the incentive effect from the plain capability effect, which is only of secondary interest. The main empirical problem arises from the fact that effort, or the difference in effort exertion $e_{imk} - e_{jmk}$, is unobservable in the data. In order to identify whether a potential effect of heterogeneity on the outcome variable merely captures a plain capability effect, or whether it really involves an incentive effect, one can exploit the different impact of heterogeneity on the optimal decisions of favorite and underdog, as is shown now. In particular, the indirect effort effect, the second term of equation (11), is unambiguously negative for both favorite and underdog, given the results from the optimal effort decisions presented before: Both players decide to exert less effort the larger the capability difference between favorite and underdog (i.e. the larger $a_{fmk} - a_{umk}$). On the other hand, the direct capability effect tends to *increase* the relative strength of play of the favorite, $\partial P_{fmk}/\partial(a_{fmk} - a_{umk}) = \alpha > 0$, while it *decreases* the relative strength of play of the underdog, $\partial P_{umk}/\partial(a_{fmk} - a_{umk}) = -\alpha < 0$. Suppose now the conjecture were true that the relative strength of play, reflected by the number of games per set a player wins, were entirely driven by capability differences, not by incentive effects. Then, larger heterogeneity would lead, everything else equal, to more games won by the favorite, and fewer games won by the *ex ante* weaker player, whose deficient capability reduces his winning probability for every single point. If incentive effects also played a role, the relative strength of play of the underdog would still be lower the larger the relative strength of the

[13] Since the direct effect of prizes on observable outcomes, $\partial \ln P_{imk}/\partial \$$ is zero, any effect of \$ on P_{imk} reflects an effort effect, which is supposedly positive, given the previous analysis: One expects $(\partial \ln P_{imk}(\cdot)/\partial e_{imk})(\partial e_{imk}/\partial \$_{mk}) > 0$, since by assumption, effort increases the games won, $\partial \ln P_{imk}/\partial e > 0$, while a higher prize in turn increases effort.

For Peer Review

favorite because both capability and incentive effects work in the same direction. Concerning the favorite the overall effect on the relative strength of play is ambiguous: a larger capability difference would tend to increase relative performance, while an opposing incentive effect that leads to lower effort exertion decreases relative performance. Hence, by looking at favorites and underdogs separately, one can identify whether heterogeneity has an effect on incentives and effort exertion in tennis matches.

Econometrically, this identification strategy can be seen as consequence of the omission of unobserved relevant variables, namely effort, in the estimation. Suppose regressing the number of games per set won by player i, on the explanatory variables of interest, relative ability and relative effort, neglecting for the moment other relevant observable controls X, such that

$$\ln P_{imk} = \alpha_i(a_{imk} - a_{jmk}) + \beta e_{imk} + u_{imk} \,, \tag{12}$$

where $(\tilde{a}_{fmk} - \tilde{a}_{umk})$ is the observed heterogeneity in terms of the rank difference between favorite and underdog. Then it is easy to show that a regression taking only into account the observable relative ability difference, $(a_{fmk} - a_{umk})$, yields biased and inconsistent estimates, with

$$E_{OLS}(\hat{\alpha}_i) = \alpha_i + \frac{\alpha Cov\left((a_{fmk} - a_{umk}), e_{imk}\right)}{Var\left((a_{fmk} - a_{umk})\right)} \,,$$

as long as the incentive effect is present, i.e., $Cov\left((a_{fmk} - a_{umk}), e_{imk}\right) \neq 0$. Note that according to the behavioral model presented before, this incentive effect is negative for favorites and underdogs, $i = f, u$. On the other hand, we have $\alpha_f = -\alpha_u$ as consequence of the direct capability effect. Hence, the omitted variable bias provides an opportunity for a testable restriction imposed by the existence of an incentive effect: any asymmetry in the effects of heterogeneity between favorites and underdogs that blurs the pure capability effect can be interpreted as the consequence of an incentive effect at work.

4 Empirical Application and Results

A direct but somewhat restrictive way of testing for the existence of an incentive effect of player heterogeneity can be obtained by estimating the following empirical model

$$\ln P_{imk} = \alpha + \alpha_1 Fav_{imk} + \alpha_2 HET_{imk} + \beta HET_{imk} * Fav_{imk} + \gamma_1 Z_{mk} + \gamma_2 X_k + u_{imk} \,, \tag{13}$$

where P_{imk} is the average number of games per set (or match) that player i won during match m in tournament k, measuring the strength of play of player i. The variable Fav_{imk} is a binary indicator variable taking the value 1 if player i is the favorite in match m of tournament k,

For Peer Review

and zero if he is the underdog. $HET_{imk} = (a_{imk} - a_{jmk})$ is a measure of heterogeneity of the relative strength of the contestants at the outset of the match. The baseline measure is the difference between the player's own ATP rank and that of his opponent. Note that favorites, that is players with a lower ranking than their opponent, are associated with negative values for HET, while underdogs exhibit positive values. Consequently, if the heterogeneity term exhibits different coefficients for favorites and underdogs the null that no incentive effect is present in the results can be rejected.

Since using the difference of ranks as heterogeneity neglects that quality differences are likely to be larger among low ranked players, i.e. higher among the top ten than among players ranked 91 to 100, I also construct the heterogeneity measure proposed by Klaassen and Magnus (2003), which takes the pyramid structure of quality prevalent in tennis explicitly into account. The measure is based on the round of the tournament in which each player is expected to drop out, and transforms linear ranks into a logarithmic measure of relative strength of player i, $R_i = K - \log_2(RANK_i)$, with K being the number of rounds played in the respective tournament k (i.e. 6 in Masters and 7 in Grand-Slams), and $RANK_i$ being the ATP-ranking. Note that the better ranked a player is, the later he is expected to drop out of the tournament, implying a higher measure R_i. The measure of relative strength of player i against opponent j in match m in tournament k is computed as $HET^{KM}_{imk} = R_j - R_i$. Given that higher R-measures indicate better players, HET^{KM}_{imk}, similar to the plain rank difference, should have a negative effect on individual effort, because it is negative if i is the favorite and positive if i is the underdog. The capability effect predicts a negative coefficient on HET for underdogs, but a positive coefficient for favorites. Z_{mk} contains information on the level of matches, such an indicator for a final as opposed to a semi-final, and X_k are tournament characteristics. In the estimations presented below, I use a full set of tournament and year dummies throughout.[14] Note that all regressors are predetermined before the respective tournament or match starts. Since the performance of the players of a particular match is interdependent, I allow the residuals u to be clustered on the level of a match, and report robust standard errors.

The hypothesis to be tested is that heterogeneity of players does not unfold an incentive effect, and that therefore any observable outcomes are exclusively driven by a capability effect.

[14] To control for effects of the rules of tennis on the number of games, I also added $Winner_{imj}$ as an indicator whether player i eventually lost the match m in tournament k as additional control, even though this does not constitute a predetermined variable. The results are qualitatively unchanged. Experiments using data on characteristics such as soil surface, season, month, outdoor dummy, tournament size in terms of the number of participants, and tournament type (grand slam or not) instead of tournament dummies lead to essentially identical results. While also leading to similar results, I do not report findings when adding the respective rank of the player as additional explanatory variable, since it is not clear why the absolute rank in addition to *relative* strength should matter for the economic performance. Unreported results are available on request.

For Peer Review

Table 2: Evidence for Incentive and Capability Effects for different player categories: Interactions

	Dependent Variable:			
	ln(Games Won per Set)		ln(Games Won per Match)	
Favorite	0.084**	-0.008	0.084**	-0.008
	[0.034]	[0.047]	[0.034]	[0.047]
Heterogeneity	-0.002***		-0.002***	
(own-opponent rank)	[0.001]		[0.001]	
Heterogeneity		-0.056***		-0.045***
(Klaassen/Magnus, 2003)		[0.014]		[0.011]
Interaction Term	0.001*	0.030**	0.001*	0.009*
	[0.001]	[0.013]	[0.000]	[0.005]
Match is final	-0.006	-0.005	0.002	0.003
	[0.017]	[0.017]	[0.006]	[0.006]
Tournament Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
N	920	920	920	920
R^2	0.08	0.09	0.08	0.1

Notes: Robust standard errors are in brackets. ***Coefficient significant at the 1-%-level. **Coefficient significant at the 5-%-level. *Coefficient significant at the 10-%-level. Coefficients for constant terms are omitted for brevity.

A coefficient estimate of β that is significantly different from zero would allow to reject the hypothesis that no incentive effect is present.

The results presented in Table 2 provide support for this conjecture, regardless of whether one uses the games won per set or per match as dependent variable, and regardless of the measure for heterogeneity used: the interaction term is positive and significant. This indicates that, for the same level of heterogeneity, underdogs do significantly better than favorites, implying that a negative incentive effect moderates the positive effect of a stronger ability on the favorites' performance relative to the underdogs.

A less restrictive specification allows all coefficients to be different for favorites and underdogs. Consequently, we also estimate a model like,

$$\ln P_{imk} = \alpha + \beta HET_{imk} + \gamma_1 Final_{mk} + \gamma_2 Z_{mj} + \gamma_3 X_j + \varepsilon_{imk} \,, \tag{14}$$

The OLS estimation results displayed in the left panel of Table 3 show a negative effect of heterogeneity on the outcome variable for underdogs as expected from both capability and incentive effects.

In contrast, heterogeneity has smaller impact on favorites' performance, which is inconsistent with an explanation based exclusively on capability, but consistent with slacking-off due to reduced effort. A possible explanation for this would be that favorites happen to win most of

For Peer Review

their matches, while the rules of the game imply that the number of games won by the winner exhibits less variation than that won by the loser of a match. The effect of heterogeneity on the outcome variable should thus be downward biased for winners relative to that of losers. To test whether this explanation plays a role, I estimate the model separately for underdogs and favorites who lost their match, and whose number of scores is not constrained by the rules of the game. The results are depicted in the right panel of Table 3. Looking at losers who were underdogs, again the effect of individual relative *ex ante* weakness is significant and negative. However, the empirical results show a positive but insignificant coefficient for heterogeneity in the favorites' sample, implying that the *larger* the predominance of a favorite, the *fewer* games he scores. This is only consistent with the incentive effect counteracting the capability effect, with both effects being apparently approximately equally strong. Moreover, simple t-tests show that the difference in the heterogeneity effects on the outcome variable between favorites and underdogs is significant. Results for the log of the share of games won per match rather than per set as outcome variable are qualitatively similar.[15]

5 Conclusion

The empirical results presented in this paper support the theoretical implication that effort exertion in tournaments is highest for homogeneous contestants. In particular, the results provide evidence for an adverse incentive effect due to heterogeneity that counteracts the capability effect for favorites, while both effects reinforce each other for underdogs. While the application used data from professional tennis, the point made in this paper is more general. The widespread use of tournaments in corporations and organizations, the substantial heterogeneity of contestants in these tournaments, and the lack of reliable information about effort exertion emphasize the importance of devising identification strategies that allow the tournament designer to evaluate whether the intended incentives actually affect individual behavior. The approach shown above is a first step in this direction. In future research, it would be interesting to apply the model to (potentially more detailed) data from different sports, contexts and environments.

[15] These results are not reported for brevity but available upon request.

For Peer Review

References

Abrevaya, J. (2002): "Ladder Tournaments and Underdogs: Lessons from Professional Bowling," *Journal of Economic Behavior and Organization*, 47(1), 87–101.

Bognanno, M. L. (2001): "Corporate Tournaments," *Journal of Labor Economics*, 19(2), 290–315.

Bull, C., A. Schotter, and K. Weigelt (1987): "Tournaments and Piece Rates: An Experimental Study," *Journal of Political Economy*, 95(1), 1–33.

Ehrenberg, R. G., and M. L. Bognanno (1990a): "Do Tournaments Have Incentive Effects?," *Journal of Political Economy*, 98(6), 1307–1324.

——— (1990b): "The Incentive Effects of Tournaments Revisited: Evidence from the European PGA Tour," *Industrial and Labor Relations Review*, 43(Suppl.), 74–88.

Eriksson, T. (1999): "Executive Compensation and Tournament Theory: Empirical Tests on Danish Data," *Journal of Labor Economics*, 17(2), 262–280.

Greene, W. (1997): *Econometric Analysis*. Prentice Hall, London, 3rd edn.

Harbring, C., and G. Ruchala (2003): "Separating the Wheat from the Chaff: An Experimental Study on the Sorting Function of Tournaments," *GEABA Discussion Paper 03-13*.

Kennedy, P. (2003): *A Guide to Econometrics*. Blackwell Publishing, Oxford, 3 edn.

Klaassen, F. J., and J. R. Magnus (2001): "Are Points in Tennis Independent and Identically Distributed? Evidence from a Dynamic Binary Panel Data Model," *Journal of the American Statistical Association*, 96, 500–509.

——— (2003): "Forecasting the Winner of a Tennis Match," *European Journal of Operational Research*, 148, 257–267.

Lazear, E., and S. Rosen (1981): "Rank Order Tournaments as Optimal Labor Contracts," *Journal of Political Economy*, 89(5), 841–864.

Prendergast, C. (1999): "The Provision of Incentives in Firms," *Journal of Economic Literature*, XXXVII(1), 7–63.

Rosen, S. (1986): "Prizes and Incentives in Elimination Tournaments," *American Economic Review*, 76(4), 701–715.

For Peer Review

Schotter, A., and K. Weigelt (1992): "Asymmetric Tournaments, Equal Opportunity Laws, and Affirmative Action: Some Experimental Results," *Quarterly Journal of Economics*, 107(2), 511–592.

Walker, M., and J. Wooders (2001): "Minimax Play at Wimbledon," *American Economic Review*, 91(5), 1521–38.

For Peer Review

Table 3: Evidence for Incentive and Capability Effects for different player categories

	Dependent Variable: ln(Games Won per Set) by							
					Losers who were			
	Favorites		Underdogs		Favorites		Underdogs	
Heterogeneity	-0.001**		-0.003***		0.001		-0.002**	
(own-opponent rank)	[0.001]		[0.001]		[0.002]		[0.001]	
Heterogeneity		-0.026***		-0.056***		0.015		-0.030*
(Klaassen/Magnus, 2003)		[0.009]		[0.014]		[0.023]		[0.018]
Match is final	-0.009	-0.012	-0.003	0.002	-0.029	-0.031	0.037	0.044
	[0.030]	[0.030]	[0.036]	[0.035]	[0.049]	[0.049]	[0.045]	[0.045]
Tournament Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	460	460	460	460	183	183	277	277
R^2	0.11	0.11	0.09	0.09	0.24	0.24	0.15	0.15

Notes: Robust standard errors are in brackets. ***Coefficient significant at the 1-%-level. **Coefficient significant at the 5-%-level. *Coefficient significant at the 10-%-level. Coefficients for constant terms are omitted for brevity.

For Peer Review