Doping and competition uncertainty*

Sebastian Bervoets† Bruno Decreuse† Mathieu Faure†

This version: November 2016

Abstract

While a large literature has focused on doping as artificially increasing one's best potential performance, an ignored aspect of doping is that it also reduces performance volatility (a "regularity" effect). We develop a model with heterogeneous players and find that doping/cheating becomes a game of strategic heterogeneity. All strategic interactions are driven by the regularity effect and not by the increase in best potential performance. In equilibrium, better players cheat more than less talented players; competition uncertainty decreases with doping costs; better players are better off when cheating is possible and less talented players are hurt.

Keywords: Game theory; Strategic heterogeneity; Economics of doping

1 Introduction

This paper focuses on doping behavior and the volatility of sports competition outcomes. It is motivated by three sets of observations. First, doping is very widespread and firmly grounded in high-level sports' culture. Ulrich et al (2016) show that between 29% and 45% of athletes present in their sample report past-year doping. Concerns about doping have been a central focus of the recent Olympic games in Rio 2016, where among others, Russian athletes were not allowed to compete.

*We thank Gérard Dine as well as the many people we interacted with on this project and seminar participants in Aix-Marseille University and Linnaeus University. We also thank Laurence Bouvard for her excellent research assistance. Special thanks to Patrick François and Yves Zenou for their advises.

†Aix-Marseille Univ (Aix-Marseille School of Economics), CNRS & EHESS. Centre de la Vieille Charité, 2 rue de la Charité, 13002 Marseille, France. The authors thank the French National Research Agency (ANR) for their support through the program ANR 13 JSH1 0009 01, as well as the Fondation Aix-Marseille University and its program "Sport, santé et développement durable".

During the paralympics that followed, handicapped athletes were caught "boosting"[1]. Cycling has long been plagued by doping scandals. The Balco case in the US revealed how athletes in baseball, basketball, American football and athletics using prohibited drugs to enhance performance; similarly evidence has been found in soccer, tennis, swimming, skiing, horse riding etc. of athletes using performance enhancing drugs (PEDs).

Doping scandals also involve sports superstars. The World anti-doping agency reports that since 2000, more than 50 Olympic medalists have lost their medals for doping. This culminated when a former seven-times winner of the Tour de France Lance Armstrong confessed to doping in 2013. These scandals feature top athletes using sophisticated doping programs involving medical doctors, high-tech doping agents and methods, and spending up to hundreds of thousands of dollars on the endeavour.

At the same time, high-level and professional sports competitions generally have highly predictable outcomes. In tennis, only seven players won at least one of the four grand slam tournaments over the decade 2006-2016. Between 2000 and 2015, the four most competitive soccer championships (UK, Spain, Italy and Germany) have seen a very limited number of teams ascending the podiums; over this period only 7 different teams account for the first three positions in the UK, 8 in Spain, 7 in Italy and 9 in Germany. In cycling, the Tour de France history reveals a list of multiple-times winners. In each of these sports the list of contenders is also very small.

The aim of this paper is to develop a new model on the general issue of cheating in contests that naturally suits the problem of doping in sports and that links these sets of observations. We consider a model where players are heterogeneous and performance is subject to uncertainty. Cheating is a continuous choice variable that has two characteristics: it increases a players' best potential performance (referred to as the *standard effect*) and the probability of performing at one's best level (referred to as the *regularity effect*).

Doping in a game theoretic setting has been previously analyzed. However, existing models do not reconcile the three observations made above. We argue that this is because the economic literature has focused on the standard effect only and we show that introducing the regularity effect leads to drastically different conclusions. The standard literature models cheating as a game with strategic complements (i.e., the individual cheating incentive increases with the opponents' cheating levels), thus contests are analyzed as typical cases of the prisoner's dilemma. This leads to three main predictions: less talented players cheat more; the possibility of cheating does not affect competitions' outcomes; and better players are unhappy that cheating pos-

[1]Boosting refers to a doping technique by which athletes mutilate some parts of their body in order to overstimulate the nervous system. Handicapped athletes use boosting on parts of their body that are insensitive to pain.

sibilities exist[2]. In contrast, we show that with the regularity effect, more talented players cheat more; as a consequence, competition uncertainty decreases when players dope; finally, better players are happy that cheating possibilities exist because this helps them secure the top positions.

We proceed in three steps. We first present a case study associating doping with reduced competition uncertainty (Section 2). We show that doping agents and methods always increase top performances (the standard effect) and concentrate the distribution of individual performances around their optimal levels (the regularity effect). We then focus on a particular sport, cross-country skiing (CCS), and a particular doping technology, synthetic erythropoietin (EPO). CCS puts the focus on endurance skills, which magnifies the regularity effect of PEDs. Moreover, there have been a number of doping scandals and medical studies provide evidence of blood manipulation in the 1990s and 2000s. Unlike cycling, teammates play a small role in individual ranking and there is a single goal in each race, i.e., being ranked as high as possible.

We compute the Spearman correlation between race-specific rankings and the final ranking for each year between 1987 and 2006, and observe two main facts. First, a large decline in the volatility of competition outcomes occurred in the late 1980s-early 1990s, coinciding with the introduction and spread of EPO. Right before then, surprising outcomes happened from time to time, with lesser ranked athletes finishing ahead of higher ranked athletes. These surprises disappeared immediately after. Though this increase in the regularity of competition outcomes could be explained alternatively - for instance, improved technologies and better training methods could differentially favor better athletes - we then observe a reversal in the early 2000s. Such an increase in competition outcome volatility is incompatible with these alternative explanations but coincides precisely with EPO detection tests and blood profiling, which made doping much riskier.

We develop a model with heterogeneous players and performance uncertainty (Section 3). Each player is in one of two states, good or bad. Cheating is a continuous and costly variable, which provides both the standard effect and the regularity effect. We show that any Nash equilibrium of the game with both effects is also an equilibrium of the game with only the regularity effect. This is because the ex-ante ranking of players in terms of their best potential performance is unchanged when players are allowed to cheat. This important result implies that the standard effect is not what drives the strategic interactions between players.

We then focus on the game featuring the regularity effect only (Section 4) and

[2]See, e.g., Bird and Wagner, 1997, Berentsen, 2002, Berentsen and Lengwiler, 2004, Kräkel, 2007, for an overview of models, and Eber and Thepot, 1999, Berentsen et al., 2008, Curry and Mongrain, 2009, for policy implications. These papers do not always feature the prisoner's dilemma. However, they all predict that anti-doping policies are Pareto-improving. In particular, the best player always prefers a world where doping is banned. In the web appendix B we present a unified framework for these models and detail these predictions.

show that it becomes a game with strategic heterogeneity: cheating acts as a strategic complement with the cheating levels of less talented agents and as a strategic substitute with the cheating levels of more talented agents. We show that there always exists a Nash equilibrium in pure strategies. Uniqueness is not guaranteed when there are at least three players, because at least one of them faces both a better opponent and a worse one. Thus this player has both strategic complements and substitutes, which is source of multiplicity.

Despite the multiplicity of equilibria, implications are clear since all equilibria share similar features. First, cheating increases with talent, because the regularity effect makes the return to cheating higher for better players. Second, less talented players suffer from the existence of cheating possibilities, while the more talented benefit from them. As they become more likely to compete at their best level, they outperform the others more often, pay more because they cheat more, and fare better overall. Competition outcomes become almost deterministic when doping costs approach zero, while uncertainty increases as doping costs increase.

These results link the observations discussed above: doping is widespread, it affects superstars and decreases competition uncertainty. They also explain why stamping out cheating is so difficult: anti-cheating policies are hard to implement because they are not supported by the best players who most benefit from cheating. Our results inform a widespread debate: many observers suggest that authorities should give free access to doping. It is argued that this will level the playing field and restore competition uncertainty. Our model suggests precisely that the opposite will happen. If offered the possibility of using PEDs, the best athletes will use them maximally, thereby further reducing less talented players' chances of winning. Competitions would degenerate into deterministic events.

Our model applies to any type of contest where performance is subject to uncertainty and cheating technologies exist. For instance, it can help understand the incentives of fund managers engaged in fraudulent accounting, of students to cheat on exams, or of scientists tempted to alter datasets. In these aspects, our paper relates to the tournament literature initiated by Lazear and Rosen (1981). This literature studies market situations where payoffs depend on relative performance. This framework naturally applies to sporting contests, but also to labor markets where it is easier to rank workers than measure their individual performances. Equilibrium efforts reflect individual skills and player heterogeneity through differential access to top positions and related economic incentives (see, e.g., Rosen, 1986, for theoretical arguments, Glisdorf and Sukhatme, 2008, and Sunde, 2009, for applications to tennis). Our contribution is to propose an explicit scenario that governs the pattern of strategic complementarity and substitutability across player types.

All the proofs are relegated to the appendix in section 6.

2 Motivating facts

This section examines the effects of PEDs on individual performances and presents our case study linking doping to competition uncertainty. We first argue that PEDs have two kinds of effects on the distribution of athletes' performances: they increase the maximum performance and concentrate the distribution around this maximum. We then focus on the CCS World Cup, a yearly competition based on 10 to 25 races. We examine the yearly race-specific rankings between 1987 and 2006 and show that rankings were relatively volatile before EPO was introduced. They then became almost deterministic in the 1990s when the use of EPO was widespread, finally becoming more volatile again after measures against EPO were introduced. We also document an increase in volatility in the late 1990s-early 2000s, right after the introduction of upper limits on hemoglobin concentration ([Hb]).

2.1 Performance-enhancing drugs and their effects

We document the two effects of doping agents and methods: the standard effect that increases the maximum performance of athletes and the regularity effect that allows them to perform more often at their best level. To illustrate, consider a professional cyclist climbing a mountain. The mean ascent speed is a random draw on some speed interval $[a, b]$. PEDs shift the upper bound b to the right - the standard effect - and they assign more weight around the upper bound - the regularity effect.

The standard effect is well documented. PEDs improve basic skills like strength or endurance. The regularity effect, however, is rarely mentioned as a key factor in understanding doping behavior. It comes from the fact that PEDs also improve recovery, decrease injury risk and duration, reduce tiredness, allow for longer training periods, etc. All of these effects reduce the odds of having a bad day and facilitate the repetition of excellent performances.

Table 1 is based on the World Anti Doping Agency (WADA) list of prohibited products and methods. This list classifies doping products into six categories depending on their biological mechanisms and also classifies two doping methods. In each case, we describe doping effects from both angles: impact on maximum performance (standard effect), and impact on the odds of having a bad day (regularity effect)[3].

Table 1 carries a key message. Almost all doping agents and methods simultaneously increase maximum performances (which is well known) but they also reduce the odds of having a bad day. This is of course true of blood doping, i.e., EPO and transfusion methods. It also holds for anabolic agents: athletes who use such agents improve their strength beyond their physiological potential. However, anabolic agents also improve recovery and reduce injury risks. These effects are

[3]We thank Gerard Dine, hematologist and expert for the French Anti-Doping Agency, for his valuable explanations.

Category	Products	Max perf	Chances of bad day
S1 Anabolic agents	Exogenous anabolic androgenous steroids	Increases strength	Decreases injury risk
	Endogenous anabolic androgenous steroids	Increases strength and tonus	Improves recovery reduces tiredness
	Other anabolic agents (non-steroids)	Increases strength	Improves recovery
S2 Peptide and growth hormones	Erythropoietin (EPO)	Increases endurance	Improves recovery reduces tiredness Allows for better training
	Human Growth Hormone	Increases strength	Reduces injury risk and injury duration Allows for better training
S3 Beta-2 agonist	Anti asthma treatments	Improves breathing (small doses) / Increases strength (high doses)	Improves recovery
S4 Hormone and metabolic modulators	Regulation factors	Improves HGH effects	Improves HGH effects
	Insulins	Improves strength and Endurance	Improves recovery
	Metabolic modulators (GW 15156 - AICAR)	Reduces weight/ Increases strength	Reduces tiredness
S5 Masking agents	Diuretics, glycerol…		
M1 Manipulation of blood	Auto, hetero, Homotransfusion	Same as EPO	Same as EPO
	Platelet Rich Plasma	None	Reduces injury duration

Table 1: Doping agents and methods and their effects. Column 1 and 2: WADA's categories and list of products. Column 3 and 4: The two effects of each product based on expert opinion.

similar to those of growth hormones and metabolic modulators.

The regularity effect is usually neglected, as the case of Platelet Rich Plasma (PRP) illustrates. PRP treatment involves extracting some of the athlete's blood and enriching it with platelets, a source of human growth factors. The blood is then reinjected into the athlete's body, and this blood manipulation helps the athlete recover faster from a muscle injury. Up to 2010, the side effects of such treatment on athletes' performances were unknown and WADA consequently banned the use of PRP. However, since 2010 medical studies have shown that PRP has no effect on the maximum potential performance of an athlete, leading WADA to authorize its use. The fact that PRP can also be used to decrease risk of injury and injury duration was not taken into account. But reducing the risk of injury clearly reduces an athlete's chances of having a bad day. WADA's position on PRP shows how institutions focus mainly on how drugs affect maximum performance, underestimating regularity of performance as a potential target for drug users.

2.2 Evidence from cross-country skiing

The CCS World Cup is organized by the Fédération Internationale de Ski (FIS). Skiers receive points based on their ranking and the final ranking is obtained by adding the points collected in each race. We use FIS data on individual rankings in different races and compute for every year how race-specific rankings are correlated with the final ranking.

A brief history of doping in CCS.—Figure 1 presents the changes in the anti-doping environment, as well as major events related to the CCS World Cup. Data are not available before 1987 and the sport was organized differently after 2006 with the creation of super races lasting several days like the Tour de Ski. These changes preclude any yearly comparison, so we restrict our attention to the period 1987-2006.

As in cycling, EPO was made available to CCS athletes in the late 1980s. No tests existed. Given the huge effects on endurance athletes' performances, EPO was widely used in the 1990s. Videman et al (2000) document the change in blood profiles observed for CCS athletes, from 1987 to 1999. The profiles show a large increase in mean [Hb] from 1989 to 1996. In the 1996-1997 season, an upper bound was imposed on [Hb], and the threshold became tighter the following season. Correspondingly, Videman et al. (ibid) show a decline in max [Hb]. However, mean [Hb] continued to increase until 1999.

Between 1999 and 2001, the anti-doping environment was characterized by uncertainty. A urine EPO test had been around for years, but it was not efficient. A test for a plasma expander and EPO-masking agent, hydroxyethyl starch or HES, became available in 2001, but athletes were not aware of this. At the 2001 Lahti World Championships, several top skiers in the Finnish team tested positive. In 2002, a test for Darbepoetin Alfa led to the disqualification of Larisa Lazutina and Olga Danilova of Russia and Johann Mühlegg of Spain from their final races in the Winter Olympic Games. Blood profiling was introduced in 2002, together with out-of-competition tests. After that, doping became much more costly. Morkeberg et al (2009) examine CCS athletes' blood profiles between 2001 and 2007 and report an overall decline in [Hb] compared with the 1990s. However, more recent years saw a change in the method of blood manipulation, i.e., skiers seem to be turning back to transfusions instead of EPO injection.

There is also evidence that the extent of blood manipulation is heavily correlated with performance. Stray-Gundersen et al (2003) analyze blood samples collected at the 2001 World Championships. They show that "of the skiers tested and finishing within the top 50 places in the competitions, 17% had "highly abnormal" hematologic profiles, 19% had "abnormal" values, and 64% were normal. Fifty percent of

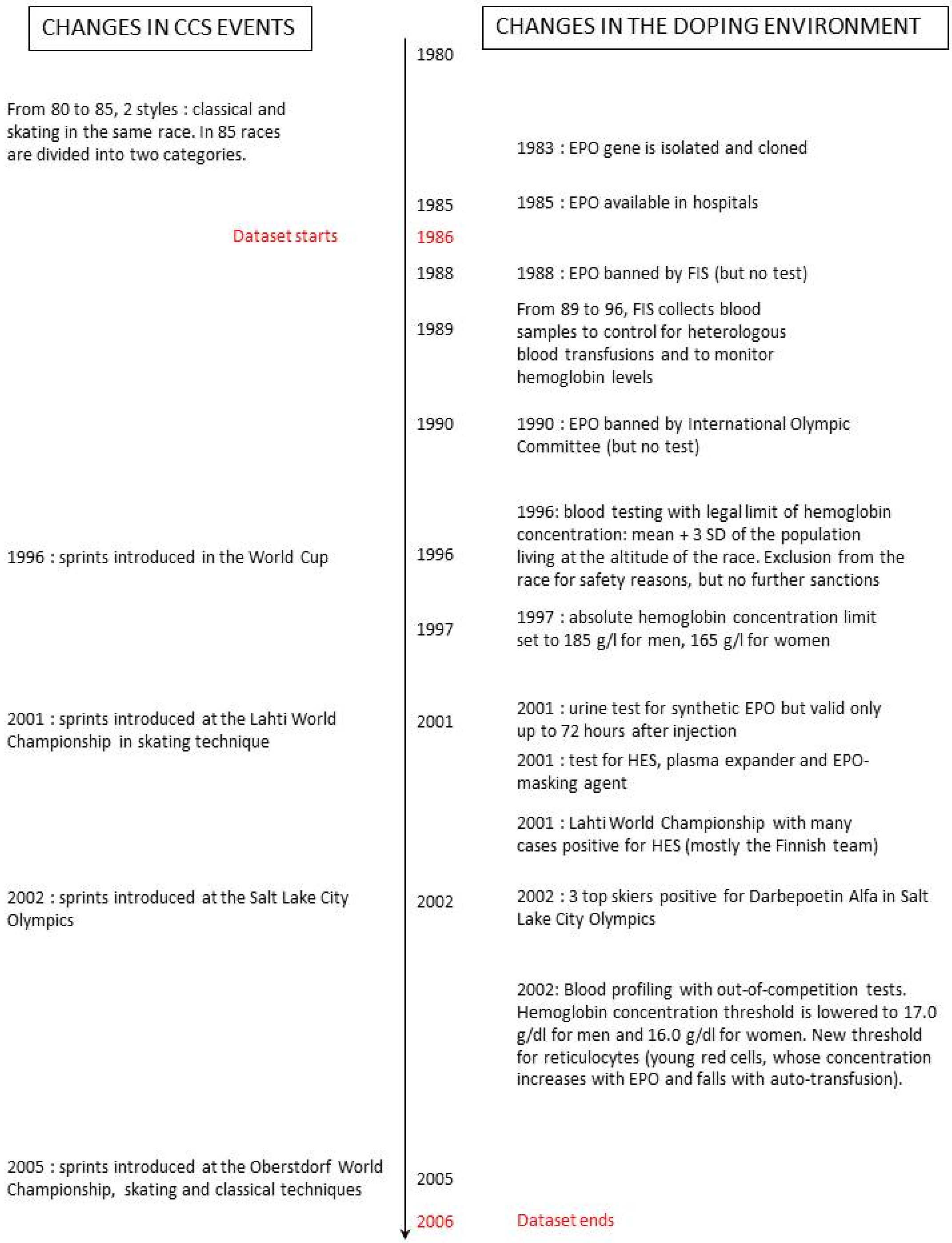

Figure 1: Chronology of changes in CCS events and in the doping environment.

medal winners and 33% of those finishing from 4th to 10th place had highly abnormal hematologic profiles. In contrast, only 3% of skiers finishing from 41st to 50th place had highly abnormal values."

Our interpretation of the facts is that doping costs dramatically decreased in the late 1980s-early 1990s, remaining low up to 1999. From 1999 to 2002, it became more hazardous to use EPO and doping costs increased. The uncertainty about EPO and EPO-masking agent detection led a number of athletes to under-estimate doping costs in the early 2000s, leading to the doping scandals of the 2001 World Championships and 2002 Olympics. After 2002, the doping costs unambiguously increased.

This non-monotonic pattern of doping costs allows us to identify the effects of doping from potential alternative explanations. For instance, improvements in training methods can also contribute to concentrating individual performances around maximum outcomes, thereby reducing competition uncertainty. However, the subsequent increase in doping costs and ranking volatility can hardly be reconciled with the emergence of such a training method.

Methodology.— For each year t, we calculate Spearman's rank correlation ρ_{kt} between the ranking of race k and the final ranking that year. This number is given by

$$\rho_{kt} = 1 - \frac{6\sum_i (fin_{ikt} - rank_{ikt})}{n_{kt}(n_{kt}^2 - 1)}$$

where fin_{ikt} is the hypothetical ranking of player i in race k that would obtain if the yearly hierarchy was fully respected, $rank_{ikt}$ is the actual rank of player i in race k and n_{kt} is the number of participants in that race. Note that all players do not participate in the same set of races, thus fin_{ikt} varies across races.

We then report the yearly average of these Spearman's rank correlations, i.e.

$$\bar{\rho}_t = \frac{\sum_k \rho_{kt}}{R_t}$$

where R_t is the number of races in year t.

The FIS rules underwent some changes over the sample period. Up to 1991, only the top 15 skiers received points in a given race; the dataset reports their rankings, but not the rankings of those finishing after the 15th position. After 1991, the top 30 skiers received points, and the dataset records the rankings of all participants. To harmonize years, we only consider the top 15 skiers in each race. This also allows us to focus on top-level skiers and escape the high volatility in rankings generated by including poorly-ranked athletes.

For each race in the sample we obtain a race-specific correlation, and we average all race-specific correlations for a given year to compute the yearly correlation. Then we match changes in yearly correlation with changes in anti-doping environment.

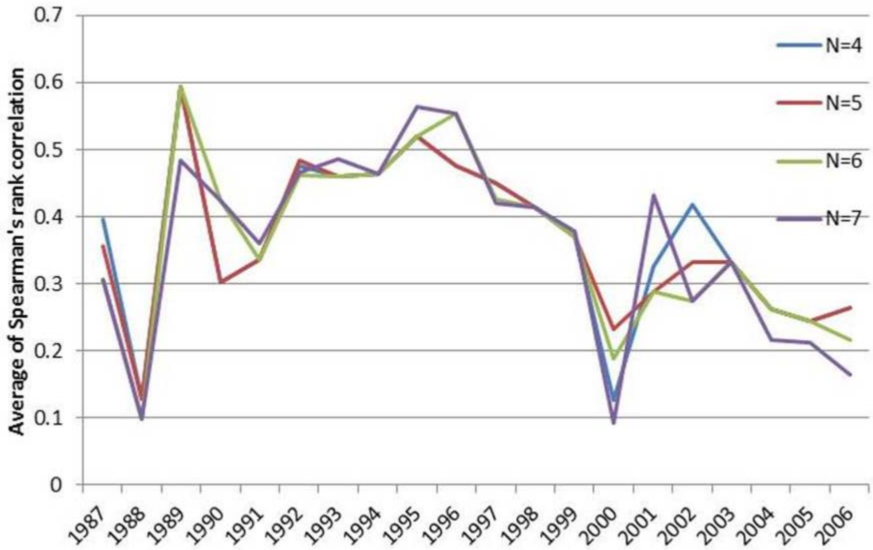

Figure 2: Average of Spearman's rank correlation as a function of the minimum number of top-15 skiers, 1987-2006

To interpret the values of $\bar{\rho}_t$, we compute p-values of the test $\bar{\rho}_t = 0$. While such statistical tests exist for the Spearman correlation, they do not exist for averages of Spearman correlations, so we constructed them by numerical simulations. The null hypothesis is that the ranking of players at each race is independently drawn at random and thus uncorrelated to their final ranking. The simulations account for the number of races that took place and the number of players in each race.

We consider FIS top 15 skiers and consider all races. The number of races varies across years and the number of top 15 skiers varies across races. We only consider races in which at least N (from 4 to 7) of the top 15 skiers participated. A lower N would mean too much variation across years in the number of top skiers considered, while a higher N would severely limit the number of races considered per year.

Results.— Our results are displayed in Figures 2 and 3. The first figure displays the value of the test statistics for every year of the sample. The higher the value, the higher the correlation between race-specific rankings and yearly final rankings. Figure 3 reports the corresponding p-values of the null hypothesis that there is no correlation between the individual race-specific rankings and the overall ranking.

Test statistics and p-values are consistent with the doping pattern previously outlined. The p-values are relatively high in the early stages of the sample, plummet in 1989 and remain very close to 0 during the 1990s, in particular in the mid-1990s. Note also that the change in rules on point attribution does not appear to affect the different curves. The sudden increase in 2000 and subsequent decline in 2001 and 2002 coincide with the uncertainty in the anti-doping environment and the 2001

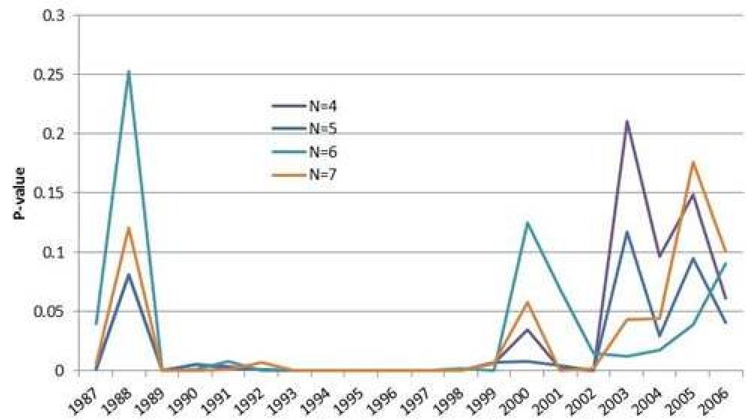

Figure 3: P-values of the zero-correlation test as a function of the minimum number of top-15 skiers, 1987-2006. The null hypothesis of the test is $\bar{\rho}_t = 0$.

World Championship and 2002 Olympics doping scandals. The increase after 2002 coincides with the introduction of blood profiling. The hierarchy among top skiers was remarkably stable during the 1990s, a period of EPO availability. The fact that p-values decrease with the appearance of EPO and increase with the introduction of blood tests suggests that doping tends to reduce competition uncertainty.

In the web appendix A, we discuss a number of variations that we tested in order to eliminate possible biases. We account for race hetereogeneity and exclude sprint races, which may require different skills. We restrict the set of top skiers to the top 10 instead of the top 15. Then, we discuss issues related to changes in the composition of the athletes' pool, in particular the fact that two superstars of the sport retired during the sample period. Lastly, we provide an alternative test by following pairs of athletes and checking whether their relative rankings change over time. In each of these variations we find the same pattern: uncertainty is low during the EPO years and relatively high before and after.

3 A game-theoretic model of cheating

In this section, we describe the main properties of the standard approach and then introduce our model featuring both the standard and regularity effects associated with cheating. We highlight a fundamental result: any equilibrium of this model is also an equilibrium of the reduced model which has only the regularity effect.

3.1 *The standard approach*

The many models on cheating in contests that can be found in the literature all follow the same mechanics. Because our model departs from these and reaches contrasting conclusions, we briefly illustrate these mechanics. In the web appendix B, we give formal statements, derived from several variations of the standard models of cheating.

Assume there are two players, a top dog and an underdog. Both players can decide to cheat by using a technology that increases their best potential performance – the standard effect. The underdog, who knows he is ex-ante less likely to win the contest than his opponent, has incentive to cheat in order to bridge the gap. It is thus on the underdog's side that the cheating process starts. As the top dog does not want to lose because his opponent cheats, he engages in cheating too. In equilibrium, this simple mechanism locks the two players into a cheating strategy; hence a prisoner's dilemma type of mechanism.

There are three direct consequences that follow. First, the top dog is unhappy with this situation. He is forced into a costly cheating strategy only to preserve his chances of winning although he is ex ante the best player. Second, because the cheating incentive is stronger for the underdog, he cheats more than the top dog, and thereby cheating possibilities increase his chances of winning. Therefore, when costs are low, cheating levels the playing field. Third, because players are locked into a bad equilibrium, it only requires some coordination to restore the non-cheating equilibrium, for instance by increasing the cost of cheating.

As we will see, our model predicts the exact opposite: the incentives for cheating start with the best players, who prefer a world in which cheating is feasible. In equilibrium, they cheat more than weaker players, and this increases their chances of winning. Thus, cheating possibilities do not level the playing field, on the contrary they deepen the gap. And last, coordination is not possible because the best players do not want to deter cheating possibilities.

3.2 The cheating game with regularity effect

There are n players who compete in a contest. Each player i is characterized by a pair of possible performance levels $(\underline{a}_i, \overline{a}_i)$, with $\underline{a}_i < \overline{a}_i$. When in a good state, player i achieves the performance level $\overline{a}_i$. When in a bad state he only achieves $\underline{a}_i$. Whether the player is in a good or in a bad state is randomly drawn, with probability p and $(1-p)$ respectively. For simplicity, we set $p = 1/2$. We denote by a_i the performance level realized by player i at the contest, and ranks are determined by the respective rankings of the a_i's.

Players are heterogeneous in their performance levels in the following way: $\overline{a}_i > \overline{a}_j$ and $\underline{a}_i > \underline{a}_j$ if $i < j$. Thus, player 1 is the best player ex ante, while player n is the worst. We assume that $\overline{a}_n > \underline{a}_1$, which guarantees that every player entering the contest has some chance, however small, of winning the contest.

The players compete for k positive prizes, with $k < n$. The prize structure is defined through positive real numbers $y_1 \geq y_2 \geq ... \geq y_k$, and $y_{k+1} = ... = y_n = 0$, where the player that ranks at the i-th position receives prize y_i. If a group of l

players end up in the same position y_i, they all receive the prize $(y_i + \dots y_{i+l-1})/l$.

To stick to the majority of examples that come to mind, we assume that the prize structure satisfies the following weak convexity condition: for any $1 < j < n$,

$$y_j \leq \frac{1}{2}(y_{j-1} + y_{j+1}). \tag{1}$$

Cheating is a continuous variable with values between 0 (no cheating) and 1 (maximal cheating). It comes at marginal cost c. This cost encompasses both the direct cost of using a cheating technology and the indirect costs that are induced, such as penalties in case cheaters are caught, reputation effects, etc. When player i exerts a cheating effort d_i, it has two effects:
(i) *the standard effect* increases the higher performance level $\overline{a}_i$ by the quantity $a(d_i)$, where the function a is non-decreasing. We denote by $\overline{a}_i(d_i)$ the quantity $\overline{a}_i + a(d_i)$;
(ii) *the regularity effect* enhances the probability of being in a good state by the quantity $h(d_i)$: without doping, each player's probability of being in a good state is $1/2$. With doping, player i has a probability of $1/2 + h(d_i)$ of being in a good state. The function h satisfies the following properties:

Hypothesis 1 *h is a continuous function on $[0,1]$, $\mathcal{C}^2$ on $(0,1]$ such that:*

- (i) *h is strictly increasing on $[0,1]$, $h(0) = 0$, $h(1) = 1/2$;*
- (ii) *$h'(1) = 0$ and $\lim_{d \to 0^+} h'(d) = +\infty$;*
- (iii) *h is strictly concave on $(0,1]$ and $h''(1) < 0$.*

Let $d = (d_1, ..., d_n)$ be a cheating profile. Because achieved performance levels are subject to uncertainty, ex-post rankings are random variables. We denote by r_i the random variable giving the ranking of individual i. Then the payoff of the cheating game for player i is given by:

$$U_i(d) = \sum_{l=1}^{n} \mathbb{P}(r_i = l|d) \left(\sum_{j=0}^{n-l} \frac{(y_l + \dots y_{l+j})}{1+j} \right) - cd_i \tag{2}$$

where $\mathbb{P}(.|d)$ is the probability conditional on d.

Notice that the payoff function is discontinuous: assume d_i and d_j are such that $\overline{a}_i(d_i) = \overline{a}_j(d_j)$. Thus individuals i and j will share prizes. Then exerting $d_i + \epsilon$ provides a lump-sump increase of payoff to individual i. Therefore, at any equilibrium, individuals will choose cheating levels such that $\overline{a}_i(d_i) \neq \overline{a}_j(d_j)$ for any pair i and j, and prizes will never be shared. The payoff will be

$$U_i(d) = \sum_{l=1}^{n} \mathbb{P}(r_i = l|d) y_i - cd_i \tag{3}$$

We illustrate with the case of two players, player 1 being the top dog and player 2 the underdog. We also set $y_1 = 1$ and $y_2 = 0$. Then the payoff for player 1 is:

$$U_1(d_1, d_2) = \begin{cases} \frac{1}{2} + h(d_1) + (\frac{1}{2} - h(d_1))(\frac{1}{2} - h(d_2)) - cd_1 & \text{if } \overline{a}_1(d_1) > \overline{a}_2(d_2) \\ \frac{1}{2}(\frac{1}{2} + h(d_1))(\frac{1}{2} + h(d_2)) + (\frac{1}{2} - h(d_2)) - cd_1 & \text{if } \overline{a}_1(d_1) = \overline{a}_2(d_2) \\ \frac{1}{2} - h(d_2) - cd_1 & \text{if } \overline{a}_1(d_1) < \overline{a}_2(d_2) \end{cases}$$

Consider for instance the first equation. If player 1's maximum performance, after cheating, is higher than that of player 2, then player 1 will win whenever he realizes his best performance (this happens with probability $1/2 + h(d_1)$), and will lose when he performs badly (with probability $1/2 - h(d_1)$) and his opponent also performs badly (with probability $1/2 - h(d_2)$). Interpretation is similar for the two other cases.

3.3 Standard effect versus regularity effect

Simplifying notation, in what follows d^* should be understood as $d^*(c)$.

Theorem 1 *Any Nash equilibrium d^* of the cheating game is also an equilibrium of the game with $a \equiv 0$.*

The regularity effect determines the characteristics of equilibrium, regardless of the standard effect. Maximum performance effects only affect existence of equilibrium[4]. Therefore, focusing on the standard effect of cheating leads to partial (no equilibrium) or non-valid conclusions (when there is an equilibrium).

The proof of this theorem relies on the fact that, at any equilibrium d^*, the ranking in terms of maximal potential performance is the same with and without cheating. Thus individuals will only focus on achieving their best performance with enhanced regularity and the standard effect becomes ineffective.

Therefore all of the strategic interactions take place because of the regularity effect.

4 Analysis of the regularity effect

In this section, we analyze the equilibrium of the reduced game with only the regularity effect. We suppose $a \equiv 0$ for all d and examine the effects of $h(.)$ on players' behavior.

Proposition 1 *The cheating game with $a \equiv 0$ is a game of strategic heterogeneity. That is, for any pair i, j such that $\overline{a}_i > \overline{a}_j$, we have*

$$\frac{\partial^2 U_i}{\partial d_i \partial d_j} > 0 \quad \textit{and} \quad \frac{\partial^2 U_j}{\partial d_i \partial d_j} < 0.$$

[4]Indeed, if the players are too close one of the other in terms of maximum performance without cheating, a standard arms race takes place and there will be no pure strategy equilibrium.

This proposition helps us understand the strategic interactions that are at play: for any player, cheating is a strategic complement with the cheating level of opponents that are ex ante less talented, and a strategic substitute with the cheating level of opponents that are ex ante more talented.

The two-player case provides good intuitions. For player 1, strategic interaction arises because he only loses when he is in the bad state and player 2 is in the good state. If player 2 cheats more, his probability of being in the good state rises. This in turn raises player 1's marginal return from cheating. Things are very different for player 2. He loses whenever player 1 is in the good state. Thus cheating efforts are wasted when player 1 makes sufficiently strong efforts.

Strategic complementarity is in line with the usual prisoner's dilemma analysis. Strategic substitutability is one central piece of our model, because it induces the radically different properties that we present below.

Before we focus on the equilibrium properties of this game, notice that a pure strategy equilibrium always exists.

Remark 1 *For any cheating cost $c > 0$, a Nash equilibrium of the cheating game with $a \equiv 0$ exists. It is unique when $n = 2$ but generally not when $n \geq 3$.*

Multiplicity can arise from the model. This is due to the fact that the cheating effort of any player is a strategic complement of the cheating effort of worse players and a strategic substitute of the cheating effort of better players. This cannot arise in the two-player case - the top dog only has a worse player against him, whereas the underdog only competes against a better player.

Proposition 2 (Ordered cheating levels) *For any cheating cost $c > 0$, and any Nash equilibrium d^*, we have that*

$$d_1^* > d_2^* > ... > d_n^* > 0$$

Corollary 1 *For any cheating cost $c > 0$, and any equilibrium d^*, we have*

$$\mathbb{P}(r_i = i | d = d^*) > \mathbb{P}(r_i = i | d = 0)$$

This proposition and its corollary are central to understanding the problem of cheating. Strategic interactions are such that, at any equilibrium, better players cheat more than weaker ones because the returns they can expect from cheating are higher than the returns expected by less talented players. Thus better players cheat more. As a consequence, when cheating is allowed (i.e., $d = d^*$), the probability that a given player ends up ex post in his ex ante ranking (i.e., $r_i = i$) is higher than when there is no cheating (i.e., $d = 0$). This conclusion departs from the standard view on cheating, whereby less talented players cheat in order to bridge the gap with better players and this levels the playing field.

Proposition 3 (Welfare at equilibrium) *Assume there are k positive prizes, while prizes for positions between $k+1$ and n are equal to 0. Then, at any Nash equilibrium d^**

(*i*) $U_j(d^*) < U_j(0)$ *for all* $j \geq k+1$;

(*ii*) *For c small enough,* $U_1(d^*) > U_1(0)$

Cheating carries a negative externality clearly captured by any model on cheating. However, in our model, although the weakest players suffer from the existence of cheating, better players benefit from cheating (at least the best player does; and generally there will also be a group of top players who benefit from cheating) when the cost is not too large (i.e., the authorities are not very repressive or cheap cheating technologies appear). They achieve their best performance more frequently and, therefore, are more likely to win than without cheating.

Proposition 3, added to the observation that $d_1^* > d_i^*$ for all $i \in \{2, ..., n\}$, explains why the best players do not want to combat cheating. They cheat more than the others and are actually happy to compete in an environment where cheating is possible. When the contest is an event relying on popularity, such as sports competitions for instance, then it is difficult to fight against cheating because the best players are those who enhance the popularity of the contest, while at the same time they are the beneficiaries of a cheating system.

Proposition 4 *Assume there are k positive prizes, while prizes for positions between $k+1$ and n are equal to 0. Then,*

(*i*) $\lim_{c \to 0^+} d_i^* = 1$ *for all* $i \in \{1, ..., k\}$;

(*ii*) $\lim_{c \to 0^+} d_j^* = 0$ *for all* $j \in \{k+1, ..., n\}$.

A very small cheating cost implies two effects. First, because cheating becomes more attractive, all players wish to cheat more. Second, the best players become unbeatable because they cheat a lot. By the substitution effect, cheating becomes less attractive for the weaker players. At the limit, the second effect overrides the first one.

This is explained as follows: player 1, by cheating maximally, secures the first position and the highest prize. Consequently, the others behave as if the first player and the first prize did not exist. For them, the game is now a competition with $n-1$ players and $n-1$ prizes. Player 2 is the best player of this competition and secures the highest prize (y_2) by cheating maximally. This goes on as long as there is a positive prize to win. Once all prizes are distributed to the best players, the rest no longer have any incentive to cheat.

This result runs counter the argument whereby free access to doping in sports would lead to a level playing field, because everyone would dope and the results of the competition would not be affected. Our result suggests the reverse, i.e., the best athletes would dope at maximal level, whereas the others would not even try. Competition results become highly predictable and absolute performances are very heterogeneous across players.

Corollary 2 *For c small enough, we have*

$$\frac{dd_i^*}{dc} < 0 \ \ \textit{for all} \ \ i \in \{1, ..., k\} \ \ \textit{and} \ \ \frac{dd_j^*}{dc} > 0 \ \ \textit{for all} \ \ j \in \{k+1, ..., n\}$$

When cheating comes at low cost, an increase in the cost reduces cheating among the k best players and increases it among the $n-k$ other players. The reason is that the increase in cost directly reduces cheating levels of the best players. By the substitution effect, this raises the incentives for lesser players to cheat, as winning one of the k prizes now comes with positive probability.

Corollary 3 *Assume the cost of cheating is low. Then,*

- c *increases* $\Longrightarrow \mathbb{P}(r_i = i|d^*)$ *decreases*
- c *decreases* $\Longrightarrow \mathbb{P}(r_i = i|d^*)$ *increases*

This second corollary shows a key implication of the previous one. When cheating comes at low cost, an increase in the cost translates into less deterministic outcomes of the contest. The k best players cheat less and therefore can be in a bad state and beaten by less talented players who are in a good state. Meanwhile the $n-k$ other players cheat more and reach the top ranks with positive probability. Overall, individual rankings become more volatile.

This corollary provides us with testable predictions and this is an important implication of our paper. Indeed, any empirical work on cheating suffers from the lack of reliable data on cheating behavior. Without such data, any empirical analysis based on the standard model without the regularity effect would have to focus on variations of the players' best performances. However, such performance levels are difficult to observe and their variation may have alternative explanations. Once we account for the regularity effect, we only need to focus on changes in individual rankings, which are easier to observe and interpret.

By this perspective, Corollary 2 offers a rationale to the findings reported in the case study of Section 2. Many factors may explain a reduction in rankings volatility. For instance, more efficient training methods could yield players performing more often at their best levels. So, a decrease in cheating costs may correlate with a decrease in rankings' volatility without explaining it. However, observed increases in rankings' volatility are harder to explain with alternative explanations; regression in training methods seem implausible. But increases in volatility coinciding with increases in cheating cost, as suggested here, seem a plausible explanation.

5 Discussions

This paper studies the incentive to cheat in contest-like situations. It focuses on the regularity effect, a generic term characterizing cheating technologies that tend to concentrate performances around the best individual-specific outcomes. The regularity effect is heterogeneous across agents and benefits the most talented agents. They become more likely to win, and performance inequality rises.

Our arguments apply to any contest with player heterogeneity. A cheating technology that increases the probability of being in a good state (closer to maximal performance) implies the best players have stronger incentive to use the technology, and as a consequence the gap between best players and weaker players widens.

Finance.—Mutual funds compete for capital. Investors allocate resources across funds according to the past performances of each fund. The managers' pay is indexed by their funds' performance, so managers can be seen as athletes competing for the highest return. However, managers differ in skills. The ability to provide lower volatility at given mean return is an attractive feature of a fund. Fraudulent accounting is one cheating technology that allows managers to reduce the volatility of their performance: temporary losses can be hidden, thereby smoothing the return trajectory of the fund. However, it is illegal due to the risk of bankruptcy. Fraudulent accounting benefits the best managers most, as they can count on their superior ability to deliver better returns at some future date so as to restore the genuine profitability of the fund.

Academia.—Scientists compete to publish in the most prestigious journals. Although better researchers have better ideas there is some uncertainty about whether an idea will be confirmed or rejected by proper empirical analysis. In other words, there is uncertainty about whether an idea is in a good or bad state. There is a random component: the history of science is full of good ideas with no empirical support. Here again, there is a cheating method that increases the probability of the idea being in a good state: modifying the data to achieve a better fit with theory. The return on such unethical behavior is actually larger for the best researchers because they have better ideas, and empirical validation of these good ideas makes very strong papers.

Students and exams.—Exam cheating provides another illustration. For mediocre students, exams are not a contest situation because it is only a question of passing or failing. However, for the more talented students, there is more at stake: they want to belong to the top of the grade distribution. For these students, exams are a contest. In this context also, there is some uncertainty about how students will perform on the day of the exam and better students have more incentives to use cheating technologies than average students, because they need to secure their placing. Several cheating technologies exist. Some consist in achieving grades that are above the student's capacity by, for example, copying results from a better student. These technologies are very costly because they are very risky. Other technologies consist in reducing the chance of being in the bad state, for example via cheat sheets or neuroenhancing drugs. Neuroenhancing drugs are increasingly used by students to reduce stress and increase memory and cognitive functions, all of which helps reduce the probability of the bad state when passing the exam (see Maier et al. (2013) and references therein). Their use has become so widespread that debates have arisen as to whether these drugs should be banned, and tested for before exams.

Finally, our paper provides alternative explanations for within-team inequalities. An interesting contribution on this issue is Candelon and Dupuy (2015), who looked at within-team performance inequality in the Tour de France (TdF). Contrary to CCS, cycling is a team sport where tactical skills are important. Moreover, individual ranking in each stage of a multi-stage event is not relevant: what matters is the final time gap between athletes and their direct opponents. The authors measured riders' skills by their speed during the prologue – an individual short-distance time trial. They show that the final average speed gap between the leader and each of his teammates increases over time, controlling for skill differential. They argue that this reflects the increasing concentration of TdF monetary rewards on the winner. Our theory provides a complementary interpretation. During the same years, EPO was very popular. Leaders had stronger incentive to dope than their teammates. They became able to maintain the same performance level throughout the TdF. When the results of the different stages were combined, the average speed gap between leaders and teammates increased.

References

Apesteguia, J., Palacios-Huerta, I., 2010. Psychological Pressure in Competitive Environments: Evidence from a Randomized Natural Experiment. American Economic Review 100, 2548-2564.

Bhaskar, V., 2009. Rational Adversaries? Evidence from Randomized Trials in One Day Cricket. Economic Journal 119, 1-23.

Berentsen, A., 2002. The Economics of Doping. European Journal of Political Economy 18, 109-127.

Berentsen, A., Lengwiler, Y., 2004. Fraudulent Accounting and Other Doping Games. Journal of Institutional and Theoretical Economics 160, 402-415.

Berentsen, A., Bruegger, E., Loertscher, S., 2008. On Cheating, Doping and Whistleblowing. European Journal of Political Economy 24, 415-436.

Bird and Wagner, 1997. Sport as a Common Property Resource: A solution to the dilemmas of doping. Journal of Conflict Resolution 41, 749-766.

Candelon, B., Dupuy, A., 2015. Hierarchical organization and performance inequality: evidence from professional cycling. International Economic Review 56, 1207-1236.

Chiappori, P.-A., Levitt, S., Groseclose, T., 2002. Testing Mixed-strategy Equilibria when Players are Heterogeneous: The Case of Penalty Kicks in Soccer. American Economic Review 92, 1138-1151.

Curry, P.A., Mongrain, S., 2009. Deterrence in Rank-Order Tournaments. Review of Law and Economics 5, 723-740.

Eber, N., Thepot, J., 1999. Doping in Sport and Competition Design. Recherches Economiques de Louvain 65, 435-446.

Glisdorf, K, Sukhatme, V., 2008. Tournament Incentives and Match Outcomes in Women's Professional Tennis. Applied Economics 40, 2405-2412.

Klaassen, F.J.G.M., Magnus, J.R., 2009. The Efficiency of Top Agents: An Analysis through Service Strategy in Tennis. Journal of Econometrics 148, 72-85.

Kräkel, M., 2007. Doping and cheating in contest-like situations. European Journal of Political Economy 23, 988-1006.

Lazear, P.E., Rosen, S., 1981. Rank-Order Tournaments as Optimum Labor Contracts. Journal of Political Economy 89, 841-864.

Maier, L., Liechti, M., Herzig, F., Schaub, M., 2013. To Dope or Not to Dope: Neuroenhancement with Prescription Drugs and Drugs of Abuse among Swiss University Students. PLoS ONE 8(11): e77967. doi:10.1371/journalpone.0077967

Malueg, D.A., Yates, A.J., 2010. Testing Contest Theory: Evidence from Best-of-Three Tennis Matches. Review of Economics and Statistics 92, 689-692.

Monaco, A., Sabarwal, T., 2016. Games With Strategic Complements ans Substitutes. Economic Theory, forthcoming.

Morkeberg, J., Saltin, B., Belhage, B., Damsgaard, R., 2009. Blood profiles in elite cross-country skiers: a 6-year follow-up. Scandinavian Journal of Medicine and Science in Sports 19, 198-205.

Rosen, S., 1986. Prizes and Incentives in Elimination Tournaments. American Economic Review 76, 701-15. Stray-Gundersen, J., Videman, T., Penttilä, I., Lereim, I., 2003. Abnormal hematologic profiles in elite cross-country skiers: blood doping or not? Clinical Journal of Sport Medicine 13, 132-137.

Sunde, U., 2009. Heterogeneity and Performance in Tournaments: a Test for Incentive Effects Using Professional Tennis Data. Applied Economics 41, 3199-3208.

Ulrich, R., Pope, Jr., Harrison, G., Cléret, L., Petróczi, A., Nepusz, T., Schaffer, J., Kanayama, G., Comstock, R.D., Rabin, O., Simon, P., 2016. Doping in Elite Sports Assessed by Randomized-Response Surveys. Mimeo

Videman, T., Lereim, I., Hemmingsson, P., Turner, M.S., Rousseau-Bianchi, M.P., Jenoure, P., Raas, E., Schönhuber, H., Rusko, H., Stray-Gundersen, J., 2000. Changes in hemoglobin values in elite cross-country skiers from 1987-1999. Scandinavian Journal of Medicine and Science in Sports 10, 98-102.

Walker, M., Wooders, J., 2001. Minimax Play at Wimbledon. American Economic Review 91, 1521-1538.

6 Appendix - Results of the general model

In many of the proofs we will use the following formulation of utilities:

If $i < j$, we denote by $N(i,j)$ the random variable equal to the number of players among players $\{i, ..., j\}$ who are in a bad state. Then, conditioning respectively on the events *"player i is in a good state"* and *"player i is in a bad state"*, the payoff of player i can be written as

$$\begin{aligned} U_i(d_1, ..., d_n) &= \left(\frac{1}{2} + h(d_i)\right) \sum_{k=0}^{i-1} y_{i-k} \mathbb{P}[N(1, i-1) = k] \\ &+ \left(\frac{1}{2} - h(d_i)\right) \sum_{l=0}^{n-i} y_{n-l} \mathbb{P}[N(i+1, n) = l] - cd_i \end{aligned} \quad (4)$$

The proof of Theorem 1 will use the following lemma.

Lemma 1 *Any equilibrium d^* of the doping game must be such that $\overline{a}_i > \overline{a}_j \implies \overline{a}_i(d_i^*) > \overline{a}_j(d_j^*)$.*

Proof. Let d^* be a Nash equilibrium of the doping game and call $\overline{a}_i^* := \overline{a}_i(d_i^*)$. We cannot have $\overline{a}_i^* = \overline{a}_j^*$ for some $i \neq j$, as there would then be trivial profitable deviations for the involved agents. Suppose that the statement of the lemma does not hold, let $i := \min\{k = 1, ..., N-1 : \ \overline{a}_k^* < \overline{a}_{k+1}^*\}$ and $j = i+1$. We wish to show that $U_i(d_j^*, d_{-i}^*) - U_i(d^*) > U_j(d^*) - U_j(d_i^*, d_{-j}^*)$, which contradicts that d^* is a Nash equilibrium.

At equilibrium, we necessarily have $\overline{a}_k^* < \overline{a}_{k'}^* \Rightarrow d_k^* < d_{k'}^*$, as a direct consequence of the first-order conditions. Thus,

$$\begin{aligned}
U_j(d^*) &= -cd_j^* + \left(\frac{1}{2} - h(d_j^*)\right) y_{B_1} + \left(\frac{1}{2} + h(d_j^*)\right) y_{G_1}; \\
U_j(d_i^*, d_{-j}^*) &= -cd_i^* + \left(\frac{1}{2} - h(d_i^*)\right) y_{B_1} \\
&+ \left(\frac{1}{2} + h(d_i^*)\right) \left[\left(\frac{1}{2} - h(d_i^*)\right) y_{G_2} + \left(\frac{1}{2} + h(d_i^*)\right) y_{G_2+1})\right]; \\
U_i(d_j^*, d_{-i}^*) &= -cd_j^* + \left(\frac{1}{2} - h(d_j^*)\right) \left[\left(\frac{1}{2} - h(d_j^*)\right) y_{B_1-1} + \left(\frac{1}{2} + h(d_j^*)\right) y_{B_1}\right] \\
&+ \left(\frac{1}{2} + h(d_j^*)\right) y_{G_3}; \\
U_i(d^*) &= -cd_i^* + \left(\frac{1}{2} - h(d_i^*)\right) \left[\left(\frac{1}{2} - h(d_j^*)\right) y_{A_1-1} + \left(\frac{1}{2} + h(d_j^*)\right) y_{A_1}\right] \\
&+ \left(\frac{1}{2} + h(d_i^*)\right) \left[\left(\frac{1}{2} - h(d_j^*)\right) y_{B_2} + \left(\frac{1}{2} + h(d_j^*)\right) y_{B_2+1})\right],
\end{aligned}$$

where B_1, G_1, G_2 and G_3 are random variables that only depend on the state of the agents, other than i and j. More precisely,

- B_1 corresponds to the ranking of agent j when he is in a bad day (notice that since $j = i + 1$ this is equal to the ranking of agent i when he is in a bad day and agent j is in a good day);
- G_1 corresponds to the ranking of agent j when he plays d_j^* and is in a good day (this does not depend on the state of agent i);
- G_2 is the ranking of agent j when he plays d_i^* and is in a good day, whereas agent i is in a bad day; this is equal to the ranking of agent i at equilibrium when he is in a good day and agent j is in a bad day. This is due to the fact that, by definition of i, $\overline{a}_k^* > \overline{a}_i^*$ for $k < i$, and $\overline{a}_j > \overline{a}_k$ for $k > j$.
- G_3 is the ranking of player i when he plays d_j^* and is in a good day.

We illustrate a typical situation with 5 players in Figure 4, where we show the upper performances of the players before and after doping. In this example, players 1 and 2 play the role of players i and j in the above computations.

We have $B_1 > G_2 \geq G_1 \geq G_3$. Denote

$$\Delta := U_i(d_j^*, d_{-i}^*) - U_i(d^*) - U_j(d^*) + U_j(d_i^*, d_{-j}^*).$$

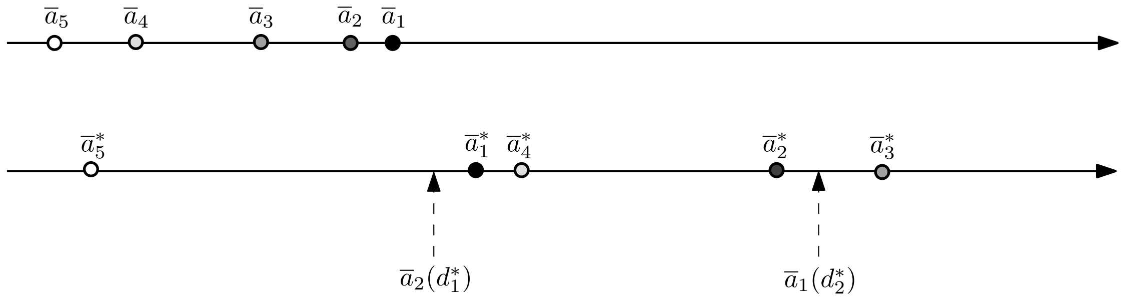

Figure 4: An example with 5 players (only high performances are represented). On the first line, high performances achieved without doping are ranked in the natural order. On the second line, rankings have changed with doping. In particular players 1 and 2's relative positions are reversed. If player 2 chose instead the doping level d_1^* his high performance would lie where the left arrow points to. The second arrow shows where player 1's high performance would lie if he chose d_2^*.

Then

$$\begin{aligned}\Delta &= \left(h(d_j^*) - h(d_i^*)\right)\left(\frac{1}{2} - h(d_j^*)\right)(y_{B_1} - y_{B_1-1}) + \left(\frac{1}{2} + h(d_j^*)\right)(y_{G_3} - y_{G_1}) \\ &+ \left(\frac{1}{2} + h(d_i^*)\right)(h_(d_j^*) - h(d_i^*))(y_{G_2-1} - y_{G_2}) \\ &\geq \left(h(d_j^*) - h(d_i^*)\right)\left[\left(\frac{1}{2} - h(d_j^*)\right)(y_{B_1} - y_{B_1-1}) + \left(\frac{1}{2} + h(d_i^*)\right)(y_{G_2-1} - y_{G_2})\right]\end{aligned}$$

Since $y_{G_2-1} - y_{G_2} > y_{B_1} - y_{B_1-1}$ by convexity assumption on prizes, we have $\Delta > 0$ ■

Proof of Theorem 1. Let d^* be a Nash equilibrium of the doping game. We prove that d^* is a Nash equilibrium with $a \equiv 0$ by induction. In other terms, we must prove that $d_i^* = Br_U^i(d_{-i}^*)$ for all i. Let V_i denote the payoff functions in the doping game and U_i be the payoff functions when $a \equiv 0$. By Lemma 1, $U_i(d_i^*, d_{-i}^*) = V_i(d_i^*, d_{-i}^*)$ for all i. Agent N always obtains a better payoff in the doping game: $V_N(d) \geq U_N(d)$ for all d. Hence, given $d_N \geq 0$ we have

$$U_N(d_N, d_{-N}^*) \leq V_N(d_N, d_N^*) \leq V_N(d_N^*, d_{-N}^*) = U_N(d_N^*, d_{-N}^*),$$

which proves that $d_N^* \in Br_U^N(d_{-N}^*)$. Now let $i \in \{1, ..., N-1\}$ and assume that for any $j > i$ we have $d_j^* = Br_U^j(d_{-j}^*)$. We need to prove that $d_i^* = Br_U^i(d_{-i}^*)$. Let $\tilde{d}_i = Br_U^i(d_{-i}^*)$. We claim that $\overline{a}_i(\tilde{d}_i) \geq \overline{a}_{i+1}(d_{i+1}^*)$. Assume by contradiction that this is not the case. Then in particular $\tilde{d}_i < d_i^*$ and we have

$$d_{i+1}^* = Br_U^{i+1}(d_{-(i+1)}^*) \leq Br_U^{i+1}(d_{-i,-(i+1)}^*, \tilde{d}_i) < Br_U^i(d_{-i}^*) = \tilde{d}_i,$$

where the first inequality follows from strategic substitution and the second inequality from the fact that $Br^i_U(d_{-i}) > Br^{i+1}_U(d_{-(i+1)})$ for all d, and taking $d = (\tilde{d}_i, d^*_{-i})$. However $\tilde{d}_i > d^*_{i+1}$ contradicts $\overline{a}_i(\tilde{d}_i) < \overline{a}_{i+1}(d^*_{i+1})$. Thus $V_i(\tilde{d}_i, d^*_{-i}) \geq U_i(\tilde{d}_i, d^*_{-i})$, as agent i's payoff in the doping game is at least as much as in the game with $a \equiv 0$, and possibly larger when $\tilde{d}_i$ is such that $\overline{a}_i(\tilde{d}_i) > \overline{a}_{i-1}(d^*_{i-1})$. Consequently

$$V_i(\tilde{d}_i, d^*_{-i}) \geq U_i(\tilde{d}_i, d^*_{-i}) \geq U_i(d^*_i, d^*_{-i}) = V_i(d^*_i, d^*_{-i}) \geq V_i(\tilde{d}_i, d^*_{-i})$$

and $\tilde{d}_i = d^*_i$.■

Proof of Proposition 1. By equation (4) we have

$$\frac{\partial U_i}{\partial d_i} = h'(d_i)\left(\underbrace{\sum_{k=0}^{i-1} y_{i-k}\mathbb{P}[N(1,i-1)=k]}_{\text{Better ranked players}} - \underbrace{\sum_{l=0}^{n-i} y_{n-l}\mathbb{P}[N(i+1,n)=l]}_{\text{Lesser ranked players}}\right) - c, \quad (5)$$

Clearly, when player j increases his cheating level, either $\overline{a}_j > \overline{a}_i$ and only the first term in the brackets is affected, or $\overline{a}_j < \overline{a}_i$ and only the second term in the brackets is affected. In both cases, the corresponding probability decreases in d_j (the probability that j is in a bad state decreases with d_j). This proves the proposition.■

Proof of Remark 1.

• Existence: The first order conditions are, for player i,

$$h'(d_i)\left(\sum_{k=0}^{i-1} y_{i-k}\mathbb{P}[N(1,i-1)=k] - \sum_{l=0}^{n-i} y_{n-l}\mathbb{P}[N(i+1,n)=l]\right) = c,$$

which we denote by $\alpha_i(d_{-i})h'(d_i) = c$. In the sequel, we will omit the dependence of α_i with respect to d when there is no ambiguity.

The map h' is strictly decreasing from $(0,1]$ to $[0,+\infty)$. Hence the inverse function $(h')^{-1}$ is well defined and strictly decreasing from $[0,+\infty)$ to $(0,1]$.

Thus best-response are given by

$$Br^i(d_{-i}) = (h')^{-1}\left(\frac{c}{\alpha_i(d_{-i})}\right)$$

Existence follows from Brouwer Theorem and the fact that $Br := (Br^i)_{i=1,\dots,n}$ is continuous and maps $[0,1]^n$ to itself.

• Uniqueness in the case of two players: we have

$$\begin{aligned} U_1(d_1,d_2) &= -cd_1 + \left(\frac{1}{2} + h(d_1)\right) + \left(\frac{1}{2} - h(d_1)\right)\left(\frac{1}{2} - h(d_2)\right), \\ U_2(d_1,d_2) &= -cd_2 + \left(\frac{1}{2} - h(d_1)\right)\left(\frac{1}{2} + h(d_2)\right), \end{aligned}$$

and

$$\begin{aligned} Br^1(d_2) &= (h')^{-1}\left(\frac{c}{1/2+h(d_2)}\right), \\ Br^2(d_1) &= (h')^{-1}\left(\frac{c}{1/2-h(d_1)}\right). \end{aligned}$$

Br^1 is increasing in d_2 and Br^2 is decreasing in d_1, since $(h')^{-1}$ is decreasing. Hence uniqueness of the equilibrium follows.

• Non-uniqueness with more than two players: Let $n = 3$ and $y := y_2 < 1/2$. We show the following: given $c > 0$ and $y \in]0, 1/2[$ there exists a function h satisfying assumption 1 and $0 < d_3 < \overline{d}_3 < \overline{d}_2 < d_2 < d_1 < \overline{d}_1$ such that

$$\alpha_i(d)h'(d_i) = \alpha_i(\overline{d})h'(\overline{d}_i) = c \; \forall i = 1, ..., 3.$$

This implies that d and $\overline{d}$ are two distinct Nash equilibria.

Let $\epsilon > 0$. Pick h_1, h_2, h_3 such that $1/4+\epsilon > h_1 > h_2 > h_3 > 1/4-\epsilon$ and define $\overline{h}_3 = h_3+\epsilon$, $\overline{h}_2 = h_2-\epsilon^2$, $\overline{h}_1 = h_1+\epsilon^3$ It is straightforward to check that for ϵ small enough we have

a) $1 > \overline{h}_1 > h_1 > h_2 > \overline{h}_2 > \overline{h}_3 > h_3 > 0$,

b) $(\overline{h}_2+\overline{h}_3)/2-(1-2y)\overline{h}_2\overline{h}_3 > (h_2+h_3)/2-(1-2y)h_2h_3$

c) $(\overline{h}_1+\overline{h}_2)/2-(1-2y)\overline{h}_1\overline{h}_2 < (h_1+h_2)/2-(1-2y)h_1h_2$

Define

$$\begin{aligned} \alpha_1 &= 3/4+y/2+(h_2+h_3)/2-(1-2y)h_2h_3, \\ \alpha_2 &= 1/2+y/2-(1-y)h_1+yh_3, \; \alpha_3 = 1/4+y/2-(h_1+h_2)/2+(1-2y)h_1h_2 \end{aligned}$$

and the $\overline{\alpha_i}$ analogously.

Clearly $0 < \alpha_3 < \overline{\alpha}_3 < \overline{\alpha}_2 < \alpha_2 < \alpha_1 < \overline{\alpha}_1 < 1$. Let $\overline{d}_1 > d_1 > d_2 > \overline{d}_2 > \overline{d}_3 > d_3 > 0$ be such that

$$\alpha_3\frac{(\overline{h}_3-h_3)}{c} < \overline{d}_3-d_3 < \overline{\alpha}_3\frac{(\overline{h}_3-h_3)}{c}$$

$$\overline{\alpha}_3\frac{(\overline{h}_2-\overline{h}_3)}{c} < \overline{d}_2-\overline{d}_3 < \overline{\alpha}_2\frac{(\overline{h}_2-\overline{h}_3)}{c}$$

$$\overline{\alpha}_2\frac{(h_2-\overline{h}_2)}{c} < d_2-\overline{d}_2 < \alpha_2\frac{(h_2-\overline{h}_2)}{c}$$

$$\alpha_2 \frac{(h_1 - h_2)}{c} < d_1 - d_2 < \alpha_1 \frac{(h_1 - h_2)}{c}$$
$$\alpha_1 \frac{(\overline{h}_1 - h_1)}{c} < \overline{d}_1 - d_1 < \overline{\alpha}_1 \frac{(\overline{h}_1 - h_1)}{c}$$

Satisfying this set of constraints is always possible. Moreover, if ϵ is picked small enough at the start of the proof, we can always have $\overline{d}_1 < 1$. We now construct our function h. First let h be such that, for $i = 1, ..., 3$,

$$h(d_i) = h_i,\ h'(d_i) = c/\alpha_i;\ \ h(\overline{d}_i) = \overline{h}_i,\ h'(\overline{d}_i) = c/\overline{\alpha_i},$$

The set of inequalities above implies that h can be extended as a twice differentiable function, strictly increasing and strictly concave, which means that it satisfies assumption 1.■

Proof of Proposition 2. We show the following: Let $n > i \geq 1$ and $d = (d_1, ..., d_n) \in [0,1]^n$. Then we necessarily have $\alpha_i(d) > \alpha_{i+1}(d)$. As an immediate consequence, if d^* is a Nash equilibrium then we have $d_1^* > d_2^* > ... > d_n^*$. Note that, conditioning on the events *"player i is in a good/bad state"*, we have

$$\begin{aligned}\sum_{k=0}^{i} y_{i-k+1}\mathbb{P}[N(1,i) = k] \quad &= \quad \left(\frac{1}{2} + h(d_i)\right)\sum_{k=0}^{i-1} y_{i-k+1}\mathbb{P}[N(1,i-1) = k] \\ &+ \quad \left(\frac{1}{2} - h(d_i)\right)\sum_{k=0}^{i-1} y_{i-k}\mathbb{P}[N(1,i-1) = k]\end{aligned}$$

As a consequence

$$\begin{aligned}&\sum_{k=0}^{i-1} y_{i-k}\mathbb{P}[N(1,i-1) = k] - \sum_{k=0}^{i} y_{i-k+1}\mathbb{P}[N(1,i) = k] \\ = \quad &\left(\frac{1}{2} + h(d_i)\right)\sum_{k=0}^{i-1}(y_{i-k} - y_{i-k+1})\mathbb{P}[N(1,i-1) = k]\end{aligned}$$

Analogously, conditioning on the events *"player i+1 is in a good/bad state"*, we have

$$\begin{aligned}\sum_{l=0}^{n-i} y_{n-l}\mathbb{P}[N(i+1,n) = l] \quad &= \quad \left(\frac{1}{2} + h(d_{i+1})\right)\sum_{l=0}^{n-i-1} y_{n-l}\mathbb{P}[N(i+2,n) = l] \\ &+ \quad \left(\frac{1}{2} - h(d_{i+1})\right)\sum_{l=0}^{n-i-1} y_{n-l-1}\mathbb{P}[N(i+2,n) = l],\end{aligned}$$

which implies

$$\begin{aligned}
&\sum_{l=0}^{n-i} y_{n-l}\mathbb{P}[N(i+1,n)=l] - \sum_{l=0}^{n-i-1} y_{n-l}\mathbb{P}[N(i+2,n)=l] \\
= &\left(\frac{1}{2}-h(d_{i+1})\right)\sum_{l=0}^{n-i+1}(y_{n-l-1}-y_{n-l})\mathbb{P}[N(i+2,n)=l].
\end{aligned}$$

Finally, we obtain

$$\alpha_i - \alpha_{i+1} = \left(\frac{1}{2}+h(d_i)\right)\sum_{k=0}^{i-1}(y_{i-k}-y_{i-k+1})\mathbb{P}[N(1,i-1)=k] \quad (6)$$

$$- \left(\frac{1}{2}-h(d_{i+1})\right)\sum_{l=0}^{n-i-1}(y_{n-l-1}-y_{n-l})\mathbb{P}[N(i+2,n)=l]. \quad (7)$$

Notice that $\sum_{k=0}^{i-1}\mathbb{P}[N(1,i-1)=k] = \sum_{l=0}^{n-i+1}\mathbb{P}[N(i+2,n)=l] = 1$. Moreover, by convexity of the family $(y_k)_k$ we have

$$y_{i-k} - y_{i-k+1} \geq y_{n-l-1} - y_{n-1} \ \forall\, 0 \leq k \leq i-1, \forall\, 0 \leq l \leq n-i-1,$$

with equality only for $l = n-i-1$ and $k = 0$. The sum in (6) is a convex combination of terms of the form $y_{i-k} - y_{i-k+1}$ multiplied by a real number greater than $1/2$, whereas the sum in (7) is a convex combination of terms of the form $y_{n-l-1} - y_{n-l}$ multiplied by a real number smaller than $1/2$. Hence we obtain that $\alpha_i - \alpha_{i+1} > 0$.■

Proof of Proposition 3. When c is close to 0, d_1^* is close to 1 (see the proof of Proposition 4 below), thus the payoff of player 1 is close to y_1. When cheating does not exist, his payoff is a convex combination of all prizes y_1 to y_n, which is trivially smaller than y_1. Conversely for player n.■

Proof of Proposition 4. We show the following: given $\epsilon > 0$, there exists c_0 such that, for any $c < c_0$ and any equilibrium d^*, we have $d_i^* > 1-\epsilon$ for all $i = 1, ..., n-k$ and $d_j^* < \epsilon$ for all $j = n-k+1, ..., n$.

Assume first that only $y_n = 0$. For $i < n$ it is sufficient to prove that α_i is bounded from below by a positive constant that does not depend on d. For $i < n$, we have $\mathbb{P}[N(i+1,n) = n-i] < 1/2$. Hence

$$\alpha_i > y_i - \left(\frac{1}{2}y_{i+1} + \frac{1}{2}y_i\right) = \frac{1}{2}(y_i - y_{i+1}) > 0$$

Now we turn to $i = n$. We have

$$\begin{aligned}
\alpha_n &= \sum_{k=0}^{n-1} y_{n-k}\mathbb{P}(N(1,n-1)=k) \\
&\leq \sum_{k=1}^{n-1} \mathbb{P}(N(1,n-1)=k) = 1 - \mathbb{P}(N(1,n-1)=0) \\
&\leq 1 - \prod_{i=1}^{n-1}(1-(1/2-h(d_i))) \\
&\leq \sum_{i=1}^{n-1}(1/2-h(d_i)) \\
&\leq \sum_{i=1}^{n-1} h'(d_i)(1-d_i) \\
&\leq \sum_{i=1}^{n-1} \frac{c(1-d_i)}{\alpha_i}
\end{aligned}$$

We proved that, for any Nash equilibrium $(d_1^*, d_2^*, ..., d_n^*)$ we have

$$h'(d_i) < \frac{2c}{y_i - y_{i+1}}, \ \forall i = 1, ..., n-1; \ \ h'(d_n) \geq \frac{\alpha_{n-1}}{\sum_{i=1}^{n-1} 1 - d_i^*}$$

This concludes the proof when $y_{n-1} > y_n$.

Now, assume that $y_{n-1} = y_n = 0$. For player $n-1$, it is as if player n was not there: if player $n-1$ is in a good state, he will beat player n regardless of his performance; if player $n-1$ is in a bad state, he will get 0 whatever happens. As a result, player $n-1$ acts as if he was competing in a race with $n-2$ opponents. Because he is the worst player, $d_{n-1}^* \to 0$ as the cost goes to zero. This reasoning trivially generalizes.$\blacksquare$