THE UNIVERSITY of ADELAIDE

School of Economics

Working Papers

ISSN 2203-6024

A Comparison of Anti-Doping Measures in Sporting Contests

Qin Wu
Ralph C-Bayer
Liam Lenten

Working Paper No. 2016-11
May 2016

A Comparison of Anti-Doping Measures in Sporting Contests

Qin Wu*, Ralph-C Bayer, Liam Lenten

May 17, 2016

Abstract

This paper proposes a new anti-doping policy. In a conditional superannuation scheme athletes have to pay a certain fraction of their season income from sports into a superannuation fund from which they can only draw if they have never been caught doping. Theoretically, this fund has two important advantages over conventional anti-doping policies such as bans and fines. It does not lose its deterrence effect when athletes get near the end of their careers such as in the case of bans and it can deal with the widespread problem that drug cheats are often only found out much later when the detection technology has caught up with doping practices. We build a model of a dynamic sporting contest implement it in the laboratory and compare the performance of our policy to that of traditional policies. Our policy compares favorable with respect to doping prevention without reducing the quality of competition more than other measures do.

*Bayer: School of Economics, The University of Adelaide, Adelaide, SA 5005, Australia (Email: ralph.bayer@adelaide.edu.au). Lenten: Department of Economics and Finance, La Trobe University, Australia (Email: l.lenten@latrobe.edu.au). Wu: School of Economics, The University of Adelaide, Adelaide, SA 5005, Australia (Email address:.qin.wu@adelaide.edu.au). Authors are alphabetically ordered by last names.

1 Introduction

The use of performance-enhancing drugs (PEDs) is rife in many competitive sports. Over the past five years, cycling, weightlifting and boxing are the three most common Olympic sports in which common PED-use by athletes exists. Especially in the cycling world, doping which was allowed at the early stage of Tour de France has been growing rapidly even after its ban since mid-1960s. With the revelations around the the drug use of the legendary cyclist Lance Armstrong, public trust in conventional prevention measures is extremely low. Currently, if a drug cheat is caught, he might lose a considerable amount of money (in loss of sponsorship, prize money, etc.) and will typically be banned from competing for a certain period of time. The current anti-doping regime has two considerable weaknesses. Since doping practices are always slightly ahead of the test technology, many cheats are only caught years after the actual doping happened. This severely reduces the deterrence effect of loss of prize money and sponsorship contracts and of bans, since the athletes can count on having finished their career by the time they are caught. Secondly, the current system of bans fails to provide strong anti-doping incentives to athletes that are close to the end of their careers.

In the light of the disadvantages of the current system, we are proposing an alternative or additional measure: conditional superannuation. Athletes have to pay part of their winnings and sponsorship moneys into a superannuation fund and will only receive payments out of the fund if they have never been found guilty of doping during their career. This measure has the potential to overcome the two disadvantages of the current system. The athlete's balance will increase during the career, which ensures that the loss from being found guilty also increases over the career. Hence, the deterrence effect is maximized at the end of the career, where bans lose their effectiveness. Moreover, it is possible to set the date when a decision on payouts is made such that enough time has elapsed and testing old samples with modern techniques has made sure that there was no cheating that has not been detected at the time.

While, in theory, advantages of conditional superannuation appear to exist, it is unclear how athletes actually react to such a measure. It considerably changes doping incentives but also incentives to exert training effort over an athletes career. The changing incentives together with general-equilibrium effects stemming from athletes competing with others, makes it hard to evaluate the likely performance of a conditional superannuation system on the base of theory alone. For this reason we implement different measures in a real-effort contest environment that models salient characteristics of sports contests and compare their performance. Three experimental subjects compete in three consecutive share contests, where the share of a prize an athlete receives is determined by a linear contest-success function with performances being the input. Subject's performances are determined by three components: the training effort, which is measured as the real effort in an adding-up task, an endogenously given time-varying ability parameter and by the doping decision. We compare the behavior under three different anti-doping regimes, a fine regime, where a doper loses his prize for the period he was caught, a ban regime where a caught doper is excluded from the next contest and the superannuation regime, where an athlete will be paid a withheld fraction of their prize moneys only if they were never caught doping. In order to make the three treatments comparable we calibrate the models by choosing enforcement parameters such that expected loss from doping is equalized across treatments. As a benchmark we add a control treatment without anti-doping policy in place.

We find that our conditional superannuation policy does clearly better at deterring athletes from doping than bans. This is mainly due to the problem that bans lose their deterrence effect once an athlete has decided to retire. The superannuation policy is also slightly better at deterring doping than fines. This is in addition to the advantage that the superannuation policy can deal with delayed detection, which is commonly observed. While more effective than other policies in reducing doping, our superannuation policy is shown not to reduce efforts and therefore the quality of competition to a larger extent than other policies.

2 Related literature

Tournaments arew used as an incentive device to allocate resources efficiently. Ehrenberg and Bognanno (1988) show the incentive effect of tournaments by studying a set of golf tournaments data. They find that by adjusting the reward structure and the prizes in tournaments, players' performance can be influenced directly. In tournaments, it is the participant's relative performance that determines their payoff. Compared to other common incentive schemes, such a noncompetitive piece rates, tournaments can often increase average effort levels, with a positive effect on output levels (Lazear and Rosen, 1979). However, unlike any other noncompetitive payment schemes, competition discourages corporation. Not only that, it can cause destructive activities such as sabotage (Lazear, 1989). By sabotaging, a player might damage other competitors' performance to increase his or her chance of winning. Lazear (1989) addresses the problem of sabotage in a theoretical framework and shows that the smaller the prize gap between winners and losers, the lower the incentive to sabotage. Sabotage is not the only drawback in tournaments. The problem of cheating (i.e. doping in sports) is another severe issue. According to Gneezy (2005), there are four categories of lies: lies that do not negatively affect anyone including the listners of lies, lies that benefit the listners even at the cost of liars, lies that harm both listners and liars and lies that benefit only liars but not the other side. He conducts a cheap talk sender-receiver game where player 1 can send player 2 either a true message or a fake message which benefits the sender at the expense of player 2. Through changing the gain of lying for player 1 and the cost imposed on player 2, he finds that people tend to cheat more with a higher gain. However, people also care about others' losses and choose to be honest if the losses are too big for others. In Fischbacher and Föllmi-Heusi (2013)s' paper, a simple dice rolling game is employd where players can self-report the number they have rolled. The result shows the nearly half of the population lies but that most of the liars did not lie to the extreme. Lying is usually harmful to the society and it is contagious based on what Robert and Arnab (2013) find in their paper. Real-world observations about doping suggest that cheating in contests follows the same pattern. In real life, people are concerned about whether the situation is fair (Ellingsen and Johannesson, 2004). Therefore, unfairness as a consequence of cheating behavior is damaging to the well-being of those who feel unfairly treated.

While there is not much experimental work that looks at doping prevention the fight against doping has been looked at theoretically. Berentsen et al. (2008) introduce a whistle-blowing mechanism into a two players game in which the loser can pay to report the winner for doping after the competition. Only the reported winner needs to take drug tests. The paper conducts a comparison of the whistle-blowing game with the normal inspection game where a third party has the final say on whether to test the winner or not. This whistle-blowing mechanism is more effectice in lowering the probabilities of doping, and whistle-blowing can provide a Pareto-optimal equilibrium. Furthermore, this mechanism is less costly as it requires less tests relative to the normal inspection game. Gilpatric (2011) models doping as a continuous instead of a dichotomous variable. He studies two aspects of enforcement and finds that, firstly, correlated audits are more effective in reducing doping compared to independent audits. Secondly, an anti-doping policy that gives losers the prize money by default if winners get caught can reduce doping incenitves. The later result is in line with what Curry and Mongrain (2009) found in their paper, which focuses on the deterrence problem. In addition, in Gilpatric (2011) the number of players plays an important role. The author suggests that cheating is more likely to happen with more participants in the contest. In one of the most recent papers which explores doping behavior, Ryvkin (2013) models a symmetric winner-take-all tournament game with an uncertain number of participants. On top of the general results that are along the lines of previous findings, his result of a non-monotonic penalty-testing frontier is of special interest. In a nutshell, the non-monotonic penalty-testing frontier describes the relation of the social optimal equilibrium in which doping does not exist and the number of participants determined endogenously in the game. The minimal size of the penalty to stop doping is a non-monotonic function of the number of contestants. Individual's chance of winning the prize becomes smaller when the number of potential contestants is large, and therefore doping is less attractive to players in a big game. However, if the number of participants is moderate, the marginal benefit of doping outweighs the cost of sharing the prize money with additional players. Hence, the incentive of doping grows in this case. In the same manner, the probability of getting caught that ensures a clean game in equilibrium as a function of number of participants is again non-monotonic.

Most of the existing empirical literature tests theoretical doping models in a controlled laboratory environment. (Cason et al., 2010; Curry and Mongrain, 2009; Faravelli et al., 2014; Fischbacher and Föllmi-Heusi, 2013; Friesen and Gangadharan, 2012; Niederle and Vesterlund, 2005) Doping is forbidden in real-word competitions, and therefore, field data on doping can barely be observed and collected. According to Van Dijk et al. (2001), abstract experiments do not require subjects to gain effort levels. Instead, they simply choose numbers as effort levels. However, in real life, "work involves effort, fatigue, boredom, excitement and other affectations not present in the abstract experiments." Two of the most famous real effort tasks are the Slider Task where subjects need to complete tasks by moving the slider bar to the middle (Gill and Prowse, 2012) and the Matrix Task at which subjects pair numbers together in given matrices to get a sum of 10 (Mazar et al., 2008). Both tasks require subjects to concentrate and do the same thing repetitively just like in a training process. However, neither of them capture the fact that as the training program continues, it progressively becomes more difficult to increase the performance as a consequence of increased training. To capture this feature we employ a task (summing up one-digit numbers), which becomes increasingly difficult, as the number of one-digit numbers that have to be summed increases over time.

The rest of the paper is structured as follows. Section 2 describes the theoretical framework of a heterogenous N-player game under four types of punishment mechanisms. Section 3 illustrates the design of the experiment based on our theoretical model. Section 4 reports the results from the analysis of the data. The paper concludes in Section 5.

3 The Underlying Model

Our model consists of four individual treatments: (i) a penalty-free system, (ii) a fine system, (iii) a ban system and (iv) a superannuation fund system. In this section, we first discuss the penalty-free model in detail and then describe the three anti-doping regimes. Generally, our environment makes it impossible to solve out for a Subgame-Perfect Nash Equilibrium for two reasons. Firstly, the subjects ability in the adding-up task is ex-ante unknown, which means that we do not know the cost of all individuals exerting effort. Secondly, in the more complex enforcement treatments, things become very complex and intractable. For this reason we derive a variety of predictions that are independent of equilibrium play. We start with the penalty-free environment.

3.1 Heterogeneous N-Player Dynamic Game without Enforcement

The baseline model is a penalty-free model in which n risk-neutral players participate in T consecutive season, which consist of one sports contest each. All players are indexed by i where $i \in I = \{1, 2, 3, \cdots, n\}$. It is assumed that athletes maximize their career expected payoffs. We assume that an athlete's performance q_i in a season depends on on his/her training effort e_i, natural ability r_i, and doping decision d_i. Intuitively, a higher training level and/or ability implies a better performance. In addition, performance can be enhanced by using performance-enhancing drugs (PEDs) regardless of the potential health damages caused by such behavior in the long run. We assume these three factors to complements, which implies that an increase of any of the factors increases the marginal impact of any of the the other factors. The performance is simply assumed to be the product of the factors:

$$q_i = e_i r_i \delta(d_i) \tag{1}$$

$$\delta(d_i) = \begin{cases} 1 & d_i = 0 \\ \delta & d_i = 1 \end{cases}$$

The ability level lies within an interval, $r_i \in [\underline{r}, \overline{r}]$ where $\underline{r}$ is the lowest ability level and $\overline{r}$ is the highest ability level among all players. A player who does not dope will have a doping multiplier of unity, while a doper receives a multiplicative performance boost of $\delta > 1$.

It is important to note that the doping decision d_i is a dichotomous choice variable with values of 0 for non-dopers and 1 for dopers. In what follows we refer to the two possible actions as "C" for clean and "D" for dope. Note that effort and doping e_i and d_i are the only two choice variables in this model.

In each season subjects compete in a share-prize contest. In many sports individual contests use winner-take-all-like compensation schemes (e.g., virtually all elite tennis tournaments). However, instead of modeling a specific sports contest, we are interested in a certain period of the athlete's career or a whole season they played, which contains many competitions. We are also not only interested in the prize money athletes can earn but also in other kinds of performance-based income during that period, such as annual salary from clubs and sponsorship and advertising income. Our share-contest setting was chosen, since there the payoff in a season depends on the relative performance of an athlete, which is very realistic. For simplicity, we call this season payoff despite its broader meaning just *prize money.* Once all individual performances are measured, an athlete's relative performances can be observed. We use the following simple linear contest function:

$$w_i = \beta(\triangle q_i) + \frac{1}{n} \tag{2}$$

where $w_i \in [0, 1]$ is the prize share received by i. The difference between player i's performance, q_i, and the average performance of all other $n-1$ athletes, $\overline{q}_{-i}$, is represented by $\triangle q_i$, where $\overline{q}_{-i} = \sum q_{-i}/(n-1)$. Parameter β denotes the reactivity to performance differences, and the last constant term, $1/n$ ensures that every player gets the same share if $\triangle q_i = 0 \forall i$. [1] The price shares of all competitors sum to one.

To derive the expected payoff $E(\pi_i)$ for player i, we have to specify training cost and total prize money. We assume that $C(e_i)$ is a subject's training cost which satisfies $C^{'}(e_i) > 0$ and

[1] Note that keeping shares between 0 and 1 requires a restriction on β.

$C(0) = 0$. The total prize money is denoted by V. The expected payoff in a season, as a result, is given by

$$E(\pi_i) = w_i V - C(e_i). \tag{3}$$

Substituting the prize share function (2) into the payoff funtion (3), the expected payoff for player i can be written as

$$E(\pi_i) = \left[\beta\left(e_i r_i \delta(d_i) - \frac{\sum e_{-i} r_{-i} \delta(d_{-i})}{n-1}\right) + \frac{1}{n}\right] V - C(e_i)$$

Given the payoff functions, player i is assumed to maximize his/her payoff by choosing a training level e_i and by making a doping decision, d_i. The first-order condition for the optimal training level, e_i requires

$$C^{'}\left(e_i^*(d_i)\right) = \beta V r_i \delta(d_i). \tag{4}$$

Hence in our simple linear setting, player i's effort decision does not depend on what other players are doing. Also, the larger the prize, V, the more effort an athlete puts into training. Additionally, the optimal effort level is higher for dopers than non-dopers, as doping increases the marginal return to effort.[2]

Now suppose that these contests are repeated multiple times and that the only parameter that varies is the athlete's ability level, r_{it}. Player i at any time t_0 maximizes his or her payoff from the future periods by choosing the optimal effort level as well as best doping decision. From period t_0 to the final period, T, player i's total payoff to be maximized is

$$\sum_{t=t_0}^{T} E(\pi_{it}) = \sum_{t=t_0}^{T}\left[\beta\left(e_{it} r_{it} \delta(d_{it}) - \frac{\sum e_{-it} r_{-it} \delta(d_{-it})}{n-1}\right) + \frac{1}{n}\right] V - C(e_{it}). \tag{5}$$

Observe that in the case without enforcement, decisions are period-wise independent. Since there is no enforcement and ceteris paribus $w_{it}(d=1) > w_{it}(d=0)$, the optimal choice is to dope in all periods and to choose the effort according to Equation 4.

Remark 1. In the penalty-free treatment subjects always dope.

[2] To see this note that $\beta V r_i \delta > \beta V r_i \Rightarrow C^{'}\left(e_i^*(d_i=1)\right) > C^{'}\left(e_i^*(d_i=0)\right) \Rightarrow e(d_i=1)^* > e(d_i=0)^*$.

3.2 Heterogenous N-Player Dynamic Game with Fines

Now suppose a fine system is introduced to prevent athletes from doping. In this dynamic game, dopers have a positive probability (p) of being caught. the fine for being caught is the loss of the total prize share. The chance of being falsely found guilty is zero.

The payoff function for player i in any period, t, is given by the equation

$$E(\pi_{it}^{f}) = \{[1 - p(d_{it})]w_{it}V\} - C(e_{it}) \tag{6}$$

$$p(d_{it}) = \begin{cases} p & d_{it} = 1 \\ 0 & d_{it} = 0 \end{cases}$$

where e_{it}, r_{it}, w_{it} and $p(d_{it})$ represent player i's effort level, ability level, prize share and probability of getting caught. The variable $p(d_{it})$ is a funcion of d_{it}, representing a positive probability of being caught doping, p, for PED-using athletes and 0 for all clean players.

As a result, player i's future payoff starting from period t_0 to the period T can be written as

$$\sum_{t=t_0}^{T} (1 - p(d_{it})) \left\{ \beta \left(e_{it}^{f} r_{it} \delta(d_{it}) - \frac{\sum e_{-it}^{f} r_{-it} \delta(d_{-it})}{n-1} \right) + \frac{1}{n} \right\} V - C(e_{it}^{f}).$$

Player i chooses the optimal training level in every period from the current period to maximize the future payoff under the fine system. Observe that as in the no-penalty regime the choices only impact payoffs in the current period. Hence optimal efforts are given by similar first-order conditions .

We have the following findings on efforts in the fine system

1. If player i does not dope, then the first-order condition is identical to Equation (4)

$$C'\left(e_{it}^{f*}(C)\right) = \beta V r_{it}. \tag{7}$$

2. If player i is a PED user, then first-order condition is

$$C'\left(e_{it}^{f*}(D)\right) = (1-p)\beta V r_{it} \delta. \tag{8}$$

The relationship between $C'\left(e_{it}^{f*}(C)\right)$ and $C'\left(e_{it}^{f*}(D)\right)$ is ambiguous. If $(1-p)\delta > 1$, player i's optimal effort is higher under the doping case (i.e.,$C'\left(e_{it}^{f*}(D)\right) > C'\left(e_{it}^{f*}(C)\right)$), and vice versa. In our experiment the probability of getting caught, p, and the doping efficiency, δ, are set tp 30% and 1.2 respectively, which implies that efforts are higher without doping (, as $(1-p)\delta = 0.7 \times 1.2 = 0.84 < 1$). Similarly we can compare the efforts of dopers and non-dopers in the penalty-free and the fine treatments.

Remark 2. For the same ability r_i a doper will exert higher effort in the penalty-free treatment than in the the fine treatments, while non-dopers will exert the same effort.

We now turn to the doping decision.

1. The expected payoff in any period, t, given player i plays C: the optimal effort for player i is $e_{it}^{f*}(C)$, given $d_{it} = 0$. Thus, the simplified maximised payoff for player i is given by

$$E(\pi_{it}^{f}|C) = \left\{\beta\left(e_{it}^{f*}(C)r_{it} - \frac{\sum e_{-it}r_{-it}\delta(d_{-it})}{n-1}\right) + \frac{1}{n}\right\}V - C\left(e_{it}^{f*}(C)\right)$$

2. Likewise, the expected payoff, given player i plays D is

$$E(\pi_{it}^{f}|D) = (1-p)\left\{\beta\left(e_{it}^{f*}(D)r_{it}\delta - \frac{\sum e_{-it}r_{-it}\delta(d_{-it})}{n-1}\right) + \frac{1}{n}\right\}V - C\left(e_{it}^{f*}(D)\right).$$

Doping pays, whenever

$$E(\pi_{it}^{f}|C) < E(\pi_{it}^{f}|D).$$

With the anti-doping fine, the difference between a doper's payoff and a non-doper's payoff depends not only on their respective performances, but also on the other competitors' average performance, which in itself depends on the competitors abilities, efforts and doping decisions. Since the effort-cost functions are unknown it is difficult to derive additional predictions. We can say the following though.

Remark 3. *Ceteris paribus, in the fine treatment the incentive to dope increases with the expected average performance of the competitors.*

This is the case since the difference between doping and non-doping payoff increases in the average performance of the others. The intuition behind this is that there is less to loose from doping if the competitors perform well and leave little of the share to the athlete.

To summarise, the fine punishment mechanism reduces both the dopers' incentives to dope and their incentives to exert effort. Also, if the probability of getting caught doping, p, and the doping efficiency, δ, are fixed values; doping decisions across different periods are independent in the fine system, meaning that whatever happened in the previous season should have no impact on this season's decisions.

3.3 Heterogenous N-Player Dynamic Game with Ban System

In the ban treatment an athlete who is caught doping will not be allowed to compete next season. The detection probability p remains the same as in the fine treatment. [3]

Player i's payoff function in the current period (t_0), $E(\pi^b_{i,t})$ is identical to that in the penalty-free case because the consequences of being caught only materialize in the next period, hence

$$E(\pi^b_{i,t}) = w_{i,t}V - C(e^b_{i,t})$$

While dopers can keep their prize money in the current period, they will be banned from next season's competitions and will receive a payoff of zero:

$$E(\pi^b_{i,t+1)}) = [1 - p(d_{i,t})]\{w_{i,t+1}V - C(e^b_{i,t+1})\}$$

We focus on efforts first. In the current period (t), player i maximizes the payoff by choosing the optimal effort, $e^{b*}_{i,t}$.

1. Given player i plays C in the current period, t, under the ban system, the first-order condition is
$$C^{'}\left(e^{b*}_{it_0}(C)\right) = \beta V r_{it}. \tag{9}$$

2. Given player i plays D in the current period, t_0, under the ban regime, the first-order

[3] In reality, some bans extend to multiple seasons. In our model, however, the basic case with a one-season ban is studied.

condition is

$$C^{'}\left(e_{it}^{b*}(D)\right)=\beta V r_{it}\delta. \tag{10}$$

Remark 4. Efforts conditional on the doping decision in the ban regime are identical to those in the penalty-free regime.

While the optimal efforts are independent of anticipated behavior in future decisions, doping decisions depend on planned own and expected behavior of others in the future. Obviously, in the final season of an athlete future bans have no deterrence effect.

Remark 5. In the final period all players dope in the ban treatment.

In the penultimate period a player will foresee that if not caught she or he will dope in the final period and therefore will compare the contemporaneous gain from doping and the expected loss from next period. So doping pays if

$$E(\pi_{it}|D)-E(\pi_{it}|C)-pE(\pi_{i,t+1}|D)>0$$

In an earlier period the condition for doping also depends on the planned doping decision in the next period. In general we can say that the doping incentive increases with the ability in the current period and decreases with the ability in the next period. This is the case as the contemporaneous gain from doping $E(\pi_{it}|D)-E(\pi_{it}|C)$ increases with the ability, while the expected loss $pE(\pi_{i,t+1}|D)>0$ increases with the ability in the next season.

Remark 6. Ceteribus paribus, we expect more doping in period $t<T$ the higher r_{it} and the lower $r_{i,t+1}$.

3.4 Heterogenous N-Player Dynamic Game with Conditional Superannuation Fund (CSF) System

The conditional superannuation fund policy refers to an arrangement whereby athletes have to make a compulsory contribution (to their super fund in each season). The contribution is a fixed proportion λ of their season's prize money. Early access to the accrued benefits is prohibited under this mechanism. Athletes can make withdrawals from their super account post-retirement if and only if they are not caught violating anti-doping policy in any period. In other words, if a doper is caught once during their career, they lose their entire fund balance. Moreover, the contribution is compulsory and needs to be made continuously until the end of their sports career. Again, dopers face the same probability of getting caught doping p as in the other regimes.

Under this regime a player who is caught doping in a certain period leads to the loss of past and expected future superannuation payments. This makes the decision for an athlete very complex. However, if an athlete had been caught at any stage in the career, then the future optimal behavior is straight-forward. Recall that the then further doping does not have an additional cost anymore but any future price money will be taxed with tax rate λ. Consequently, all caught dopers will dope in the future and the optimal effort is determined by $C^{'}\left(e_{it}^{s*}(D)\right) = (1-\lambda)\beta V r_{it}\delta$.

On the other hand, if player i has never been caught previously and is coming up to decide in period t_0 the the relevant future payoff is

$$EU \quad = \quad \sum_{t=t_0}^{T}(1-\lambda)\,w_{it}V - C(e_{it}^{s}) + \left(\prod_{t=t_0}^{T}(1-p(d_{it}))\right)\left(\sum_{t=1}^{T}\lambda w_{it}V\right)$$

If a player is planing not to dope in the future, then the optimal effort in all future periods is equal to the non-doping efforts in the other environments.

Remark 7. An athlete, who has not yet been caught and plans never to dope again, will choose the same effort as all non-dopers in the fine and ban treatments.

More generally, the first-order condition for the effort of an athlete in period t_0 with a plan for future doping decision under the CSF regime becomes

$$C^{'}\left(e_{i,t_0}^{s*}(d_{i,t_0})\right) = \left[(1-\lambda) + \left(\prod_{t=t_0}^{T}(1-p(d_{it}))\right)\lambda\right]\beta V r_{i,t_0}\delta(d_{i,t_o}). \tag{11}$$

Remark 8. In the superannuation treatment, for a given doping decision in the current period the current effort declines with the number of planned future dping seasons.

The intuition behind this is simple. The more often future doping is planned the higher the probability of being caught at least once and therefore the lower the expected superannuation return from this periods prize money. Whether a doper optimally exerts more effort than a non-doper is ambiguous (similar to what is found in the fine system). In our experiment, with $p = 0.3$, $\delta = 1.2$ and $\lambda = 0.35$ it turns out that a once-off doper exerts more while repeat dopers should exert less effort than a player who plans to always stay clean.

The doping decision in the CSF treatment is extremely complex as the expected loss from doping in a certain period, depends on the future plans of doping and on efforts for all possible contingencies but also on the superannuation already accrued. Since the contemporaneous gain from doping increases with the own ability, we can at least make the following remark.

Remark 9. In the CSF treatment, ceteribus paribus, the likelihood of doping increases with the ability.

4 Experimental Design

Based on the model, our experiment includes four treatments, namely (i) *a punishment-free* treatment, (ii) a *fine* treatment, (iii) a *ban* treatment and (iv) a *conditional superannuation fund (CSF)* treatment. To complement the above theoretical analysis of the different anti-doping enforcements, we study and compare their effectivenesses in the laboratory setting. All subjects started with a non-competitive real-effort game. Then, the same real-effort task was used in a three-period sports-competition game with a doping option. Subject's payoff was measured in Experimental Currency Units (ECUs) during the experiment and converted into Australian Dollars at a fixed exchange rate at the end of the experiment.

We employ a real-effort task where subjects have to add strings of single-digit numbers mentally to demonstrate their effort level. This task is different from most existing mental arithmetic tasks used in previous research, as our task questions become increasingly difficult. Participants are given a limited time to choose whether to complete the questions. The first question involves adding two single-digit numbers together. After the participant solves this question correctly, the subject moves on to the following question, which involves three new single-digit numbers to be added, and so on. All of these single-digit numbers are randomly generated. We use this arithmetic task because it is easy to understand and requires very little background knowledge.

In each treatment, two different payment schemes, a piece rate and a share-prize contest, are used to reward subjects according to the number of questions they solved. Under the non-competitive piece rate scheme, which forms the baseline task, each individual receives a fixed rate per question solved correctly. In the contest incentive scheme, participants are divided into groups and they share a fixed money prize within the group. The share is determined according to the theoretical contest-success function detailed above, where effort is determined by the number of solved additions. The cost of effort are implemented by paying a fixed rate per second that is not used for solving addition questions. Since the difficulty of addition exercises increases, subjects require more time for an additional solution, which leads to increasing marginal opportunity cost of effort.

4.1 Parameter Selection

In the real-effort task, a total time of 240 seconds is provided to players which can be used to complete additions tasks or for earning an outside wage of 1 ECU per second not used for solving sums. Since the questions become more and more difficult and time consuming to solve, the marginal cost of doing the summation task increases. By contrast, the marginal benefit of completing an extra question is fixed in both payment schemes. As a result, there is an optimal stopping time, where subjects should switch from solving sums to earning outside money.

A typical experiment was conducted as follows. All subjects started the real-effort task with the piece rate scheme first. This piece rate system is supposed help the subjects to understand the core component (the real-effort task). It also served as a reference for subjects' effort decisions in the contest game that they would play later. Every solved sum paid 15 ECUs and every second saved not saved paid 1 ECU. At the end of the piece rate section, subjects' payment in ECUs was recorded and shown to them.

After they completed this part, subjects played three share-prize contest in which the same real-effort task was employed. All participants were randomly assigned into groups of three that remained constant over three consecutive periods. In every period, a prize of 600 ECUs was to be competed for by the three competitors. At the end of every period, each individual received a performance score, which was used to calculate the share of the prize. As in the theoretical model, this performance score q_{it} was multiplicative determined by three factors: *the time-varying ability level* (r_{it}), *the cheating bonus* ($\delta(d_{it})$) and *the effort* (e_{it}), which was equal to the number of solved sums:

$$q_{it} = r_{it}\delta(d_{it})e_{it}.$$

To differentiate players within a group, in each period we assign three different ability levels to the group members: low ($r_{it}^{L} = 2$), medium ($r_{it}^{M} = 3$) and high ($r_{it}^{H} = 4$). Every subject would have a different ability level in each period. So the average ability was identical in each period and all subjects had each of the ability levels exactly once. Subjects knew the ability levels of all three group members for all three periods in advance.

Given the consequences of violating different anti-doping rules in the instructions, players were asked if they want to cheat or not (to improve their scores in that round). Those who chose to cheat received a cheating bonus and their score value increased by 20%, while the score for those who decided not to cheat did not change (i.e., $\delta(d_i) = 1$ if $d_i = 0$ and $\delta(d_i) = 1.2$ if $d_i = 1$). Following that, the subjects had to work in order to establish their effort level. They could choose when to stop solving sums. Once the performances were determined, the shares of the prize were determined. The reactivity to the performance difference, β, was set to a value of 1/100, which gives"

$$Prize\ Money\ in\ ECUs = 200 + 6 \times (Own\ Score - Average\ Score\ of\ Others).$$

In addition to the share of the prize money, players were also paid 1 ECU per second saved from the effort stage, just as in the piece rate game. The same process was repeated in next two periods but with altered ability parameters.

In order to make the three different anti-doping enforcement schemes comparable, it was necessary to choose the parameters such that the expected losses from detected doping were equalized. In the *fine*, *ban* and *CSF* treatments, the probability of detection was set to 0.30. Recall that any subject's gain in the *share-prize competition* consists of two parts: the prize share and the saved time, or:

$$Payoff = Prize\ Share + Time\ Saved.$$

In the *fine* treatment, if one gets caught, he or she loses all the prize money in that round, but could still keep the money from saving their time at the rate of 1 ECU per second. The expected loss for dopers in any period is

$$E(Loss_{fine}) = p \times Prize\ Share = 0.3 \times Prize\ Share.$$

4.1.1 Making the Ban and Fine Treatments Comparable

In the second treatment, the captured doper is banned from the following period and is paid merely a fixed base salary, S, during the period of suspension. Given that dopers who are found guilty of using PEDs are not allowed to participate, their share prize is automatically forsaken and therefore, the expected loss for first-time dopers in this case is

$$E(Loss) = p \times (Prize\ Share + Time\ Saved) - pS = 0.3 \times (Prize\ Share + Time\ Saved) - 0.3S.$$

Under the assumption of identical doping behavior and effort exertion in the *fine* and *ban* treatments, setting expected losses under both schemes equal allows us to determine a residual payment for a banned subject. I turns out that using the behavior of subjects in the *fine* treatment as the common behavior a residual payment of $S = 110$ ECU equalizes expected losses. Note that these expected losses are only accruing in the first two periods, since a ban has no deterrence effect in the last period.[4]

4.1.2 Making the CSF and Fine Treatments Comparable

In the *CSF* treatment, the compulsory contribution, λ, needs to be determined. Assuming that θ is the probability that an individual dopes in any period the total expected loss under the *CSF* system can then be expressed as the probability of getting caught at least once multiplied by the total contribution over the three periods:

$$E(Total\ Loss_{super}) = (1 - (1 - p\theta)^T)(T\lambda Prize\ Share) = (1 - (1 - 0.3\theta)^3)(3\lambda Prize\ Share)$$

and that in the *fine* system, that is:

$$E(Total\ Loss_{fine}) = Tp\theta \times Prize\ Share = 3 \times 0.3\theta \times Prize\ Share.$$

To solve for the contribution rate, the two equations are equalized (i.e., $(1 - (1 - 0.3\theta)^3)\lambda = 0.3\theta$). According to the underlying assumption of identical behaviors in the *fine* and *CSF* treatments, the probability that a player dopes in any period, θ, is therefore taken directly from the *fine* treatment data. There are 11 dopers out of 63 participants in every period under the *fine* system, producing a probability of 11/63 if a random player and period is picked. Solving out,

[4] The problem of a missing competitor in a group due to a ban was solved by substituting the performance of another player in the session into the share-price calculation.

the period CSF contribution rate becomes 35% of the prize money.

4.2 Laboratory Implementation

The experiment used subjects drawn from the population of the University of Adelaide Experimental Lab (AdLab) Database. Sessions for this experiment were conducted at the computer laboratory of the University of Adelaide from November 2014 to December 2015. There was a total of 270 participants in twelve sessions for four treatments, which were recruited using Ben Greiner's ORSEE. Each of them were allowed to register for one session only. Before the experiment started, subjects were randomly assigned to a computer in the lab. Communication was forbidden once the subjects entered the lab. The experiment were programmed in Urs Fischbachers *Z-tree*. The instructions were provided at the beginning of each part were read aloud by the experimenter. Before the competition part of the actual experiments began subjects had to answer a set of control question regarding the contest success function. Subjects were paid at the end of the experiment at the exchange rate of one AUD for 70 ECU, and subjects earned about 17 AUD on average for one hour of their work.

5 Results

This section presents our results. Our aim is to identify which anti-doping regime does best in a) preventing doping without b) overly reducing efforts. First, we begin with aggregate results that provide an overview over aggregate differences across treatments. A deeper analysis of doping behavior follows. Finally, compare effort levels in four treatments.

5.1 Aggregate Results

In our experiment, group members remain fixed for all three periods, which makes every group an independent observation point.[5] As a result, group average values are used in the non-parametric two-sample Wilcoxon rank-sum test used to investigate treatment differences. Table 1 shows the group doping fraction, group average effort level and group average payoffs in our four treatments, (i) *No punishment* system; (ii) *Fine* system, (iii) *Ban* system, and (iv) *Conditional Superannuation Fund* treatment, under both *piece rate* and *share-prize contest* payment schemes. The fraction of doping is calculated as the total number of periods experimental subjects doped divided by the

[5] Except for those groups that had suspended subjects in the *ban system*.

total number of doping decisions within a group over the three periods. In treatments (i), (ii) and (iv), the total number of doping decisions made is nine. In the *Ban* treatment, group average values are calculated by excluding the replacements for suspended players'. Here the total number of doping decisions depends on the actual number of non-banned participants within a group.

Table 1: Fraction of Doping, Average Effort and Average Payoffs in the Four Treatments

Types of Punishments	N Participants	Probability of Doping	Average Effort	Average Payoffs
Piece Rate				
No punishment	78		9.385	250.859
Fine	63		9.206	240.349
Ban	60		10.117	237.433
CSF	69		10.101	218.420
Share-prize Contest				
No punishment	78	0.872	9.350	974.518
Fine	63	0.175	8.899	936.518
Ban	60	0.306	9.852	953.453
CSF	69	0.106	9.725	900.920

According to Table 1, the probability of doping in three treatments with enforcement is significantly lower compared to that in the *No-punishment* treatment ($p < 0.0001$ for all pairwise comparisons). In terms of the doping fraction in the three treatments with punishments, the *Fine* and *CSF* treatments tend to have a lower probability of doping than the *Ban* treatment, with $p = 0.0095$ and $p = 0.0002$, respectively. The lowest probability of doping is observed in the *CSF* treatment. However, it is not significantly different from that in the *Fine* treatment ($p = 0.1649$). On aggregate, our anti-doping policies discourage subjects' doping effectively compared to the *No-punishment* treatment. Furthermore, the *CSF* and *Fine* systems are more effective than *Bans*.

The average effort level is higher in the *piece rate* game than in the *contest* for all treatments. This effort difference between the two incentive schemes is, however, not significant given the p-values are 0.9220 *for* the *No-punishment*, 0.7803 *for* the *Fine*, 0.4800 for the *Ban*, and 0.9343 for the *CSF* treatments. In the *piece-rate* game, all subjects were provided with identical instructions regardless of which treatment they participated in, and therefore, no treatment effects should be observed. If there are any, they are credited to natural individual differences. In our case, treatments have no significant impact on average effort under the *piece rate* scheme (p-values of 0.6215 for *No-punishment* versus *Fine,* 0.4038 for *No-punishment* versus *Ban,* 0.2639 for *No-*

punishment versus *CSF,* 0.1826 for *Fine* versus *Ban* and 0.8401 for *CSF* versus *Ban,* respectively). However, there is a weakly significant effect comparing the *CSF* to the *ban* treatment ($p = 0.0974$). In the *share-prize contest,* in which treatment differences could be discovered, surprisingly, no effects on effort level can be established except for the *CSF* treatment, which has a weakly significant positive effect on effort level compared to the *Fine* system ($p = 0.0602$), in line with that illustrated under the *piece-rate* scheme. It appears to be the case that the sample that participated in the *CSF* treatment was a bit better at solving addition tasks than those sampled into the fine treatment. Similarly there are no big differences in payoffs in the share contests. Aside from the *Ban* policy effect ($p = 0.0602$ when compared to *Fine* treatment), payoff differences in the *share-prize contest* across treatments are remarkably small and not significant.

Figure 1 shows the evolution of the proportion of dopers over the three periods in all four treatments. Unsurprisingly, the *No-punishment* treatment has the highest proportion of dopers with values above 80% for all three periods. Based on our model, all players should dope under this system because there are no consequences of doping. The reason behind the existence of some non-dopers might be some moral cost. On the other hand, the probability of doping within the *CSF* treatment is always lowest among all punishments in each period and lies in the range from 10% to 17%. The proportion of dopers under the *Fine* system is slightly larger than that under the *CSF.* As for the *Ban* treatment, doping fractions in the first two periods are similar to those observed in the *Fine* and *CSF* treatments, while this proportion soars to a much higher level (60%), in the last period. This is as predicted by the theory. With respect to doping prevention fines and conditional superannuation are preferable to bans, since they do not suffer from the lack of deterrence in the last season of a competitor.

Figure 1 about here

Figure 2 illustrates an overview of the effort level in the four treatments. Subjects' average effort levels typically lie within the interval $[8.5, 10]$. A monotonically-increasing effort level over the three periods is found under the *No-punishment, Fine* and *Ban* treatments, indicating that subjects from these three treatments became more and more competitive in the *contest.* Moreover, subjects in the *Ban* treatment exerted the highest effort level in all the periods compared to those who were in the *No-punishment* treatment. Participants under the *Fine* system had the lowest effort level in all periods. As far as the *CSF* treatment goes, the average effort level decreases from period 1 to period 2, but then climbs to a new higher level in period 3.

Figure 2 about here

In Figure 3, the trends of average period profit in the *contest* for the four treatments are presented. Firstly, the *No-punishment,* the *Fine* and the *Ban* system have very similar average profits and all of them experience a moderate decrease in profits over the three periods. This profit decline is consistent with the increasing competitiveness in the *contest* as displayed in Figure 2. Recall that subjects in the *CSF* treatment were required to make a contribution out of their prize money into their super account and the account balance will be returned to them as a lump sum in period 3 if they are not caught cheating at all. Due to this property, profits are much lower in the first two periods of the *CSF* treatment but much higher in the final period, compared to the other three treatments.

Figure 3 about here

To understand subjects' individual differences in treatments, we regressed effort, optimality and profits in the *piece rate* on *innate mental calculation ability* and the treatment dummies: (i) *No-punishment,* (ii) *Fine,* (iii) *Ban* and *(iv) CSF;* controlling for possible differences in demographic factors: *age, gender* and *highest education level.* The dependent variable, optimality, is calculated as: $Marginal\ Benefit - Marginal\ Cost$ for every subject. In theory, the payoff is maximised when subjects choose to equalise their marginal benefit with their marginal cost. Marginal benefit in the *piece rate* is fixed at 15 ECUs while marginal cost varies given the different amount of time individuals spend on the real-effort task. Recall that for every second used for solving the task questions, 1 ECU that could have been saved is forgone. Thus, marginal cost is equivalent to the seconds spent on an additional question. As a consequence, a positive (negative) optimality suggests an under-exertion (over-exertion) of effort. One of the control variables, *innate mental calculation ability*, measures subject's innate ability to do mental arithmetic. To derive this variable, an individual's accumulated time to reach their final effort level is obtained in the *piece rate*. Then, for every effort level, there is a quickest person with the smallest amount of accumulated time spent and this time is used as the benchmark for the subject's innate ability. The variable is

$$\frac{accumulated\ time\ of\ the\ quickest\ person\ who\ solved\ effort\ level\ x}{accumulated\ time\ of\ an\ individual\ with\ effort\ level\ x}.$$

In our case, the quickest subject has an *innate mental calculation ability* of one and the slowest subject tends to have a *mental calculation ability* closer to zero.

Table 2 lists the result of Tobit estimations of effort level and OLS estimations of optimality and profits in the *piece rate*. Obviously, *innate mental calculation ability* has a strong positive effect on effort level, optimality and profits in the *piece rate*. Subjects in the *CSF* and *Ban* treatments exerted a significantly-higher effort level but received a lower profit (significant only for the *CSF* treatment). These treatment differences can either stem from heterogeneity, such as different personal preferences in attempting the real-effort task and different attitudes towards competitions or could be due to the different enforcement system. The former explanation is in line with the rank-sum test result, as subjects in both the *Ban* and *CSF* treatments were more competitive by default than those in the *Fine* treatment (p-values of 0.0547 and 0.0904, respectively).

Table 2: Tobit/OLS/OLS Estimations of Effort Levels, Optimality and Profits in the Piece Rate

Model	Tobit	OLS	OLS
Dependent	Effort	Optimality	Profits
Intercept	5.6346***	-18.9866***	128.8115***
	(0.7376)	(4.0534)	(7.6695)
Innate Mental Calculation Ability	12.7335***	37.8998***	253.699***
	(1.3875)	(7.5680)	(14.3195)
Fine	0.3141	0.3057	-1.9702
	(0.5011)	(2.7740)	(5.2487)
Ban	1.1567**	-4.3296	-9.2513*
	(0.5031)	(2.7832)	(5.2660)
CSF	1.6997***	-3.8490	-20.9683***
	(0.4920)	(2.7228)	(5.1517)
N	266	266	266
Pseudo R^2	0.0672		
N Left-Censored	4		
N Right-Censored	1		
Adjusted R^2		0.1024	0.6468

***Significant at the 1 per cent level.
**Significant at the 5 per cent level.
*Significant at the 10 per cent level.

5.2 Doping Behaviours

We conduct a random-effect probit regression of a binary variable, *dope*, and a random-effect tobit regression of *effort*[6] on *innate mental calculation ability*, ability dummies: *low*, *medium* and *high*, period dummies and treatment dummies, controlling for demographic factors as well as trends over time. By employing a random-effect model, the effect of unobserved heterogeneity is captured and controlled for. Under the *Ban* system, we are mainly interested in its deterrence effect in the first two periods since the doping behavior in the final period is unchecked. Consequently, the final period effect in the *Ban* is isolated by the use of a dummy.

Table 3 lists the results of these two regressions. Firstly, the probit regression shows a strong negative treatment effects of the *Fine, Ban* and *CSF* on *doping* compared to the baseline *No-punishment* treatment (with the same three p-values of 0.0000), which confirms the rank-sum test results as well as the theoretical prediciton. On average, the *Fine, Ban* and *CSF* treatments decrease the probability of doping by 51.01%, 43.40% and 57.47% respectively.

Result 1: Comparing to the *No-punishment* treatment, the predicted doping probability is significantly lower in the *Fine, Ban* and *CSF* treatments.

Among the *Fine, Ban*, and *CSF* treatments, the only significant difference in the treatment effect on *doping* is observed between the *Fine* and *CSF* which a p-value of 0.0658. Namely, there is a (weakly) significant smaller probability of doping in the *CSF* treatment compared to the *Fine* treatment. As expected, being in the third period under the *Ban* system increases the probability of doping significantly (by 34.63%).

Result 2: The *CSF* system is more effective in deterring doping compared to the *Fine* system. The *Ban* system does not have a significantly different deterrence than the other two policies when we consider only the first two periods. However, after taking the last period effect into consideration, the *Ban* system shows the highest doping rate among three anti-doping policies.

Surprisingly, both the *innate mental calculation ability* as well as the ability level assigned have no significant effect on the doping decision. In most cases theory predicts that doping should be most likely when ability is high or when effort cost are low. Also, no trend of doping can be observed over the three periods.

[6]There are three extra control variables in the effort tobit regression: the *benchmark effort* (effort in the *piece rate*), *dope* and the *interaction* variable of *treatment* and *dope*.

Effort levels are positively affected by the benchmark effort level in the *piece rate* which indicates subjects' ability in and enjoyment of the real-effort task. We also notice a strong positive effect of the *innate mental calculation ability* on effort level. More specifically, on average, the quickest subject completed 5.27 more questions than the slowest subject. As theory predicts, there is a strong positive effect of the assigned ability on effort level. This shows subjects react to the incentives provided by the opportunity cost in the expected way. Moreover, there is a possible learning effect, which lead to an increase of competitiveness in the real-effort task over time, indicated by the time dummies with significantly positive marginal effects.

We test the differences between coefficients of treatment dummies and find that non-dopers have the same effort levels across different treatments (i.e., No significant effect is observed between any pair of the treatment dummies). This is in line with the theory predictions and also tells us that none of the anti-doping policies has a specifically negative impact on the performances of those who comply.

In addition, *doping* and its interaction with *treatment* is tested for impacts on effort decisions. Dopers in the *Ban* treatment put in the highest effort (on average, they solved 2.68 more questions than dopers in the *No-punishment* treatment). This is followed by dopers in the *CSF* treatment with an average of 2.48 more questions completed than *No-punishment* dopers. The dopers in the *Fine* treatment solved 1.16 more questions. We discover that dopers in the *Ban* and *CSF* treatments exerted significantly higher effort than dopers in the *Fine* and *No-punishment* treatments, while no significant differences are noticed between the *Ban* and *CSF* treatments and the *No-punishment* and *Fine* treatments (*p-values* of 0.0239 for *No-punishment* versus *Ban* treatments, 0.0505 for *No-punishment* versus *CSF* treatments, 0.0078 for *Fine* versus *Ban* treatments and 0.0473 for *Fine* versus *CSF* treatments). This result contradicts theory which predicts that dopers in the *No-punishment* treatment exert the same effort level as those in the *Ban* treatment.

Result 3: As our model predicts, non-dopers have the same average effort level regardless of what kind of anti-doping policies they face. Hence, anti-doping policies do not impact those who stay clean. On the other hand, dopers' effort levels are influenced by the different anti-doping systems. Specifically, an effort-level ranking among different treatments of $Ban = CSF > Fine = No - punishment$, is observed.

Finally, we test the difference between the treatment dummies and the interaction variables to investigate the difference in efforts between the non-dopers and dopers within a treatment. In the

No-punishment treatment, on average, dopers completed approximately one question less than the non-dopers (*p-value*=0.077). Theory predicts the opposite. In the *Ban* treatment, dopers were more competitive than non-dopers such that dopers solved two real-effort questions more than non-dopers (*p-value*=0.0527). In the *Fine* and *CSF* treatments, dopers on average completed two and one task, respectively, more than non-dopers. This difference is not significant though. Therefore, we can conclude that our punishment mechanisms (*Fine, CSF* and *ban*) do not reduce the intensity of the competition to a degree not predicted by theory. Note that reductions of effort of compliant athletes compared to dopers in the *ban* treatment is in line with predictions.

Result 4: In the *Ban* treatment, dopers exert significantly higher efforts than non-dopers, whereas dopers in the *No-punishment* surprisingly exert significantly lower effort levels than non-dopers. In the other two treatments, the effort differences between dopers and non-dopers are not significant.

Therefore, concerns that a conditional superannuation fund could reduce efforts more than other policies when inducing athletes to become compliant are not warranted.

Table 3: Random-Effect Probit/Random-Effect Tobit Estimations of Dope and Effort in the Contest

Model	Random-Effect Probit	Random-Effect Tobit
Dependent	Doping	Effort
Intercept		4.1496***
		(0.9460)
Benchmark Effort		0.3921***
		(0.0605)
Innate Mental Calculation Ability	-0.1289	5.2680***
	(0.1043)	(1.5340)
Medium-Ability	-0.0239	0.9879***
	(0.0225)	(0.1836)
High-Ability	0.0290	1.2808***
	(0.0236)	(0.1842)
Period 2	0.0078	0.3394*
	(0.0234)	(0.1814)
Period 3	0.0173	0.4736**
	(0.0266)	(0.1852)
Dope		-0.9795***
		(0.5534)
Fine	-0.5101***	-0.9119
	(0.0242)	(0.6931)
Ban	-0.4339***	-0.7695
	(0.0225)	(0.7024)
CSF	-0.5748***	-0.5705
	(0.0194)	(0.6882)
Fine#Dope		-0.1771
		(0.7919)
Ban#Dope		1.6957***
		(0.6964)
CSF#Dope		1.5025*
		(0.8356)
Ban Last Period	0.3463***	
	(0.0791)	
N	810	810
N Left-Censored		20
N Right-Censored		7

In the probit model of dope, instead of coefficients, we report marginal effects.
*** Significant at the 1 percent level
**Significant at the 5 percent level
*Significant at the 10 percent level

Finally, Figure 4 displays the variable optimality for low-, medium-, and high-ability players. Within these ability groups, there are two subgroups: compliant and players and dopers. Recall that a zero value of the variable indicates optimal effort exertion, while a positive (negative) value means under (over) exertion of effort. We notice an obvious trend that optimality increases with ability level, from a negative value for low-ability players to a positive value for high-ability subjects. In other words, on average, low-ability players over-exerted effort whereas high-ability players did not put in as much effort as they should have. Medium-ability players, on the other hand, are almost at the margin. So subjects react to the incentives provided when choosing their effort levels in the expected way. However, the reactivity is lower than what theory would predict.

Figure 4 about here

6 Conclusion

Empirical evidence on doping behavior is very rare, as reliable field data is unavailable. Doping behavior cannot be observed easily in real-world competitions since it is always conducted in secret. Our solution to this observability problem is the use of laboratory experiments. This approach is clearly second-best, as the external ability depends on the artificial laboratory environment emulating the salient elements of the real world.

This paper finds that the probability of doping is significantly less in treatments with punishments compared to *No-punishment* treatment. Moreover, the probability of doping in the newly proposed anti-doping *CSF* treatment is significantly lower than in the *Fine* treatment. By controlling for the last period effect in the *Ban* treatment, the *CSF* and *Ban* treatments are not significantly different in deterring subjects from doping. However, after including the last period of the *Ban*, the *CSF* treatment is more effective. We conclude that our newly proposed anti-doping policy of a compulsory superannuation system, where an athlete forfeits all claims if he is caught doping at any time in his career, can potentially improve compliance. Beyond the good performance in the experiments it has also the advantage that it can cope with the fact that dopers are often only found out many years after the act, when detection technology has caught up with doping practices. Fears that the conditional superannuation system could reduce training efforts and therefore the quality of competition to a larger extent than other policies proved to be unfounded.

References

Berentsen, A., E. Bruegger, and S. Loertscher (2008). On cheating, doping and whistleblowing. *European journal of political economy 24*(2), 415–436.

Cason, T. N., W. A. Masters, and R. M. Sheremeta (2010). Entry into winner-take-all and proportional-prize contests: An experimental study. *Journal of Public Economics 94*(9), 604–611.

Curry, P. A. and S. Mongrain (2009). Deterrence in rank-order tournaments. *Review of Law & Economics 5*(1), 723–740.

Ehrenberg, R. G. and M. L. Bognanno (1988). Do tournaments have incentive effects?

Ellingsen, T. and M. Johannesson (2004). Promises, threats and fairness*. *The Economic Journal 114*(495), 397–420.

Faravelli, M., L. Friesen, and L. Gangadharan (2014). Selection, tournaments, and dishonesty. *Journal of Economic Behavior & Organization.*

Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in disguise—an experimental study on cheating. *Journal of the European Economic Association 11*(3), 525–547.

Friesen, L. and L. Gangadharan (2012). Individual level evidence of dishonesty and the gender effect. *Economics Letters 117*(3), 624–626.

Gill, D. and V. Prowse (2012). A structural analysis of disappointment aversion in a real effort competition. *The American Economic Review 102*(1), 469–503.

Gilpatric, S. M. (2011). Cheating in contests. *Economic Inquiry 49*(4), 1042–1053.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review*, 384–394.

Lazear, E. P. (1989). Pay equality and industrial politics. *Journal of political economy*, 561–580.

Lazear, E. P. and S. Rosen (1979). Rank-order tournaments as optimum labor contracts.

Mazar, N., O. Amir, and D. Ariely (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of marketing research 45*(6), 633–644.

Niederle, M. and L. Vesterlund (2005). Do women shy away from competition? do men compete too much? Technical report, National Bureau of Economic Research.

Robert, I. and M. Arnab (2013). Is dishonesty contagious? *Economic Inquiry 51*(1), 722–734.

Ryvkin, D. (2013). Contests with doping. *Journal of Sports Economics 14*(3), 253–275.

Van Dijk, F., J. Sonnemans, and F. Van Winden (2001). Incentive systems in a real effort experiment. *European Economic Review 45*(2), 187–214.

A Figures

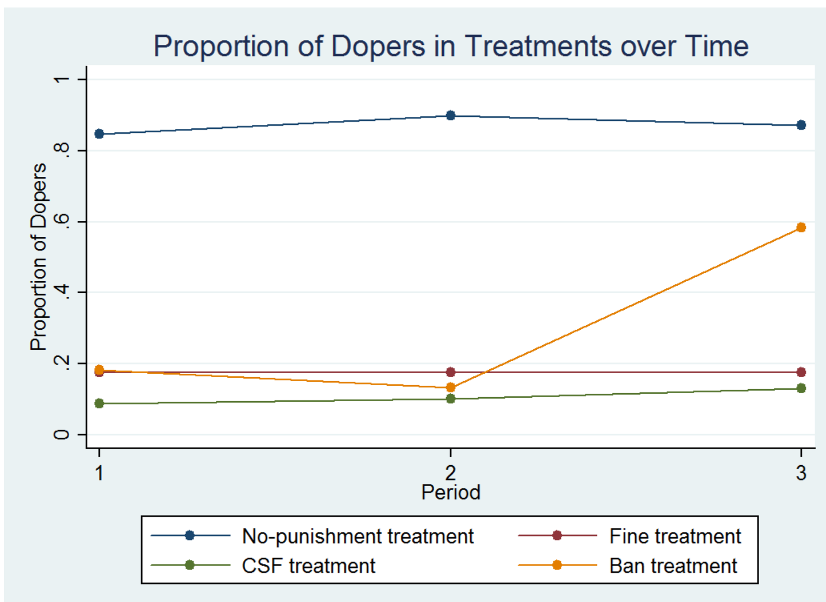

Figure 1: Probability of Doping Over Time

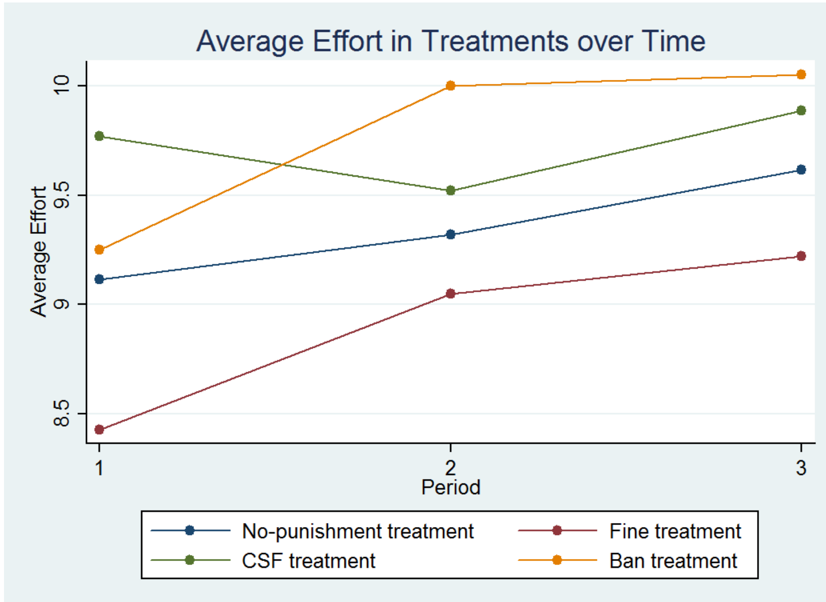

Figure 2: Average Effort Levels Over Time

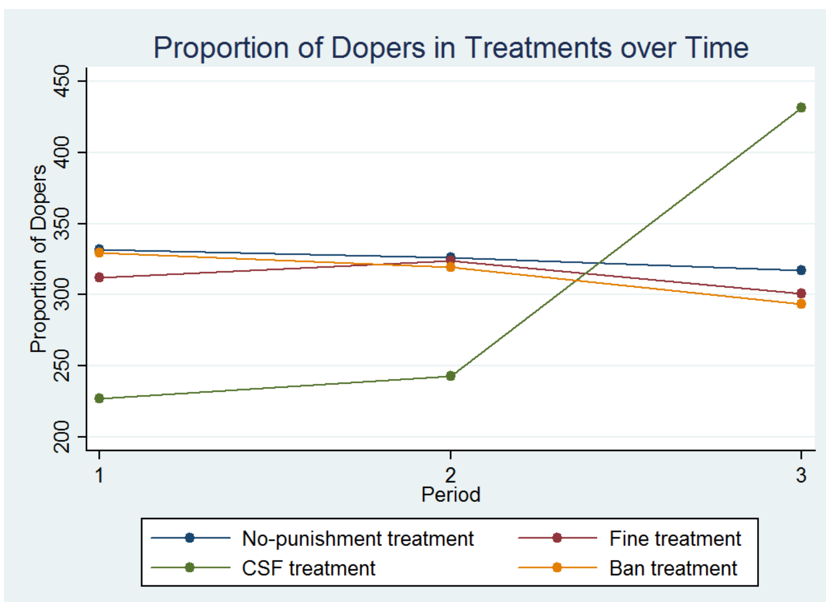

Figure 3: Average Period Profit Over Time

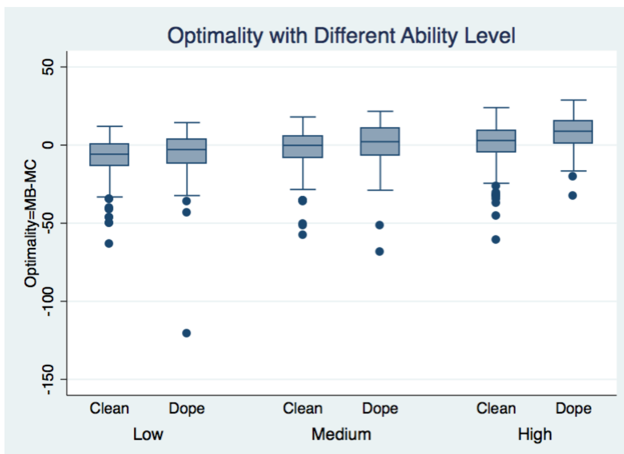

Figure 4: Optimality for Different Ability Levels