Performance-Enhancing Drugs and the Production of Season Best Performances in Track and Field

Allison Fisher and Michael C. Seeborg*

Abstract. This paper explores trends in the production of post WWII track and field performances in six events. Of particular concern is how performances deviated from long-term trend from 1962 through 1990, a period when elite track and field athletes were known to systematically use steroids in an environment of ineffective testing. Log-linear trend regression for each event and their residuals are used to develop an indicator variable of whether the actual performances of elite athletes were better than trend. Probit models are then used to predict the probability of better than trend performance for three time periods. The results show very high probabilities of above trend performance during the heavy use period and much lower probabilities of above trend performance after 1990, a period when testing was more effective and international anti-doping efforts more coordinated. (J24)

I. Introduction

Recent history demonstrates that many track and field athletic feats are tarnished by the use of performance enhancing drugs. While the motivations for taking these drugs are subject to much public speculation, one rarely views such decisions through an economics lens. Each track and field event can be thought of as a production process with several immeasurable inputs producing one precisely measured output. This precise output, performance of the athlete, can be measured to the thousandth of a second for running events or to the millimeter for field events. However, measuring the quantities and qualities of productive inputs such as anabolic steroids is much more difficult. Despite the measurement difficulties, significant evidence indicates that track and field performances have been affected by steroid use among elite athletes, especially during a period beginning in the 1960s (Yesalis, 2000).

It is also clear that there has been increasing enforcement of rules against doping over the past two decades. A well publicized recent occurrence was in 2007 when sprinter Marion Jones pleaded guilty to lying to federal investigators regarding allegations of steroid use (Wilson, 2008). She joins a long list of track stars including C.J. Hunter, Tim

*Illinois Wesleyan University. E-mail: mseeborg@iwu.edu

Montgomery, Ben Johnson, and Justin Gaitlin that have been stripped of medals or records and banned from competing (International Association of Athletics Federations, 2008). Do these recent high profile cases of elite athletes being caught mean that testing and anti-doping efforts by national and international organizations are finally becoming effective? Evidence of the effectiveness of these recent anti-doping efforts can be inferred by trends in performance in individual events.

The economics literature on the effects of performance enhancing drugs by elite athletes in track and field is limited. However, there is a general body of theoretical literature on the economics of superstars (Rosen, 1981) and the economics of crime (Becker, 1968) that suggests why elite athletes might use illegal and potentially dangerous performance enhancing drugs. Sherwin Rosen (1981) explores the economics behind the extraordinary earnings of superstars. Performing at the highest level can result in extraordinary rewards with incomes distributed very unequally between the few who achieve superstar status and the many very good performers who do not (Fort, 2002). In track and field events, only the very top performers can expect to receive widespread fame and the corresponding large income from sponsorships and advertising contracts. For some of these, the use of performance enhancing drugs could be a necessary condition for achieving superstar status.

The difference between the superstars and the rest of the field is often measured in fractions of a second for runners or millimeters for throwers and jumpers. Thus, a small improvement in performance could result in large pecuniary and non-pecuniary benefits to the athlete. Elite athletes who use performance enhancing drugs likely believe that these drugs can give them enough of a performance edge that they might reach superstar status. These athletes may choose to combine illegal performance enhancing drugs with legal inputs, such as coaching, training programs, shoe quality, track and field training surfaces, to produce a performance output sufficient to reach superstar status.

While elite track athletes may be ethically opposed to breaking rules against use of performance enhancing goods, many have obviously succumbed to the temptation. Gary Becker (1968), in his landmark work on the economics of crime, posits that a crime will be committed when the benefits are high relative to the costs and probability of apprehension. Thus, according to Becker's cost-benefit approach to crime, we can conclude that even athletes with a natural aversion to cheating might still chose to use performance enhancing drugs.

Economic literature indicates that anti-doping policy might be ineffective where there are problems with policy credibility and where strategic doping behavior could push athletes to choose the use of performance enhancing drugs. Eber (2002) argues that anti-doping policy faces an inherent credibility problem unless conducted by international bodies that are independent and with authority to enforce penalties. Haugan (2004) uses a game theory framework to point out a disturbing strategic outcome of a two player doping game where the decision to use performance enhancing drugs can lead to a "prisoners dilemma" where both players choose to take drugs even when weak anti-doping measures are in place.

Taken together, the economic theories of superstars (Rosen, 1981), crime (Becker, 1968), policy credibility (Eber, 2002) and strategic doping behavior (Haugan, 2004) suggest that athletes might rationally choose to take performance enhancing drugs in their quest to reach superstar status, unless anti-doping policies significantly increase the likelihood of apprehension and imposes sufficiently severe penalties on violators. These theories are useful in thinking about secular trends in track and field performances since 1949, a period that is long enough for new doping technologies to be introduced and for the intensity of anti-doping efforts to change.

During the 1960s, technological changes increased the effectiveness and affordability of performance enhancing drugs, especially anabolic steroids. Since the probability of apprehension was very low and punishments weak until the 1990s, Becker's (1968) theory would suggest heavy use of performance enhancing drugs by elite athletes who strive to reach "superstar" status. The basic idea is that the high potential gains and relatively low risk of apprehension provides incentives for players in the drug-use game to move to a "prisoner's dilemma" outcome where elite performers choose to use performance-enhancing drugs in hope of achieving superstar status (Haugen, 2004). During this period of heavy steroid use, track and field fans should be treated to above trend performances from elite athletes.

However, beginning around 1990, coordinated anti-doping enforcement became more technologically sophisticated and public disapproval towards doping increased (Yesalis, 2000). The probability of apprehension increased and the penalties were often severe with elite athletes being banned from participation and often stripped of medals and records. Becker's (1968) economics of crime theory suggests that use of illicit performance enhancing drugs should decrease during this period because of the increased probability of apprehension and the high penalties being levied by international governing bodies.

The main purpose of our study is to explore how actual performance of elite track and field athletes in six events deviated from trend performance over three time periods: a period of limited availability of anabolic steroids (1949-1961); a period of increasing availability of performance enhancing drugs linked with ineffective anti-doping efforts (1962-1990); and a period of more effective testing and increasing coordination between anti-doping programs (1991-2007). The procedure used to determine the start and end date of the second time period is described in Section IV of the paper.

A second purpose of this study is to determine whether gains in performance during the heavy steroid use period are less for distance runners when compared to middle distance runners, sprinters and field event athletes. Research shows that muscles consist of two types of fiber, slow twitch fiber and fast twitch fiber. It also has established that anabolic steroids cause twice the amount of growth for fast twitch fiber than in slow twitch fiber (Dare, 1979; Yesalis, 2000). Sprinters typically have only 8-24 percent of slow-twitch fibers with the rest being fast-twitch. Distance runners, on the other hand, typically have at least 80 percent slow-twitch fibers. Between these two extremes are throwers and jumpers with 40-50 percent slow-twitch fibers and middle distance runners with 40-65 percent slow-twitch. (Dare, 1979). Since distance runners have a higher percentage of slow twitch muscle fiber, they are not likely to benefit as much as athletes in the other events from anabolic steroid use. Additionally, Taylor (2002) found that steroids improve an athlete's reaction time. Clearly, this is another benefit to sprinters since they rely heavily on their ability to get out of the blocks faster than the competition. Dropping even one hundredth of a second from the reaction time can be important for a sprinter, whereas for other event groups, it would not be as significant. Thus, research in sports medicine supports the claim that sprinters should benefit the most from anabolic steroid use, distance runners the least and field event specialists and middle distance runners somewhere in between.

Based on the history of steroid use patterns and the medical evidence indicating that steroids influence sprint and power events more than endurance events we posit the following three hypotheses:

1. Track and field performances of elite athletes should be consistently better than long-term trend during the heavy steroid use period (1962 through 1990)
2. Endurance events, like the 5000 meter run, should realize less of an improvement in performance during the heavy steroid use period than middle distance runs, sprint events and field events.
3. Track and field performances of elite athletes should be consistently worse than long-term trend during the period of effective testing and coordinated anti-doping programs (1991 through 2007).

In the following section we review literature that deals with trends in the performance of track and field athletes in light of how those trends can be affected by the use of performance enhancing drugs. Section III describes the season best performance data set and Section IV develops the OLS and probit models that are used to analyze performance trends. Section V presents results that show significant patterns of above-trend performance during the heavy steroid use period.

II. Measuring Track and Field Performances over Time

Several studies conduct longitudinal analyses of track and field performances of elite athletes. Munasinghe, O'Flaherty and Danninger (2001), for example, explore the effects of globalization and technological progress on the number of records broken in track and field from 1886 through 1994. Although they are not able to identify specific sources of technological progress, they mention a set of factors, including improvements in running surfaces, shoes, equipment, training techniques, nutrition, and so on. They find that the number of records broken in track and field remained fairly high over the decades and attribute this to technical change and globalization. Surprisingly absent from their analysis is any consideration of the effects of performance enhancing drugs on record breaks.

Liu (2004) argues that it is best to use season best performance data rather than the more idiosyncratic and discontinuous world record data for analysis and prediction. Season best data is non-zero every year and representative of the performance of elite athletes. Schutz and Liu (1998) analyze the trends of season bests of seven different events for track and field using linear and exponential models, but do not take into account the effects of performance enhancing drugs.

Literature on the history of steroid use in sports indicates that the beginning of the heavy-use period is in the 1960s and that the end is around 1990. While the history of steroids dates back to ancient Greece with testosterone use, anabolic steroids make their appearance much later (Hoberman and Yesalis, 1995). Daniel Rosen (2008), in his history of doping in sports, established that elite weight lifters were systematically using testosterone and other anabolic steroids at the beginning of the 1960s. Since steroids clearly had a positive effect on the performance of these weight lifters, steroid use quickly spread to sports where weight training was important, including track and field. Yesalis (2000) and Dimeo (2007) suggest that by the 1964 Olympics a substantial number of track and field athletes were using steroids, especially throwers and power lifters, and by the end of the decade steroids were being used by runners as well.

Steroid use regimens became more sophisticated as early as the 1960s with some countries such as East Germany actively supporting these programs in order to compete more effectively in international events such as the Olympics (Pound, 2004). According to Yesalis, Cousin and Wright (2000), by the late 1960s, dosages were much higher than therapeutic recommendations for replacement therapy, and athletes began using a variety of anabolic steroids simultaneously (stacking) to maximize their effects. By the 1968 Olympics, the use of anabolic steroids was widely acknowledged in the press and seems to have been generally accepted by athletes and fans. Francis and Coplon (1990) conjectured on the basis of insider estimates that by 1972 about 80 percent of the top male athletes were using steroids.

It was not until 1975 that the International Association of Athletics Federation banned steroids from competition and systematically added more performance enhancing drugs to the list (Francis and Coplon, 1990). However, testing techniques were not very effective during the 1970s and 1980s. For example, Ben Johnson passed approximately two or three drug tests per year despite being on a heavy regimen of multiple steroids (Francis and Coplon, 1990).

There is some evidence that the use of steroids has been on the decline since the early 1990s. For example, the 1996 Atlanta Olympics were very clean with only two of the two thousand athletes checked for steroids testing positive ("Superhuman Heroes," 1998). This decline could be, in part, because of increasing coordination of anti-doping efforts across countries over the past twenty years. For example, the Anti-Doping Convention of the Council of Europe was established in December, 1989, and the World Anti-Doping Agency was established in 1999. These coordinated efforts received significant public support because of increasing concerns regarding the safety of steroid use, ethical concerns over fairness, and concerns that the use of steroids by elite role model athletes could influence millions of young athletes (Yesalis, 2000).

Thus, it appears that heavy use of anabolic steroids in track and field began in the early 1960s and that effective testing and sanctions were in place around 1990. We chose the precise break dates (1962 and 1990) econometrically using a procedure that is described in Section IV.

An important assumption underlying our analysis is that improvements in performance due to changes in non-drug related productive inputs are likely to be continuous and steady. For example, coaching and physical training inputs per athlete have likely increased steadily as more athletes train in well equipped specialized training facilities. There also appears to have been steady technological improvements in the quality of inputs. For example, coaching methods seem to evolve steadily as do innovations of track surfaces and shoe technology.

If our assumption of steady increases in non-drug related inputs and technology is reasonable, a technological change most likely to cause dramatic deviations from steady improvements in performance is the rapid introduction of performance enhancing drugs, especially anabolic steroids. The introduction of anabolic steroids could have a pronounced non-linear effect on performance because it represents both a new production input and a new technological innovation that can increase the efficiency of other production inputs (e.g., training and nutrition). Therefore, for the purpose of this study, we make the assumption that percentage rate changes in track and field performances in the absence of steroids should be relatively constant. That is, we expect a log-linear change in performance before performance enhancing drugs become widely used. If the widespread introduction of performance enhancing drugs caused spikes in performance, we would expect performance during this period to be consistently better than the performance predicted by a log-linear trend.

To check on the reasonableness of the assumption that there would be log-linear performance improvement in the absence of steroids, we did simple log-linear trend regressions of performance for each of six track and field events for the period of 1949 through 1960, a period of time before the widespread introduction of steroids. This model performed well with statistically significant coefficients and high R square statistics for all six events.

III. Data

We use annual performance data from 1949 through 2007 that is gathered by the International Association of Athletics Federations (2008). They provide top performance statistics for each year in all events and it is published in *Track and Field News* annually. The sample we selected is restricted to men only as women's data is not complete in all events for this time period. Six events were chosen for inclusion as they are representative of short, middle and long distance running events as well as the field events. The 100 meter and 400 meter runs are representative of short and long sprints respectively. The mile run is a middle distance race while the 5000 meter run represents distance races. The long jump and shot put were included in our sample to represent jumpers and throwers, respectively.

For each of these six events, we gathered statistics for the top ten performances for each year. The data was gathered manually from *Track and Field News* for the years 1949 through 2007. From these statistics it was possible to identify three annual performance measures for each event: the season best performance, the median of the top 10 performances and the average of the top 10 performances. Extensive analysis was done using all three of these measures. This paper reports analysis of the median of the top 10 performances. This measure has the advantage of representing the "typical" performance of the elite athletes in each event and is not influenced by "outlier" performances of the top athletes. It is important to note, however, that all three measures, when used as dependent variables, generated very similar patterns of results.

One data problem to be addressed is the use of fully automated timing or FAT. It was not introduced to the track and field world until 1975 (Francis and Coplon, 1990). Therefore, times before 1975 may be compromised, especially in the short sprints. Before 1975, manual timing involved a person starting and stopping the watch based on their perception of the gun being fired and the athlete crossing the finish line. With the old manual timing system, all one hundredths of a second beyond zero were rounded up to the nearest tenth of a second. In addition to errors from rounding, there were two potential human errors: reacting slow to the starting gun and anticipating the athlete crossing the finish line ahead of the actual crossing. Track and field statisticians have determined that for the sprints the combined manual timing errors from these three factors were recorded times that, on average, were 0.24 seconds too fast. FAT removed the human reaction time factor altogether since it involves a system that is linked to the starting gun. The clock automatically starts when the gun fires, and a camera captures the precise moment when the athlete crosses the finish line. Since experts have determined that measured performance is slower under FAT than manual timing, 0.24 seconds are added to pre-1975 manual times for events up to 1600m (Francis and Coplon, 1990). Thus, times in the 100 meter run and 400 meter run before 1975 were adjusted for this fact. The mile and 5000m were not adjusted, since 0.24 seconds is not considered significant enough in these events to require correction. Officials feel that this adjustment takes into account the documented delayed reaction in starting the watch and stopping early (Francis and Coplon, 1990, p. 162).

An important decision in this study was to determine the years of heavy-steroid use. Based on the institutional research discussed above, we define three post World-War II time periods:

- Period 1 is the pre-heavy steroid use period and is defined as 1949 through 1961.
- Period 2 is the period of systematic steroid-use period and lasted from 1962 through 1990. Anti-doping efforts during this period were not well coordinated.
- Period 3 is the aggressive anti-doping enforcement period from 1991 through 2007. During this period, testing became more effective and there were more aggressive and coordinated campaigns to stop the distribution and use of anabolic steroids and other performance enhancement drugs, especially to elite athletes.

The performance measure is designed to capture the typical performance of elite athletes. For each year we computed the median for the top ten performances in each event. Table 1 is a summary table that presents the mean of median annual performances for three time periods for each of the six events.

TABLE 1–Mean of Median Annual Performances for Three Time Periods by Event

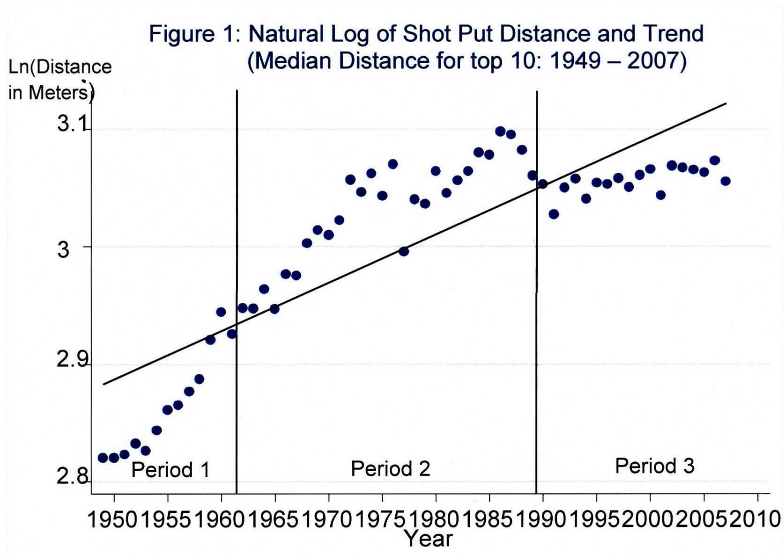
	Period 1: Pre (1949 - 1961)	Period 2: Heavy (1962 - 1990)	Period 3: Post (1991 - 2007)
100 Meter Run (seconds)	10.57	10.22	9.97
400 Meter Run (seconds)	46.62	45.20	44.48
Mile Run (minutes)	4.05	3.90	3.85
5K Run (minutes)	14.07	13.40	12.97
Long Jump (meters)	7.72	8.18	8.38
Shot Put (meters)	17.57	20.76	21.27

Table 1 shows that the performances improved in period 2 compared to period 1. For example, the performance measure for the 400 meter run improved from 46.62 seconds to 45.20 seconds, and the performance measure for the shot put improved from 17.57 meters to 20.76 meters. The degree of improvement between period 2 and period 3, on the other hand, appears more moderate.

Figure 1 shows a scatterplot of actual performance values (in natural logs) around a trend line for the shot put. The pattern is dramatic with actual values exceeding the trend for almost all years during period 2, the period of unrestricted drug use. Distances tend to be below trend during periods 1 and 3.

Scatterplots for the other five events have similar patterns, showing actual performance that is predominantly better than trend during period 2 and worse than trend during periods 1 and 3.

The next section develops an empirical framework to explore these trends in more detail. The main purpose of the empirical analysis is to determine if track and field performances were significantly above trend during the period of heavy use (period 2: 1962-1990) and whether performances fell below trend during the period of more effective testing and more coordinated anti-doping programs (period 3: 1991-2007). The analysis should also determine whether the performance of elite distance runners (5 kilometer) show less improvement relative to trend compared to sprinters, middle distance runners, shot putters and long jumpers.



IV. Empirical Framework

To address these questions we adopt a four stage procedure. First, we run the following log-linear trend regressions with robust standard errors for each event in order to establish trend over the 1949 through 2007 time period:

$$LN(PERFORMANCE_i) = \alpha + \beta_i(YEAR), \quad (1)$$

where $PERFORMANCE_i$ is the performance measure for event i and is computed as the median performance of the top ten individual performances in the year of observation. Separate log-linear trend regressions are run for each of the following six events: 100 meter dash, 400 meter run, mile run, 5000 meter run, long jump, and shot put. It is important to note here that our goal with this trend regression is not to use time series techniques to fit performance patterns, but instead, to observe deviations of performance from a log-linear trend. As argued above, systematic deviations from a log-linear trend could indicate widespread use of performance enhancing drugs during period 2. Since serial correlation is expected, we estimate the log-linear trend regressions using robust standard errors.

Second, we calculate residuals from each of the log-linear trend regressions. Each residual represents the extent that the actual performance deviates from trend. During period 2, we expect performances to have *negative* residuals for running events and *positive* residuals for field events as this indicates better than trend performances. We expect the opposite, *positive* residuals for running events and *negative* residuals for field events, during period 3 as performances should be worse than trend due to the improvements in testing and greater public awareness.

Third, we use the residuals to create a performance related dummy variable (BETTER) that indicates whether the actual performance was "better" than the performance predicted from the trend regressions. Note that BETTER is coded 1 for running events when actual performance is *less than* the predicted performance while BETTER is coded as 1 for the long jump and shot put when actual performance that is *greater than* the predicted performance. Since the indicator variable BETTER is in the same units (zero or one) we are able to pool this variable across the six events and over the 59 year time period. The pooled sample consists of 354 observations. Converting the six disparate performance measures into a zero-one indicator variable (BETTER) and pooling the data also has the advantage of allowing us to use simple panel methods to estimate the effect of the heavy steroid use period (PERIOD 2) on the probability of having a better than predicted performance. Pooling of performance data is not possible using the raw performance measures because those measures are in different units (e.g., feet for shot put vs. seconds for running events).

Fourth, to predict the probability that a performance will be greater than trend, we use a marginal effects probit model with robust standard errors to predict BETTER as a function of a set of dummy variables. Detailed definitions of these variables are given in Table 2. The following two models are estimated:

$$\text{BETTER} = \alpha + \beta_1(\text{PERIOD2}) + \beta_2(\text{PERIOD3}) + \beta_1(\text{OLYMPIC}) \qquad (2)$$

$$\text{BETTER} = \alpha + \delta_1(\text{100\_PERIOD2}) + \delta_2(\text{400\_PERIOD2}) + \delta_3(\text{MILE\_PERIOD2}) + \delta_4(\text{5K\_PERIOD2}) + \delta_5(\text{LJ\_PERIOD2}) + \delta_6(\text{SP\_PERIOD2}) + \delta_7(\text{PERIOD3}) + \delta_8(\text{OLYMPIC}) \qquad (3)$$

Equations 2 and 3 are simple panel models. They allow the pooling of data across the six events and include time dummies to estimate the effect of the three time periods on performance and event dummies to estimate effects of specific events on performance (Equation 3).

Equation (2) is the base model that regresses BETTER against dummies for two time periods. PERIOD 2 is the heavy steroid use period and PERIOD 3 is the post 1990 period of more coordinated testing. We also include OLYMPIC to control for the recurring seasonal effect of the Olympics on performance. Athletes often orient their training to achieve peak performance during Olympic years.

To determine the precise beginning and ending date for period 2, we ran 64 probit analyses using Model 2 and combinations of a set of potential beginning dates from 1961 through 1968 and a set of potential ending dates from 1989 through 1996. These ranges were chosen from anecdotal evidence from the literature on the history of performance enhancing drugs in track and field that suggested that relatively unregulated use began in the 1961 to 1968 period and that coordinated international anti-doping policy began in the early 1990s. The next step was to use an econometric procedure to identify the specific start date. Marginal effect probit models were run for each of the 64 combinations of potential beginning dates and ending dates to determine the "best" start date and end date. The goodness of fit measure to determine the best beginning and ending date for period 2 is the Wald Chi Square statistic from the STATA results. The period 2 break points that maximize the Wald test statistic are a start year of 1962 and an end year of 1990.

The research hypotheses that track and field performances of elite athletes should show significantly more improvement during the heavy steroid use period (1962 through 1990) than during the pre-heavy use period (1949-1961) and the doping-control period (1991 through 2007) can be restated in terms of the coefficients of PERIOD2 and PERIOD3, with the pre-heavy use period as the omitted category. In terms of equation 2, we hypothesize that $\beta_1 > 0$ and $\beta_1 > \beta_2$.

Equation (3) systematically replaces PERIOD2 with a set of interaction terms that multiply PERIOD2 times the six track and field event dummies. For example, 100_PERIOD2 is an interaction term that assumes the value of 1 if the event is the 100 meter run and takes place during period 2 which is the heavy use period. The reference period remains PERIOD 1, just as it was in the short model (Equation 2). In terms of equation 3, we hypothesize that the coefficients to all interaction terms (δ_1, δ_2, δ_3, δ_4, δ_5, δ_6) will be positive and statistically significant.

TABLE 2–Variable Definitions

Variable	Definitions
Dependent Variable	
BETTER	Dummy variable indicating that the event median performance of the top ten performances was better than predicted by the event specific trend regression.
Independent Variables	
YEAR	Measured as actual year and used in the trend regressions.
PERIOD1	Dummy variable indicating time frame 1949 through 1961. This period is the control period before the use of anabolic steroids in track and field became systemic.
PERIOD2	Dummy variable indicating time frame 1962 through 1990. During this period anabolic steroids were frequently used by track and field athletes and testing was not very effective. This period is referred to as the heavy use period in the paper.
PERIOD3	Dummy variable indicating time frame 1991 through 2007. Testing of elite track and field athletes for anabolic steroid use became more effective during this period. It is referred to as the post heavy use period in the paper.
OLYMPIC	Dummy variable indicating that the performance was during an Olympic year.
100_PERIOD2	Interaction dummy variable that assumes the value of 1 if the performance was the 100 meter run and the performance was during Period 2 (i.e., 1962 through 1990).
400_PERIOD2	Interaction dummy variable that assumes the value of 1 if the performance was the 400 meter run and the performance was during Period 2 (i.e., 1962 through 1990).
MILE_PERIOD2	Interaction dummy variable that assumes the value of 1 if the performance was the mile run and the performance was during Period 2 (i.e., 1962 through 1990).
5K_PERIOD2	Interaction dummy variable that assumes the value of 1 if the performance was the 5 kilometer run and the performance was during Period 2 (i.e., 1962 through 1990).
LJ_PERIOD2	Interaction dummy variable that assumes the value of 1 if the performance was the long jump and the performance was during Period 2 (i.e., 1962 through 1990).
SP_PERIOD2	Interaction dummy variable that assumes the value of 1 if the performance was the shot put and the performance was during Period 2 (i.e., 1962 through 1990).

We also hypothesize that elite distance runners will have more moderate performance gains during period 2 than elite athletes in the other events. In terms of equation 3, we expect that δ_4 will be of lower magnitude than δ_1, δ_2, δ_3, δ_5 and δ_6.

V. Results

A. TREND REGRESSIONS AND RESIDUALS

Results from the basic trend regression (Equation 1) are shown in Table 3. All regressions were run with robust standard errors because of the presence of autocorrelation. Since the performance measure is in natural logs, the log-linear estimates produce coefficients that are interpreted as "percent change" in performance per year. The trend coefficients are statistically significant for each event and range from negative 0.11 percent in the 400 meter run to 0.41 percent in the shot put. All of the signs are correct since negative coefficients in the running events indicate decreasing times and positive coefficients for the field events (long jump and shot put) indicate longer distances over time.

TABLE 3–Event Trend Regressions (1949-2007) Estimating Annual Percent Change in Natural Log of Event Performance (Robust Standard Errors)

Event	Coefficient	T	R^2	N
100 Meter	-0.13	-24.5	.92	59
400 Meter	-0.10	-16.3	.86	59
Mile Run	-0.11	-10.7	.77	59
5K Run	-0.17	-23.8	.93	59
Long Jump	0.17	13.5	.83	59
Shot Put	0.41	11.3	.72	59

Table 4 provides a summary of residuals for each of the three time periods by giving the proportion of "better than trend" performances for each event in each of the three time periods.

TABLE 4–Proportion Better than Trend Performance by Event for Three Time Periods

Event	Period 1 (1949-1961)	Period 2 (1962-1990)	Period 3 (1991-2007)
100 meter	.23	.79	.47
400 meter	.23	.86	.35
Mile Run	.38	.90	.12
5K Run	.38	.72	.18
Long Jump	.08	.97	.35
Shot Put	.08	.93	.00

For the running events, where the performance is measured by time, "better than trend" means a negative residual where actual performance is less than trend. For the two field events, where the performance is measured by distance, "better than trend" is a positive residual where actual performance is above trend. The results are striking. During the heavy use period (Period 2), the proportion of better than trend predictions range from 0.79 for the 5,000 kilometer run to 0.97 for the 400 meter run and the long jump. During Period 3, on the other hand, when testing became more effective and anti-doping enforcement efforts more coordinated, the proportion of "better than trend" performances fell sharply, ranging from 0.00 for the shot put to 0.47 for the 100 meter dash. These patterns suggest that the relatively widespread and uncontrolled use of performance-enhancing drugs during period 2 had a marked effect on performance. Additionally, this suggests that more effective testing and coordinated anti-doping activities since 1990 could have reduced the amount of drug use and caused performances to be consistently less than trend.

Table 5 shows the probit results from estimation of equations (2) and (3) above. The dependent variable (BETTER) is an indicator that assumes the value of one if the observed performance is better than the log-linear trend estimate. Because of the presence of serial correlation, the analysis was run with robust standard errors. The coefficients are reported as marginal effects and can be interpreted in probability terms.

Model 1 presents results for the simple model (Equation 1) that includes dummies for whether the performance took place during the heavy-use period (PERIOD2) or the effective anti-doping control period (PERIOD3). Since the omitted category is PERIOD1, the coefficients to PERIOD2 and PERIOD3 are interpreted relative to the first time period (1949 through 1961). The coefficient to PERIOD2 shows that the probability of having a better than trend performance increased by almost 62 percent during the second period relative to the first period. However, there is a reversal in period 3. The coefficient to PERIOD3 is negative and not statistically significant. This means that during period 3, the probability of a better than predicted performance is not significantly different than in period 1, the omitted time category. To summarize, Model 1 results are consistent with expectations. They indicate that elite track and field athletes were much more likely to have a better than trend performances during the heavy use period (PERIOD2) relative to periods 1 and 3.

The control variable OLYMPIC is a dummy variable that indicates whether the year was an Olympic year. The coefficient indicates that performances in our sample were about 21.7 percent more likely to be above trend performance during Olympic years.

Model 2 replaces the PERIOD2 dummy variable with a set of interaction terms. Each interaction term is derived by multiplying an event dummy with PERIOD2. The results are presented in the last two columns in Table 5 and are consistent with expectations. The coefficient for each interaction term is positive, indicating that each event had a higher percentage of better than trend performances during period 2 relative to the period 1. For example, the coefficient to PERIOD2_100 indicates that 100 meter dash athletes had about a 41 percent higher probability to have a better than trend performance during period 2 in reference to period 1. All of the coefficients to the interaction terms are positive and statistically significant.

As expected, the coefficient to the 5,000 meter run interaction term (PERIOD2_5k) is smaller than the other interaction terms. While this is consistent with our hypothesis that distance runner performances are less likely to be better than predicted during period 2 in comparisons to the other 5 events, the difference between the 5k coefficient and the others is not great. Thus, we do not find very strong evidence for our expectation that distance runners benefitted less than athletes in the other five events during the heavy steroid use period.

TABLE 5–Binary Probit Marginal Effects Model of BETTER[a]
for Pooled Sample (Robust Standard Errors)

Variable	Model 1 (without interactions)		Model 2 (with interactions)	
	dF/dx	Robust Standard Error	dF/dx	Robust Standard Error
OLYMPIC	0.217***	0.065	0.219***	0.065
PERIOD2	0.623***	0.053	-	-
PERIOD3	-0.033	0.080	-0.039	0.080
PERIOD2_100	-	-	0.406***	0.045
PERIOD2_400	-	-	0.439***	0.039
PERIOD2_Mile	-	-	0.452***	0.038
PERIOD2_5k	-	-	0.373***	0.052
PERIOD2_LJ	-	-	0.488***	0.034
PERIOD2_SP	-	-	0.469***	0.035
Observations	354		354	
Log pseudolikelihood	-163.77		-158.08	
Wald Chi^2	131.16		130.55	
$Prob>Chi^2$	0.000		0.000	
Pseudo R^2	0.33		0.35	

[a]Dependent is a dummy variable that indicates better than trend performance using median of top ten performances in the event
***indicates significance at $\alpha = .01$ level
**indicates significance at $\alpha = .05$ level

VI. Conclusions

The results of this study provide strong support for two of our three hypotheses. First, the hypothesis that performances of elite track and field athletes will be consistently better than long-term trend during the heavy anabolic steroid use period (1962-1990) receives strong support from the results. For example, Table 4 provides direct support for our expectation of above trend performance during the heavy steroid use period by showing that period 2 had a much higher proportion of better than trend performances than either period 1 or period 3. Additional support comes from estimating better than trend performance using marginal effects probit models (Table 5). In both models, the indicator variable for period 2 (PERIOD2) is a highly significant predictor of better than trend performance.

A conceptual difficulty of this study is whether change in non-drug related inputs could cause the observed patterns of above trend performances during period 2. One reason that we are confident in our conclusion that the introduction of anabolic steroids is the primary cause of these patterns is that the results are remarkably consistent across the six diverse events. We know from historical observations that steroids were taken by elite athletes in all of these events. However, changes in other performance related inputs are not likely to affect performances in all groups simultaneously. For example, while an improvement in shoe quality or track surface could benefit runners; it is unlikely to explain the large improvements in performance observed field events like the shot put where shoes and running surfaces are not as important. Thus, finding the same pattern of residuals over diverse events supports our claim that the introduction of anabolic steroids to athletes in all events is the major cause of the high percentage of better than trend period 2 performances.

Second, our results provide at best limited support for the hypothesis that 5,000 meter runners experience less performance improvement during period 2 than sprinters receives limited support. We expected this result because medical research shows that distance runners in events like the 5,000 meter run have a high percentage of slow twitch muscle fiber relative to athletes in the other five events and that slow twitch fiber responds less to anabolic steroids than fast twitch fiber. However, while the probit results (Table 5) do show 5,000 meter runners with fewer better than trend performances during period 2 than athletes in the other 5 events, the difference between events does not appear to be that great.

Finally, our results support the third hypothesis. Table 4 shows the proportion of better than trend performances falling markedly from period 2 to period 3. Also, in sharp contrast to the strongly positive coefficients for PERIOD2, the coefficients to PERIOD3 in the probit analyses are not statistically significant (Table 5). This suggests that the systematic efforts to control doping during the 1990's reduced the use of anabolic steroids among elite track and field athletes. We believe that this evidence indirectly supports the claim that these efforts are having an effect. Additionally, some elite track and field athletes are volunteering to be randomly drug tested in order to restore confidence in the sport. Hopefully, the efforts of anti-doping agencies and the increasing awareness of athletes will return track and field to a steroids-free sport.

References

Becker, Gary. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy.* 76(2): 169-217.

Dare, Bernie. 1979. *Running and Your Body: Applying Physiology to Track Training*. Los Altos, CA: Tafnews Press.

Dimeo, Paul. 2007. *A History of Drug Use in Sport: 1876-1976.* London: Routledge.

Eber, Nicolas. 2008. "The Performance-Enhancing Drug Game Reconsidered." *Journal of Sports Economics.* 9(3): 318-27.

Fort, Rodney. 2002. "The Value of Sports Talent." Pp. 177-219 in *Sports Economics*. New York, NY: Prentice Hall.

Francis, Charlie, and Jeff Coplon. 1990. *Speed Trap: Inside the Biggest Scandal in Olympic History*. New York, NY: St. Martin's Press.

Haugen, Kjetil. 2004. "The Performance-Enhancing Drug Game." *Journal of Sports Economics*, 5(1): 67-86.

Hoberman, John, and John Yesalis. "The History of Synthetic Testosterone." *Scientific American.* 272(2): 76-81.

International Association of Athletics Federations. 2008. *IAAF Athletics*. http://www.iaaf.org.

Liu, Yuanlong. 2004. "Track and Field Performance Data and Prediction Models: Promises and Fallacies," in Sergiy Butenko, Jaime Gil-Lafuente, and Panos Pardalos, eds., *Economics, Management and Optimization in Sports,* Berlin, Germany: Springer, 225-233.

Munasinghe, Lalith, Brendan O'Flaherty, and Stephan Danninger. 2001. "Globalization and theRate of Technological Progress: What Track and Field Records Show." *The Journal of Political Economy*. 109(5): 1132-49.

Pound, Richard. 2004. *Inside the Olympics: A Behind-the-Scenes Look at the Politics, the Scandals, and the Glory of the Games."* Canada: John Wiley & Sons Ltd.

Rosen, Daniel. 2008. *Dope: A History of Performance Enhancement in Sports from the Nineteenth Century to Today.* Westport, CT:

Rosen, Sherwin. 1981. "The Economics of Superstars." *American Economic Review.* 71(5): 845-58.

Schutz, R. and Yuanlong Liu. 1998. "Statistical Modeling in Track and Field," in Bennett, J. ed., *Statistics in Sport,* London: Arnold, 173-97.

"Superhuman Heroes." 1998. *The Economist* 347: 10-13.

Taylor, William. 2002. *Anabolic Steroids and the Athlete*. Jefferson, NC: McFarland and Company, Inc.

Yesalis Charles, ed. *Anabolic Steroids in Sport and Exercise*. Champaign, IL: Human Kinetics, 2000.

Yesalis, Charles, Stephen Courson and James Wright. "History of Anabolic Steroid Use in Sport and Exercise." *Anabolic Steroids in Sport and Exercise*. Ed. Charles Yesalis. Champaign, IL: Human Kinetics, 2000.