

fair play from foul

By Andrew Leigh

With every major sporting event – from World Cup soccer to the Olympic Games – officials must ready themselves for the inevitable claims of foul play. In each new tournament, the same questions keep coming back. How can we ensure that referees are scrupulously fair? Is there a way of stopping teams from “throwing” games when they are out of contention? And when can you hazard a guess that drugs are helping a team’s performance?

In sports from soccer to sumo wrestling, and even in the Olympics, these challenges loom large. For the most part, efforts to weed out corruption have been left to sporting officials, while drug testing has been handed over to the scientists. Successes have been mixed.

Now, a new approach has come to the starting blocks. From the nascent field of “sports economics”, a small group of academics are demonstrating that they can help discern foul play from fair. In a series of little-noticed papers, they are beginning to show that referee bias, player shenanigans, and drug use can be systemically highlighted.

Andrew Leigh is a PhD Student at the John E Kennedy School of Government, Harvard University. He thanks Justin Wolfers and Lisa Middlebrook for their assistance.

Blowing the whistle on the referees

To understand just how much of a difference small refereeing decisions can make, it helps to go back to one of the most famous soccer games of 1999 – a face-off between two of the world’s best teams, Manchester United and Bayern Munich, in the European Champions League Cup Final. At the end of the allocated 45 minutes for the second half, the English side trailed the Germans 0-1. All that remained was a few minutes of “injury time”.

Italian referee Perluigi Collina, adhering to soccer rules that require him to determine the duration of injury time, announced that the game would run for an additional three minutes. In one of the greatest upsets of modern soccer, Manchester United went on to score two goals, winning 2-1. Collina’s decision to extend the game by three minutes proved crucial. Had injury time been extended by only one or two minutes, the result might have been reversed.

Allegations of referee bias are common enough in soccer. There were probably some Bayern Munich fans convinced that Collina, one of soccer’s most respected referees, showed bias towards the English side. But in reality, it is almost impossible to be sure that a referee has acted improperly on any given day.

What economics can tell us, however, is whether there is systematic bias by referees. And it appears that this is exactly what has occurred. In another recently published paper, Luis Garicano (Chicago University), Ignacio Palacios (Brown University) and Canice Prendergast (Chicago University)¹ analysed the amount of injury time that referees added to soccer games, depending on whether the home team was ahead or behind. Because the data was readily available, they looked at games in the Spanish football league – but the same analysis could readily be performed on any other soccer tournament in the world.

The trio concluded that favouritism by referees did exist. When the home team was behind by one goal at the end of normal time, injury time was 35% higher. When they were ahead by one goal, injury time was 29% lower. In other words, referees adjusted injury time to boost the chances of a home team win. And the result was not a statistical anomaly – controlling for factors that might influence the intensity of the game (disciplinary actions by the referee, player substitutions, or the



Photo: AAP

relative rankings of the two teams), the evidence of favouritism remained.

The researchers also discovered several factors that affected the size of referee bias. First, home team favouritism increased when the stakes went up. When the Spanish Soccer Federation boosted the number of points per win from two to three, bias increased. Second, favouritism rose during the season. At the end of the season, it was 40% higher than at the beginning. Third, for all they might like to deny it, referees are affected by the roar of the crowd. There was a direct correlation between the number of home team supporters in the stadium and the likelihood that injury time would be adjusted to benefit their team.

But how much does such bias really matter? The three economists estimate that this form of bias probably changed the results in about 1 out of 50 games. But they stress that this is not the only form of home team bias. Referees may also boost the chances of a home team victory through penalties and fouls – we are simply unable to measure this form of favouritism. If Spanish soccer authorities want to ensure a fairer sport (and they may well be happy with the status quo), perhaps they ought to choose referees who are less susceptible to the baying of the locals.

A little help from my friends

In the sporting arena, favouritism need not come from referees. Sometimes one's opponent can help out – especially if more is at stake for you than for him. While this can happen in team sports, the most common incidents are in one-on-one matches. And according to the boffins, one of the sports that suffers most is sumo wrestling.

One of Japanese sumo's renowned bouts was the fight between champions Keisuke Itai and Yokozuna Akebono in 1991. Prior to the match, the older, more experienced Itai was firm favourite to win. Yet to the surprise of many fans, Itai hit the sand, and Akebono took the bout. Years later, Itai went public with a startling allegation. He claimed that the match had been rigged. Moreover, he told a packed press conference, it was one of just many bouts that he had "thrown" in his sumo career. Itai's accusations rocked the austere Japan Sumo Association, and helped to send the sport into a decline from which it is yet to recover.

We will probably never know whether the 1991 Itai-Akebono bout was rigged. But two prominent economists from the University of Chicago – Professors Mark Duggan and Steven Levitt – have recently looked into the issue of match-fixing in sumo wrestling, and returned with some disturbing findings.²

The temptation for wrestlers arises from a sharp discontinuity. Japanese wrestling tournaments are comprised of fifteen bouts. Wrestlers who win eight or more bouts benefit dramatically more than those who win seven or less (in terms of ranking and salary). Thus for those on the margin of winning eight bouts, the payoff is disproportionately large.

Analysing the bouts, the two economists find some intriguing results. To begin with, they point out that the proportion of wrestlers who end each tournament with eight wins should be the same as the proportion who end with seven wins. But in fact, substantially more wrestlers manage to end with eight wins. Moreover, wrestlers who are on the margin of eight wins on the final day of the tournament are around 25 percent more successful than one would expect.

But mightn't those who are on the margin of getting the critical eight wins simply put in more effort than their opponents? To test this question, Duggan and Levitt proceed to look at what happens when the same two wrestlers next meet. They find that a wrestler who won his eighth bout on the final day of the tournament is 10 percent more likely to lose when he next meets the same opponent. In other words, part of the payoff for throwing a match is doing the same for one's opponent the next time. As Itai told *Time* magazine last year, "If you didn't have enough wins, you would borrow a win from another wrestler at that tournament, and then you would have to pay them back at the next tournament."

Interestingly, public attention seemed to have an effect on match-fixing. When Duggan and Levitt looked at the three sumo tournaments that followed media coverage of allegedly rigged matches, they found that the win-loss ratio for wrestlers on the margin was far closer to what would ordinarily be expected. For those who would like to see sumo become a more honest sport, this is good news indeed.

Thanks to a pair of Chicago economists, we know that Itai was telling the truth when he said that match-rigging (or "yaocho" in Japanese) was widespread in sumo wrestling. Probably no great surprise to fans, who have come to accept the phenomenon. But perhaps the stonewalling Japan Sumo Association, which continues to refuse to acknowledge any similarities between sumo wrestling and the WWE, could benefit from heeding the lessons in Duggan and Levitt's research.

Since drug-testing first began at the 1968 Olympics, 57 athletes have returned positive tests, including 9 at the Sydney Olympics

Substantiating substance abuse

Finally to drugs, and the explosive question of drug use in the modern Olympics. Since drug-testing first began at the 1968 Olympics, 57 athletes have returned positive tests, including 9 at the Sydney Olympics. During the Cold War, East Germany was rumoured to have provided performance-enhancing drugs to many of its athletes. In the late-1990s, China's swimming team came under suspicion when many of its athletes suddenly seemed to improve. And in many other countries, competitors have been willing to put their health at risk to improve their chances of tasting Olympic gold.

With "blocking" drugs improving as rapidly as drug-testing, it seems the chances of cracking down on Olympic cheats are unlikely to improve anytime soon. But where economists may be able to supplement the work of biochemists is in working out where to focus scarce testing resources.

Stefan Szymanski (Imperial College, London)⁹, believes a fairly straightforward test can be applied. He begins with a model developed by two US economists to predict Olympic medal tallies, based on population, GDP per capita, whether the nation is hosting the games, and whether the country was formerly a planned economy or under Soviet influence. As he points out, the model was surprisingly effective, producing a 95% correlation between predicted and actual results at the Sydney Games. Indeed, the model correctly predicted that the US would win 97 medals.

But what Szymanski is interested in is not where the model works, but where it doesn't. In other words, which countries performed better than their economic and demographic strength would lead us to predict? Of course, some nations will always do better or worse in a

given year, but an anomalous performance in 2000 might at least raise some suspicion of foul play.

So who over-achieved in 2000? Top of the list is Russia, which won 29 more medals than predicted. Next is China, which despite withdrawing 27 of its swimmers, still beat its expected tally by 10, and the UK (also +10). Further down the list are Cuba (+9), Romania (+9), Australia (+6), Jamaica (+6), the Netherlands (+6), Belarus (+5) and Greece (+5).

Those who underachieved most were Canada (-9), Spain (-7) and Germany (-6). If underperformance is a sign of honesty, Szymanski declares, these countries should receive the gold, silver and bronze medals for their morals. On the other hand, it ought to come as no surprise that there have been plenty of doping allegations against Russia and China in recent years, and that a number of high-profile British athletes were accused of drug use around the time of the Sydney Games.

Economics and sport

So economists may have helped uncover some shenanigans on the field. But can they help reduce unfair play in the future? The answer, most surely, is yes. For sports officials and sports fans, there are three clear take-aways from the economists.

First, athletes respond to incentives. In most situations, sporting honour and the joy of victory will be the dominant incentive, keeping our beloved teams honest. But the incentive for corrupt conduct is greatest in situations in which one team cares much more about the result than another. Thus, "dead rubbers" – games played after the series has been decided – are ripe for corruption. Similarly, it is notable that players on the cusp of retirement have been at the

centre of the recent cricketing scandals. Another area in which the incentives are distorted is "points shaving".

Teams who deliberately win a game by less than the "bookies' start" – as US college basketball teams have sometimes been accused of doing – can potentially both win the game *and* make money by backing their opponents. Authorities may want to closely monitor attempts to manipulate points-start betting and betting on "dead" games.

Second, public attention can help reduce bias. Just as sumo tournaments held in the wake of increased media focus on match-rigging were fairer, soccer referees tended to be less biased when the number of home and away fans was more evenly matched. This seems to give credence to US Supreme Court Justice Louis Brandeis's nostrum that "sunlight is the best disinfectant". When suspicion is up, we may have good reason to hope that corruption is down.

Finally, analysing corruption at an aggregate level can help root out corruption at an individual level. Szymanski has turned the spotlight on Russia, China and the UK after the Sydney Games. If drug-testing resources are scarce, authorities may want to use the tools of economics when deciding which teams should be most carefully scrutinized. Likewise, it would be comforting to know that the bodies that govern Japanese sumo and Spanish soccer were responding to the evidence of misbehaviour in their respective sports.

Until now, the stars at uncovering foul play on the sports field have been courageous whistleblowers and investigative journalists. Yet if we truly want to stamp out cheating and corruption – not just in the World Cup, but in all sporting tournaments – it may be time to put a few economists on the starting bench. ¹⁰²