Forensic Sports Analytics: Can We Detect Match-Fixing in Tennis?

Oliver Hatfield¹, Jonathan Tawn¹, Chris Kirkbride¹, Tim Paulden², David Irons², Grace Stirling²

¹. STOR-i, Lancaster University. ². ATASS Sports.

What is Match-Fixing?
Match-fixing is a scourge on sport that involves pre-determining the outcome of matches for personal gain. It can be split into two main types:
- **Fixing to lose**: a player or team deliberately underperforms in a sporting contest, usually so that others have an edge on gambling markets.
- **Fixing to win**: a player or team illegally increasing their chance of winning. This can involve bribing officials or opponents, or other foul play.

We consider the first type and how it can be detected in betting markets.

Background - Betting Markets
One way of detecting match-fixing is to look for irregularities in betting markets. To do this, we need to understand how bookmakers’ odds move.

Consider a match between two players being represented by a see-saw. If a player’s end of the see-saw is high, then the odds available on that player winning are large, so that the winnings on a successful gamble are also large.

How Does Match-Fixing Affect This?
- In order to persuade a player to lose a match, substantial persuasion is needed - often in the form of bribes.
- The fixer will want to recoup this loss by gambling.
- If enough is gambled, this may affect the betting odds in a suspicious manner. But how much is suspicious?

The mere presence of large gambles in itself is not suspicious - sometimes, large irrational gambles are made anyway. We need to combine this with information from the matches themselves.

Literature
Most literature focusses on using bookmakers’ pre-match odds and sports models as two separate forecasts of match outcomes. However, only betting odds can be affected by match-fixing, so big differences can be suspicious.

**Issues with this approach:**
- Relies on good sports models.
- Betting odds are better forecasts than any sports models.

Additionally, using forecasts at a single point in time can be wasteful of information, since odds fluctuate over time. Focussing only on pre-match odds can also be restrictive - we wish to also consider in-play betting.

Anyone For Tennis?
Currently, we are considering match-fixing in tennis for several reasons.
- Individuals compete, not teams, making it easier to model.
- Most players are poorly paid, which heightens the risk of match-fixing.
- Suspicious betting patterns are seen comparatively commonly in tennis.

A Markov Model
In order to look at in-play match-fixing in tennis, we need to consider how to model a tennis match as it is played.

To win a match, a player needs to win 3 (or sometimes 5) sets, each of which involves winning 6 games, each of which requires winning at least 4 points.

Is iid a good assumption? It has been shown that assuming points are independent in this manner is not entirely accurate - some dependence exists, and server advantage can change at important points. Nevertheless, it has also been shown to be a reasonable approximation to make. (Klaassen, F. J. and Magnus, J. R. (2001), Journal of the American Statistical Association)

Next Steps
- Current methods rely on known serve parameters $p_1$ and $p_2$. This is impractical - want to make good pre-match priors and update in a Bayesian manner.
- If the market behaves rationally, match probabilities should behave very similarly to market data. Want to quantify what counts as suspicious.
- Combine market data with match data to look for suspicious activity.

Tennis Probabilities
Under this model, and known serve parameters $p_1$ and $p_2$, the probability a player wins the match can then be calculated by recursively conditioning from the end on the next point’s winner.