NOTTINGHAM TRENT UNIVERSITY

# DISCUSSION PAPERS IN ECONOMICS

No. 2019/3 ISSN 1478-9396

# HOW WELL DO ELO-BASED RATINGS PREDICT PROFESSIONAL TENNIS MATCHES?

LEIGHTON VAUGHAN-WILLIAMS

CHUNPING LIU

HANNAH GERRARD

JUNE 2019

**DISCUSSION PAPERS IN ECONOMICS**

At Nottingham Business School, our Working Papers Series in Economics include a wide variety of approaches reflecting our research interests and academic expertise.

This paper is part of the new series, ***Discussion Papers in Economics***.

Earlier papers can be found at:
https://www.ntu.ac.uk/research/research-at-ntu/academic-schools/research-at-nottingham-business-school/nbs-working-papers

Please contact the Editor (maternity cover) for enquiries concerning any of our Discussion Papers:

King Yoong Lim
Division of Economics
Nottingham Trent University
50 Shakespeare Street
Nottingham NG1 4FQ
UNITED KINGDOM

Email: king.lim@ntu.ac.uk
Tel: + 44 (0)115 848 6071

# How well do Elo-based ratings predict professional tennis matches?

Leighton Vaughan Williams[1], Chunping Liu[2], Hannah Gerrard[3]


## Abstract

This paper examines the performance of five different metrics for forecasting men's and women's professional tennis matches. We use data derived from every match played at the 2018 Wimbledon tennis championships, the only grass court Grand Slam tournament. The metrics we use are the betting odds, the official tennis rankings, the overall Elo ratings, the surface-specific Elo ratings (Elo based in this case only on matches played on grass), and a composite of some of the above. The Elo rating system is a method of ranking players based on their past matches, weighted by the ratings of the players they competed against. The performance indicators we use are prediction accuracy, calibration and model discrimination. For men's tennis we find that the betting odds outperform the other measures in terms of prediction accuracy and calibration. A weighted composite of overall and surface-specific Elo performs best in terms of model discrimination. For women's tennis, we find that a weighted composite of overall and surface-specific Elo performs best in terms of prediction accuracy, while a weighted composite of the betting odds, overall Elo and surface-specific Elo performs best in terms of calibration and model discrimination.


**Key words:** Forecasting, Elo, calibration, prediction accuracy, model discrimination, Wimbledon, tennis.


[1] Nottingham Business School
[2] Nottingham Business School
[3] Nottingham Business School

## 1. Introduction

Sport is bound up with forecasting. The bookmaking industry exists because of disagreements between forecasts. It is thus clear that forecasting is a central aspect of sport and of sports betting. The purpose of this paper is to examine the performance of different forecasting methodologies for both men's and women's professional tennis matches. The measures we use are the betting odds, the official men's (ATP) and women's (WTA) tennis rankings, the overall Elo ratings, the surface-specific Elo ratings (Elo based in this case only on matches played on grass), and a composite of some of the above. The Elo rating system is a method of ranking players based on their past matches, weighted by the ratings of the players they competed against. The performance indictors we use are prediction accuracy, calibration and model discrimination.

We focus on both men's and women's singles matches for the 2018 Wimbledon tennis championships, employing data derived from every match played at Wimbledon. Originating in 1877, it is the third of the four annual Grand Slam tournaments of the tennis season, and the only major tournament played on grass courts.

Both the men's and women's singles consist of 128 players, with direct entries based on the official ATP rankings of 104 males and the official WTA rankings of 108 female competitors. Another eight players of each gender are then chosen as 'wild card' entries, decided by the Committee of Management based on a player's previous performances during the season or by being a competitor of public interest to increase publicity for the event, with the remaining spots being filled by the winners of qualifying matches held in the week prior to the main competition (Wimbledon, 2019). The strongest 32 players of each gender are 'seeded' so that the best players do not play each other too early in the tournament. The rest of the players are then randomly assigned their matches, both against themselves and the seeded players. For female competitors, lower-ranked players can be put forward for seeding by the committee if it is considered that the official rankings do not correctly reflect the true current ability of the player. The players compete in a "single elimination tournament modus (knockout system)" (Leitner et al., 2009, p. 278).

## 2. Literature

Stekler et al. (2010) provide a review of sports forecasts – see also Vaughan Williams and Stekler (2010) – noting that many seek to evaluate the profitability of a forecasting method when used to place bets, rather than to evaluate forecasts per se. They also note that if we view betting odds as forecasts, then standard tests of forecast efficiency are also tests of information efficiency. Such studies have been common over the years – seminal papers include Snyder (1978), Asch et al. (1984) for horse race betting; Pope and Peel (1989) for football betting.

Many forecasting methods are evaluated according to whether they would achieve positive betting returns – early papers include Vergin and Scriabin (1978) for American football, Bolton and Chapman (1986) for horse racing, while much more recently Angelini and De Angelis (2019) assess betting market efficiency for eleven European football leagues.

Among statistical forecasting models, a common approach is to rank participants based on historical performance. Many sports run official ranking systems, and in addition Elo (1978) proposed a rating system for chess that has been used in a range of sports. Hvattum and Arntzen (2010) test Elo ratings against bookmakers and econometric models as a forecasting tool for English Premier League matches, finding that bookmakers outperform Elo ratings, but that Elo ratings are superior to econometric models, while Leitner et al. (2010) use Elo ratings among other methods when attempting to forecast outcomes from the 2008 European Championships football tournament. Ryall and Bedford (2010) create an Elo-based model for Australian Rules football, and Carbone et al. (2016) do so for rugby league. Kovalchik (2016) evaluates an Elo-based prediction system created by the website FiveThirtyEight.com (Silver and Fischer-Baum, 2015; Morris et al., 2016) and finds that this comes closer than other forecasting methodologies to beating bookmaker prices in tennis. Kovalchik and Reid (2019) extend this method for in-play tennis betting.

## 3. Data and Methodology

Table 1 summarises the source and sample size of the data including men's Association of Tennis Professionals (ATP) rankings and Women's Tennis Association (WTA) rankings, betting odds, and Elo ratings. The data used for each of the models is based on the 256 Wimbledon main draw entrants (128 men and 128 women). The construction of the data set is summarised in Table 1.

**Table 1: Summary of the data set**

| Data set | Source | Sample size |
|---|---|---|
| ATP Rankings | ATP World Tour, 2018 | Top 200 |
| WTA Rankings | WTA Tennis, 2018 | Top 300 |
| Betting ATP odds | Oddschecker, 2018a | 222 match odds |
| Betting WTA odds | Oddschecker, 2018b | 222 match odds |
| ATP Elo ratings | Tennis Abstract, 2018a | Top 169 |
| WTA Elo ratings | Tennis Abstract, 2018b | Top 173 |

Data was collected for the ATP and WTA rankings and for the Elo ratings at the start of the tournament and for the betting odds before the beginning of play on each day of the tournament. The ATP and WTA rankings were collected from the official websites, atpworldtour.com and wtatennis.com respectively, with the rankings of the top 300 women and 200 men shown in Table 1. The Elo and surface-specific Elo ratings were collected from www.tennisabstract.com. The data available was gathered for the top 173 rated WTA tour competitors and the top 169 ATP tour competitors. To find the best betting odds available, the betting comparison website, www.oddschecker.com was used as it collates all the data from a wide range of betting operators to give the most competitive odds.

## 3.1 ATP/WTA ranking

The Association of Tennis Professionals (ATP) and Women's Tennis Association (WTA) official world rankings are used within professional tennis to determine tournament eligibility. They both follow a 52-week cumulative rolling points system, with the results from the four Grand Slam tournaments having the highest points weighting. The weighting of the points decreases with the prestige of the tournament, as well as the round of the tournament reached. The points accrued from 19 ATP and 16 WTA tournaments out of all those played (weakest tournament scores drop out) are totalled to create the overall rankings of the players (Dingle et al. 2012).

## 3.2 Elo

The Elo rating system, originally developed by Arpad Elo (Elo, 1978) as a method of ranking chess players, takes the relative skill level of players based on their past performances to establish a prediction for a head-to-head outcome, and then updates the ratings after each match result.

The method works by allocating more points to a player when defeating a stronger opponent and deducting points when losing to a weaker opponent (Hvattum and Arntzen, 2010).

As a general rule, a 100-point difference is the equivalent of a 64% chance of winning, a 200-point difference equivalent to 75%, and 300-point difference to an 85% chance (Walkofmind, 2018) - see Equation (1).

$$p(A) = \frac{1}{1 + 10^{(R_B - R_A)/400}} \qquad (1)$$

where $R_A$ and $R_B$ are the ratings for player A and B. The Elo ratings differences were then converted to win probabilities for each player in a match. However, this formula does not work in a different rating structure, such as the ATP and WTA rankings.

Three types of Elo ratings were used within the methodology.

1. Standard Elo for ATP and for WTA.

2. Surface-specific Elo. Wimbledon is played on a grass court, so a surface-specific Elo only accounts for games played by the competitors on a grass surface. Other surfaces are clay and hard court.

3. An adjusted/combined Elo, which weights both standard Elo and surface-specific Elo. As Wimbledon is played on a grass court, the grass surface ratings are chosen to best reflect the player's abilities within this match scenario. We firstly construct an adjusted Elo rating to reflect both Elo and surface ratings, which is shown in Equation (2).

$$Adjusted\ Elo = (1 - \lambda) * Elo + \lambda * Surface \qquad (2)$$

The simplest adjustment is to weight each type of Elo equally, so taking the midpoint of the standard Elo and surface-specific Elo for each player (Adjusted Elo ratings 1). However, the equal weight of Elo and surface-specific Elo may not produce the optimal return. Considering this, we set $\lambda$ to be varying between 0 and 1. For each $\lambda$, we calculate the prediction accuracy, calibration and model discrimination. We choose the maximum value (best performance) of the three measures. The corresponding $\lambda$ is the optimal weight on surface-specific Elo. Instead of placing equal weights on Elo and surface Elo, we have calculated the adjusted Elo ratings (Adjusted Elo ratings 2), which uses the optimal weights.

As the forecasting performance of betting odds is another important indicator, we construct another rating in the Equation (3) incorporating the betting odds.

$$Adjusted\ Elo\ 3 = (1 - \lambda_1 - \lambda_2) * Elo + \lambda_1 * Surface + \lambda_2 * Betting\ Odds \qquad (3)$$

We set $\lambda_1$ and $\lambda_2$ to be varying between 0 and 1 but the sum of them cannot exceed 1. For each combination, we calculate the calibration and model discrimination. We choose the maximum value of the three measures.

The idea of developing a weighting-based or rule-based combination of methods to improve forecasting accuracy in sport has been previously explored by, for example, Spann and Skiera (2009).

## 3.3 Betting

To find the best odds available for the analysis, the odds comparison site, Oddschecker, was used as it collates all the data from a range of betting operators to give the best odds.

The odds were deflated by the over-round (the excess of the sum of the odds over 1) to give the implied probabilities for each player in a match.

Regarding the fractional odds, the method in which the implied probabilities were calculated is given in Equation (4), which follows Graham and Stott (2008). See also Clarke et al. (2017).

$$\frac{denominator}{denominator + numerator} * 100 \qquad (4)$$

# 4. Model performance

To test the performance of the models, three measures were used: prediction accuracy, calibration and model discrimination. When looking at the predictive power of a model, although accuracy may be viewed as the most desirable characteristic, the sensitivity to bias within the model is also important (Irons et al. 2014), hence the choice of these different measures.

Prediction accuracy is a measure of the number of correctly predicted matches that the player with the higher probability won. It is calculated by finding the number of matches that were correctly predicted divided by the total number of predictions and is expressed as a percentage.

$$Prediction\ accuracy = \frac{total\ number\ of\ correctly\ predicted\ matches}{total\ number\ of\ predictions} * 100 \qquad (5)$$

Calibration can be defined as how well the forecasted probabilities correspond to the actual outcomes (Tetlock and Gardner, 2015). In this paper, a calibration ratio is used, calculated as the sum of the probabilities of the higher ranked player winning divided by the number of matches the higher ranked player won.

$$Calibration = \frac{\text{sum of the probabilities of the higher ranked player wins}}{\text{total number of matches the higher ranked player won}} * 100 \quad (6)$$

The closer the ratio is to 1, the better calibrated and less biased the model is. If the model puts more weighting on the higher ranked players to win, the calibration will be more than 1, with a model underestimating the higher ranked players having a ratio less than 1.

Model discrimination is calculated as the mean probability of matches the higher-ranked player won minus the mean probability of when they lost (upsets).

$$\begin{aligned} &Model\ discrimination \\ &\quad = mean\ prediction\ for\ matches\ higher\ ranked\ player\ won \\ &\quad - mean\ prediction\ for\ matches\ they\ lost \end{aligned} \quad (7)$$

This is equal to the integrated discrimination improvement (IDI) measurement used by Pencina, D'Agostino and Vasan (2008). Higher values of the IDI and model discrimination reflect a higher discriminatory power, indicating that the probabilities are more certain for wins than upsets within the matches.

## 5. Results

Table 2 shows the forecasting performance of different rating methods. For men's tennis we find that the betting odds outperform the other metrics in terms of prediction accuracy and calibration. A simple weighted average of overall and surface-specific Elo performs best in terms of model discrimination. Looking at women's tennis, we find that betting odds perform the best in terms of prediction accuracy and calibration, while a simple weighted average method and surface Elo outperforms the others in terms of model discrimination.

As there are no probabilities associated with the ATP and WTA rankings, we are not able to calculate calibration and model discrimination.

**Table 2: Summary of prediction by method type**

| | **Rating methods** | **Prediction accuracy** | **Calibration** | **Model Discrimination** |
|---|---|---|---|---|
| ATP | Betting odds | 76.6% | 74.9% | 7.6% |
| | ATP Rankings | 64.9% | NA | NA |
| | Elo ratings | 70.3% | 72.1% | 6.0% |
| | Surface Elo ratings | 69.4% | 72.0% | 6.3% |
| | Adjusted Elo ratings 1 ($\lambda$=0.5) | 70.3% | 71.1% | 7.7% |
| WTA | Betting odds | 70.3% | 71.3% | 4.8% |
| | WTA Rankings | 63.1% | NA | NA |
| | Elo ratings | 68.5% | 71.3% | 4.0% |
| | Surface Elo ratings | 66.7% | 69.1% | 6.4% |
| | Adjusted Elo ratings 1 ($\lambda$=0.5) | 67.6% | 69.7% | 6.4% |

Setting the Elo and surface-specific Elo equally may not produce the best performance. We then set $\lambda$ to be varying between 0 and 1. For each $\lambda$, we calculate the prediction accuracy, calibration and model discrimination. We choose the maximum value (best performance) of the three measures. The corresponding $\lambda$ is the optimal weight. Table 3 summarises the prediction by this method. Based on this search, almost all the forecasting measures are improved compared with the Elo rating itself. The optimal weights are different if we choose to maximise different forecasting measures. For example, if we use prediction accuracy as our target, we should set 5.6% on Elo rating for ATP but 65.3% on Elo rating for WTA.

**Table 3: Summary of prediction by weighted Elo and Grass surface ratings**

| **Rating methods** | **Adjusted ATP Elo ratings 2** | **Adjusted WTA Elo ratings 2** |
|---|---|---|
| **Prediction accuracy** | 72.1% | 71.2% |
| Optimal weight on Elo | 5.6% | 65.3% |
| Optimal weight on surface | 94.4% | 34.7% |
| **Calibration** | 73.1% | 71.3% |
| Optimal weight on Elo | 74.0% | 100.0% |
| Optimal weight on surface | 26.0% | 0.0% |
| **Model discrimination** | 11.0 | 6.6% |
| Optimal weight on Elo | 40.5% | 3.7% |
| Optimal weight on surface | 59.5% | 96.3% |

As the role of betting odds is important in forecasting the performance, we construct another rating in the Equation (3) incorporating the betting odds. The corresponding optimal weights are shown in Table 4. For example, we should set the weight on Elo to be 57.7%, 12.9% on surface Elo and 29.4% on betting odds to achieve the highest model discrimination in men's tennis.

It should be noted that there is no prediction accuracy calculated, as the only way to construct this adjusted Elo is through the weighted average of probabilities of winning. We need to convert Elo, Elo surface and betting odds into probabilities first. Therefore, the adjusted Elo is a weighted average of winning probabilities. It is not possible to calculate prediction accuracy using these probabilities.

**Table 4: Summary of prediction by weighted Elo, Grass surface ratings and booker makers odds**

| **Rating methods** | **Adjusted ATP Elo ratings 3** | **Adjusted WTA Elo ratings 3** |
|---|---|---|
| **Calibration** | 74.7% | 71.7% |
| Optimal weight on Elo | 1.9% | 69.8% |
| Optimal weight on surface | 0.0% | 6.1% |
| Optimal weight on betting odds | 98.1% | 24.1% |
| **Model discrimination** | 10.9 | 8.0% |
| Optimal weight on Elo | 57.7% | 63.6% |
| Optimal weight on surface | 12.9% | 11.9% |
| Optimal weight on betting odds | 29.4% | 24.5% |

Table 5, 6 and 7 summarise methods with the best forecasting performance. For men's tennis, Betting odds are still the best in terms of prediction accuracy and calibration. Adjusted Elo are better in terms of model discrimination. For women's tennis, a weighted composite of the betting odds, overall Elo and surface-specific Elo performs best in terms of calibration and model discrimination.

**Table 5: Best performance in terms of prediction accuracy**

| ATP | Weights | WTA | Weights |
|---|---|---|---|
| Betting ATP odds | NA | Adjusted WTA Elo ratings 2 | 65.3% (Elo)<br>34.7% (surface) |

**Table 6: Best performance in terms of calibration**

| ATP | Weights | WTA | Weights |
|---|---|---|---|
| Betting ATP odds | NA | Adjusted WTA Elo ratings 3 | 69.8% (Elo)<br>6.1% (surface)<br>24.1% (betting odds) |
| Adjusted ATP Elo ratings 3 | 1.9% (Elo)<br>0.0% (surface)<br>98.1% (betting odds) | | |

**Table 7: Best performance in terms of model discrimination**

| ATP | Weights | WTA | Weights |
|---|---|---|---|
| Adjusted ATP Elo ratings 2 | 40.5% (Elo)<br>59.5% (surface) | Adjusted WTA Elo ratings 3 | 63.6% (Elo)<br>11.9% (surface)<br>24.5% (betting odds) |
| Adjusted ATP Elo ratings 3 | 57.7% (Elo)<br>12.9% (surface)<br>29.4% (betting odds) | | |

## 6. Conclusion

This paper seeks to compare and evaluate the performance of different metrics (official world rankings, Elo-based ratings and betting odds) against three indicators, i.e. prediction accuracy, calibration and model discrimination. For men's tennis we find that the betting odds outperform the other metrics in terms of prediction accuracy and calibration. A weighted composite of overall and surface-specific Elo performs best in terms of model discrimination. For women's tennis, we find that a weighted composite of overall and surface-specific Elo performs best in terms of prediction accuracy, while a weighted composite of the betting odds, overall Elo and surface-specific Elo performs best in terms of calibration and model discrimination. Consistently, therefore, we find that the official ranking system proved to be the worst performing measure, highlighting a case for a change in the method by which the official rankings are calculated (see also Reid et al., 2010).

The findings of this paper complement those of earlier studies, notably Kovalchik (2016), who studied the predictive ability of previously published tennis prediction models.

Kovalchik finds that no approach was able to match the predictive ability of the bookmaker, although the standard Elo was the closest competitor (the study did not include the combined Elo approach employed in this paper).

Overall, the findings of this study add to the case for a wider use of Elo-based approaches within sports forecasting, as well as within the player rankings methodologies.

Further work could extend the general approach applied to Elo-based forecasting to evaluate the performance of this approach against additional indicators, such as expected return as well as on different surfaces and at different tournament levels. Brier scores (Brier, 1950) could also be calculated. The Elo scores, and official rankings, could be updated in-running during the tournament, and other identified biases, such as the favourite-longshot bias, could be adjusted for (see Abinzano et al., 2016). Additional focus could also be applied to explaining the gender differences identified in the results (Paserman, 2007; Wozniak, 2012; Kovalchik and Ingram, 2018). Differently adjusted Elo-based ratings could also be used, such as employed by FiveThirty.com (Silver, 2018; Kovalchik and Reid, 2018).

Finally, an issue that was not addressed in this paper is the use of in-match updates of the pre-play expectations of match outcomes (Kovalchik and Reid, 2018).

## References


Abinzano, I., Muga, L. and Santamaria, R., 2016. Game, set and match: the favourite-longshot bias in tennis betting exchanges, *Applied Economics Letters*, 23(8), pp. 605-608.

Angelini, G. and De Angelis, L., 2019. Efficiency of online football betting markets. *International Journal of Forecasting*, 35 (2), pp. 712-721.

Asch, P., Malkiel, B.G. and Quandt, R.E., 1984. Market efficiency in racetrack betting, *Journal of Business*, 57 (2), pp. 165-175.

ATP World Tour, 2018. ATP Rankings. Available at: www.apttour.com/en/rankings/singles/

Bolton, R.N. and Chapman, R.G., 1986. Searching for positive returns at the track: A multinomial logit model for handicapping horse races. *Management Science*, 32 (8), pp. 1040-1060.

Brier, G.W., 1950. Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78 (1), pp. 1-3.

Carbone, J., Corke, T. and Moisiadis, F., 2016. The Rugby league Prediction Model: Using an Elo-based approach to predict the outcome of National Rugby League (Nrl) matches. *International Educational Scientific Research Journal*, 27 (9), pp. 26-30.

Clarke, S., Kovalchik, S. and Ingram, M., 2017. Adjusting bookmaker's odds to allow for overround. *American Journal of Sports Science*, 5 (6), p. 45.

Dingle, N., Knottenbelt, W. and Spanias, D., 2012, July. On the (page) ranking of professional tennis players. *European Workshop on Performance Engineering*, pp. 237-247. Springer, Berlin, Heidelberg.

Elo, A.E., 1978. *The rating of chessplayers, past and present*. Arco Pub.

Hvattum, L.M. and Arntzen, H., 2010. Using Elo ratings for match result prediction in association football. *International Journal of Forecasting*, 26 (3), pp. 460-470.

Irons, D.J., Buckley, S. and Paulden.T., 2014. Developing an improved tennis ranking system. *Journal of Quantitative Analysis in Sports*, 10 (2), pp. 109-118.

Kovalchik, S.A., 2016. Searching for the GOAT of tennis win prediction. *Journal of Quantitative Analysis in Sports*, 12 (3), pp. 127-138.

Kovalchiok, S. and Reid, M., 2018. A calibration method with dynamic updates for within-match forecasting of wins in tennis. *International Journal of Forecasting*, 35 (2), pp. 756-766.

Leitner, C., Zeileis, A. and Hornik, K., 2009. Is Federer stronger in a tournament without Nadal? An evaluation of odds and seedings for Wimbledon 2009. *Austrian Journal of Statistics*, 38 (4), pp. 277-286.

Leitner, C., Zeileis, A. and Hornik, K., 2010. Forecasting sports tournaments by ratings of (prob)abilities: A comparison for EURO 2008. *International Journal of Forecasting*, 26 (3), pp. 471-481.

Lisi, F., 2017. Tennis betting: can statistics beat bookmakers? *Electronic Journal of Applied Statistical Analysis*, 10 (3), pp. 790-808.

Morris, B., Bialik, C. and Boice, J., 2016. How We're Forecasting the 2016 U.S. Open. FiveThirtEight.com. August 28. Available at: https://fivethirtyeight.com/features/how-were-forecasting-the-2016-us-open/

Oddschecker, 2018a. Men's Wimbledon Betting Odds. Available at: www.oddschecker.com/tennis/wimbledon/mens/

Oddschecker, 2018b. Women's Wimbledon Betting Odds. Available at: www.oddschecker.com/tennis/wimbledon/womens/

Paserman, M.D., 2007. Gender differences in performance in competitive environments: evidence from professional tennis players.

Pencina, M.J., D'Agostino, R.B. and Vasan, R.S., 2008. Evaluating the added predictive ability of a new marker: From area under the Roc curve to reclassification and beyond. *Statistics in medicine*, 27 (2), pp. 157-172.

Pope, P.F. and Peel, D.A., 1989. Information, prices and efficiency in a fixed-odds betting market. *Economica*, pp. 323-34.

Reid, M., McMurtrie, D. and Crespo, M., 2010. The relationship between match statistics and top 100 ranking in professional men's tennis. *International Journal of Performance Analysis in Sport*, 10 (2), pp. 131-138.

Ryall, R. and Bedford, A., 2010. An optimized ratings-based model for forecasting Australian Rules football. *International Journal of Forecasting*, 26 (3), pp. 511-517.

Silver, N., 2018. How Our NFL Predictions Work. FiveThirtyEight.com Available at: https://fivethirtyeight/com/methodology/how-our-nfl-predictions-work/

Silver, N. and Fischer-Baum, R., 2015. How We Calculate NBA Elo Ratings. FiveThirtyEight.com. May 21. Available at: https://fivethirtyeight.com/features/how-we-calculate-nba-elo-ratings/

Snyder, W.W., 1978. Horse racing: Testing the efficient markets model. *The Journal of Finance*, 33 (4), pp. 1109-1118.

Spann, M. and Skiera, B., 2009. Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters. *Journal of Forecasting*, 28 (1), pp. 55-72.

Stekler, H.O., Sendor, D. and Verlander, R., 2010. Issues in Sports Forecasting. *International Journal of Forecasting*, 26 (3), pp. 606-621.

Tennis Abstract, 2018a. Current Elo ratings for the ATP tour. Available at: www.tennisabstract.com/reports/atp_elo_ratings.html

Tennis Abstract, 2018b. Current Elo ratings for the WTA tour. Available at: www.tennisabstract.com/reports/wta_elo_ratings.html

Vaughan Williams, L. and Stekler, H.O., 2010. Sports forecasting. *International Journal of Forecasting*, 26 (3), pp. 445-447.

Vergin, R.C. and Scriabin, M., 1978. Winning strategies for wagering on national football league games. Management Science, 24 (8), pp. 809-818.

Walkofmind, 2018. Elo rating vs. winning probabilities. Available at: https://www.walkofminf.com/prgramming/chess/elo.htm

Wimbledon, 2019. The Championships, Wimbledon 2018- Seeding. Available at: https://www.wimbledon.com/prgramming/chess/elo.htm

Wozniak, D., 2012. Gender differences in a market with relative performance feedback: Professional tennis players. *Journal of Economic Behavior and Organization*, 83 (1), pp. 158-171.

WTA Tennis, 2018. WTA Rankings. Available at: www.wtatennis.com/rankings

## DISCUSSION PAPERS IN ECONOMICS

2019/3 Leighton Vaughan-Williams, Chunping Liu, and Hannah Gerrard, How Well Do Elo-based Ratings Predict Professional Tennis Matches?

2019/2 Pengfei Jia and King Yoong Lim, Police Spending and Economic Stabilization in a Monetary Economy with Crime and Differential Human Capital.

2019/1 King Yoong Lim and Diego Morris, Modeling the Drugs and Guns Trade in a Two-Country Model with Endogenous Growth.

2018/3 Pengfei Jia and King Yoong Lim, Tax Policy and Toxic Housing Bubbles in China.

2018/2 Yousef Makhlouf, Trends in Income Inequality.

2018/1 Dimitrios Bakas and Athanasios Triantafyllou, The Impact of Uncertainty Shocks on the Volatility of Commodity Prices.

2017/9 Eghosa Igudia, Rob Ackrill and Simeon Coleman, *Entrepreneurial Responses to Austerity: The Role of the Informal Sector*.

2017/8 Rahmi Cetin and Robert Ackrill, *Openness and Growth in Challenging Times: Analysing the Trade-Growth Nexus in Slovakia*.

2017/7 Dimitrios Bakas and Yousef Makhlouf, *Can the Insider-Outsider Theory explain Unemployment Hysteresis in the OECD Countries?*

2017/6 Dimitrios Bakas, Georgios Chortareas and Georgios Magkonis, *Volatility and Growth: A not so Straightforward Relationship.*

2017/5 Will Rossiter, *Prospects and Challenges for City Region Devolution in Nottingham and the East Midlands.*

2017/4 Chunping Liu and Zhirong Ou, *What determines China's housing price dynamics? New evidence from a DSGE-VAR.*

2017/3 Morakinyo O. Adetutu and Thomas G. Weyman-Jones, *Fuel subsidies versus market power: is there a countervailing second best welfare optimum?*

2017/2 Robert Mullings, *Do Institutions moderate Globalization's effect on Growth?*

2017/1 Anthony J Glass, Karligash Kenjegalieva, Victor Ajayi, Morakinyou Adetutu, Robin C. Sickles, *Relative Winners and Losers from Efficiency Spillovers in Africa with Policy Implications for Regional Integration*.

2016/4 Sara Ornati, *International Environmental Agreements and Domestic Politics*.

2016/3 Yanhui Zhu, Jingwen Fan and Jon Tucker, *The impact of UK monetary policy on gold price dynamics*.

2016/2 John Ebireri and Alberto Paloni, *Bank Development and a Lower Degree of*

*Sophistication and Diversification of Developing Countries' Exports*.

2016/1 Robert Mullings and Aruneema Mahabir, *Growth by Destination: The Role of Trade in Africa's Recent Growth Episode*.

2015/1 Andrew Atherton, João R. Faria, Dongxu Wu and Zhongmin Wu, *Human Capital, Entrepreneurial Entry and Survival*.

2014/5 Leighton Vaughan Williams, *The US Supreme Court and the 'Affordable Care Act': An Exercise in Closed-Door Forecasting*.

2014/4 Juan Carlos Cuestas and Barry Harrison, *Unemployment Persistence in the EU15*.

2014/3 Francesco Marchionne and Sunny Parekh, *Growth, Debt, and Inequality*.

2014/2 Bo Jiang, Zhongmin Wu, Bruce Philp and Simeon Coleman, *Macro Stress Testing in the Banking System of China*.

2014/1 Eghosa Igudia, Rob Ackrill, Simeon Coleman and Carlyn Dobson, *Austerity Measures or Entrepreneurial Development? The case of the Nigerian Informal Economy*.

2013/5 Barry Harrison and Theodorus Wisnu Widjaja, *Did the Financial Crisis impact on the Capital Structure of Firms?*

2013/4 Geetha Ravishankar and Marie Stack, *The Gravity Model and Trade Efficiency: A Stochastic Frontier Analysis of Potential Trade*.

2013/3 Chunping Liu and Patrick Minford, *How Important is the Credit Channel? An Empirical Study of the US Banking Crisis.*

2013/2 Chunping Liu and Patrick Minford, *Comparing Behavioural and Rational Expectations for the US Post-War Economy.*

2013/1 Dan Wheatley, *Is it good to share? Debating patterns in availability and use of job share*.

2012/3 Simeon Coleman and Vitor Leone, *Time-Series Characteristics Of UK Commercial Property Returns: Testing For Multiple Changes In Persistence*.

2012/2 Otavio Ribeiro de Medeiros and Vitor Leone, *Multiple Changes in Persistence vs. Explosive Behaviour: The Dotcom Bubble*.

2012/1 Rob Ackrill and Simeon Coleman, *Inflation Dynamics In Central And Eastern European Countries*.

2011/4 Simeon Coleman, *Inflation Dynamics and Poverty Rates: Regional and Sectoral Evidence for Ghana.*

2011/3 Dan Wheatley and Zhongmin Wu, *Work, Inequality, And The Dual Career Household*.

2011/2 Simeon Coleman and Kavita Sirichand, *Fractional Integration and the Volatility Of UK Interest Rates*.

2011/1 Simeon Coleman, *Investigating Business Cycle Synchronization In West Africa*.

2010/11 Marie Stack and Eric Pentecost, *A Gravity Model Approach To Estimating Prospective Trade Gains in The EU Accession And Associated Countries*.

2010/10 Vitor Leone And Bruce Philp, *Surplus-Value And Aggregate Concentration In The UK Economy, 1987-2009*.

2010/9 Robert Ackrill and Adrian Kay, *WTO Regulations and Bioenergy Sustainability Certification – Synergies and Possible Conflicts.*

2010/8 Paul Alagidede, Simeon Coleman and Juan Carlos Cuestas, *Persistence Of Inflationary Shocks: Implications For West African Monetary Union Membership.*

2010/6 Bruce Philp and Dan Wheatley, *The time scarcity and the dual career household: competing perspectives*

2010/5 Juan Carlos Cuestas, Sebastián Freille and Patricio O'Gorman, *The media and public agendas: testing for media effects in Argentina Turing the Kirchner administration*

2010/4 Vitor Leone, *From property companies to real estate investment trusts: the impact of economic and property factors in the UK commercial property returns*

2010/3 Juan Carlos Cuestas and Paulo José Regis, *Purchasing power parity in OECD countries: nonlinear unit root tests revisited*

2010/2 Juan Carlos Cuestas and Bruce Philp, *Exploitation and the class struggle*

2010/1 Barry Harrison and Winston Moore, *Nonlinearities in Stock Returns for Some Recent Entrants to the EU*

2009/7 Joao R. Faria, Le Wang and Zhongmin Wu, *Debts on debts*

2009/6 Juan Carlos Cuestas and Luis A. Gil-Alana, *Unemployment hysteresis, structural changes,non-linearities and fractional integration in Central and Eastern Europe*

2009/5 Juan Carlos Cuestas and Javier Ordóñez, *Unemployment and common smooth transition trends in Central and Eastern European Countries*

2009/4 Stephen Dobson and Carlyn Ramlogan, *Is there a trade-off between income inequality and corruption? Evidence from Latin America*

2009/3 Juan Carlos Cuestas and Luís Alberiko Gil-Alana, *Further evidence on the*

*PPP analysis of the Australian dollar: non-linearities, structural changes and fractional integration*

2009/2 Estefanía Mourelle and Juan Carlos Cuestas, *Inflation persistence and asymmetries: Evidence for African countries*

2009/1 Juan Carlos Cuestas and Barry Harrison, *Further evidence on the real interest rate parity hypothesis in Central and Eastern European Countries: unit roots and nonlinearities*

2008/16 Simeon Coleman, *Inflation persistence in the Franc Zone: evidence from disaggregated prices*

2008/15 Juan Carlos Cuestas and Paulo Regis, *Nonlinearities and the order of integration of order prices*

2008/14 Peter Dawson and Stephen Dobson, *The influence of social pressure and nationality on individual decisions: evidence from the behaviour of referees*

2008/13 Juan Carlos Cuestas and Barry Harrison, *Testing for stationarity of inflation in Central and Eastern European Countries*

2008/12 Juan Carlos Cuestas and Dean Garratt, *Is real GDP per capita a stationary process? Smooth transitions, nonlinear trends and unit root testing*

2008/11 Antonio Rodriguez Andres and Carlyn Ramlogan-Dobson, *Corruption, privatisation and the distribution of income in Latin America*

2008/10 Stephen Dobson and Carlyn Ramlogan, *Is there an openness Kuznets curve? Evidence from Latin America*

2008/9 Stephen Dobson, John Goddard and Frank Stähler, *Effort levels in contests: an empirical application of the Tullock model*

2008/8 Juan Carlos Cuestas and Estefania Mourelle, *Nonlinearities in real exchange rate determination: Do African exchange rates follow a random walk?*

2008/7 Stephen Dobson and John Goddard, *Strategic behaviour and risk taking in football*

2008/6 Joao Ricardo Faria, Juan Carlos Cuestas and Estefania Mourelle, *Entrepreneurship and unemployment: A nonlinear bidirectional causality?*

2008/5 Dan Wheatley, Irene Hardill and Bruce Philp, *"Managing" reductions in working hours: A study of work-time and leisure preferences in the UK industry*

2008/4 Adrian Kay and Robert Ackrill, *Institutional change in the international governance of agriculture: a revised account*

2008/3 Juan Carlos Cuestas and Paulo José Regis, *Testing for PPP in Australia: Evidence from unit root test against nonlinear trend stationarity alternatives*

2008/2 João Ricardo Faria, Juan Carlos Cuestas and Luis Gil-Alana, *Unemployment and entrepreneurship: A Cyclical Relation*

2008/1 Zhongmin Wu, Mark Baimbridge and Yu Zhu, *Multiple Job Holding in the United Kingdom: Evidence from the British Household Panel Survey*

## DISCUSSION PAPERS IN POLITICAL ECONOMY

2006/3 Ioana Negru, *On Homogeneity and Pluralism within Economics Schools of Thought*

2006/2 David Harvie and Bruce Philp, *Learning and Assessment in a Reading Group Format or Reading Capital… For Marks*

2006/1 David Harvie, Bruce Philp and Gary Slater, *Regional Well-Being and 'Social Productivity' in Great Britain'*

2004/2 Massimo De Angelis and David Harvie, *Globalisation? No Question: Foreign Direct Investment and Labour Commanded*

2004/1 David Harvie, *Value-Production and Struggle in the Classroom, or, Educators Within, Against and Beyond Capital*

## DISCUSSION PAPERS IN APPLIED ECONOMICS AND POLICY

2007/2 Juan Carlos Cuestas, *Purchasing Power Parity in Central and Eastern European Countries: An Analysis of Unit Roots and Non-linearities*

2007/1 Juan Carlos Cuestas and Javier Ordóñez, *Testing for Price Convergence among Mercosur Countries*

2006/2 Rahmi Cetin and Robert Ackrill, *Foreign Investment and the Export of Foreign and Local Firms: An Analysis of Turkish Manufacturing*

2006/1 Robert Ackrill and Adrian Kay, *The EU Financial Perspective 2007-2013 and the Forces that Shaped the Final Agreement*

2004/5 Michael A. Smith, David Paton and Leighton Vaughan-Williams, *Costs, Biases and Betting markets: New evidence*

2004/4 Chris Forde and Gary Slater, *Agency Working in Britain: Character, Consequences and Regulation*

2004/3 Barry Harrison and David Paton*, Do 'Fat Tails' Matter in GARCH Estimation? Stock market efficiency in Romania and the Czech Republic*

2004/2 Dean Garratt and Rebecca Taylor, *Issue-based Teaching in Economics*

2004/1 Michael McCann, *Motives for Acquisitions in the UK*

2003/6 Chris Forde and Gary Slater, *The Nature and Experience of Agency Working in Britain*

2003/5 Eugen Mihaita, *Generating Hypothetical Rates of Return for the Romanian Fully Funded Pension Funds*

2003/4 Eugen Mihaita, *The Romanian Pension Reform*

2003/3 Joshy Easaw and Dean Garratt, *Impact of the UK General Election on Total Government Expenditure Cycles: Theory and Evidence*

2003/2 Dean Garratt, *Rates of Return to Owner-Occupation in the UK Housing Market*

2003/1 Barry Harrison and David Paton, *The Evolution of Stock Market Efficiency in a Transition Economy: Evidence from Romania.*