

Erasmus
School of
Economics

Sports Economics: Introduction

Dr. Sam Hoey
sportecon@ese.eur.nl

Erasmus University Rotterdam



Topics - Schedule

1. Introduction to the Economics of Sports
 - Tuesday 28 October 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Sam Hoey
2. Match prediction and Analytics
 - Thursday 30 October 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Sam Hoey
3. Betting Behaviour and Strategy
 - Tuesday 4 November 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Sam Hoey
4. Player Labor and Transfer Market
 - Thursday 6 November 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Thomas Peeters
5. Impact of Coaches
 - Tuesday 11 November 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Sam Hoey
6. Labor market discrimination
 - Thursday 13 November 2025, 13h00-15h00
 - Theil CB-5
 - Lecturer: Sam Hoey
7. Individual sport contests
 - Tuesday 18 November 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Sam Hoey
8. Club Revenues and Sharing
 - Tuesday 25 November 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Sam Hoey
9. Salary Caps and Financial Fair Play
 - Tuesday 2 December 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Thomas Peeters
10. Prediction Competition Results and Exam Questions
 - Tuesday 9 December 2025, 13h00-15h00
 - Van der Goot M1-17
 - Lecturer: Sam Hoey



Today and Thursday

Learning objectives:

1. Develop and tune a data-driven model to predict match outcomes, evaluate its performance using out-of-sample metrics (e.g., Brier score, accuracy), and compare it against relevant benchmarks.
2. Translate model predictions into a transparent betting strategy and assess its profitability and Sharpe ratio using out-of-sample testing.
3. Create and share a clean replication package (code + data + README) enabling third-party replication.

A stylized, handwritten-style logo of the word "Erasmus" in a dark blue or black color, located in the bottom right corner of the slide.

Today (and possibly Thursday)

Questions:

1. What forms of Sport Economics and Analytics are there? (*course overview*)
2. What methods can be used to predict match outcomes?
3. How can we tune (hyper) parameters of our model?

Practical issues:

- Instructor info
- Planning lectures and Data Clinics
- Assignments and exam(s)

→ I will use some mentimeter quizzes throughout the lecture(s).

A handwritten signature in black ink, appearing to read "Erasmus", located in the bottom right corner of the slide.

Readings

- Academic papers
 - Holmes, B., & McHale, I. G. (2024). Forecasting football match results using a player rating based model. *International Journal of Forecasting*, 40(1), 302-312.
 - Peeters, T. (2018). Testing the Wisdom of Crowds in the field: Transfermarkt valuations and international soccer results. *International Journal of Forecasting*, 34(1), 17-29.
 - Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of forecasting*, 26(3), 460-470.
 - Dixon, M. J., & Coles, S. G. (1997). Modelling association football scores and inefficiencies in the football betting market. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 46(2), 265-280.

Find them in the literature section in canvas!

A stylized, handwritten-style logo of the word "Erasmus" in a dark blue or black color.

Sport Economics? Sport Analytics?



Erasmus

Menti time

Join at menti.com | use code 5605 4055

 Mentimeter



SH 

What do you associate with sports economics and analytics?

fast bold
creative
inspiration leader focus
transpiration

Menti

New presentation



Choose a slide to present

What can we use to predict match outcomes?

One concept still fuzzy + one you can explain to a friend.



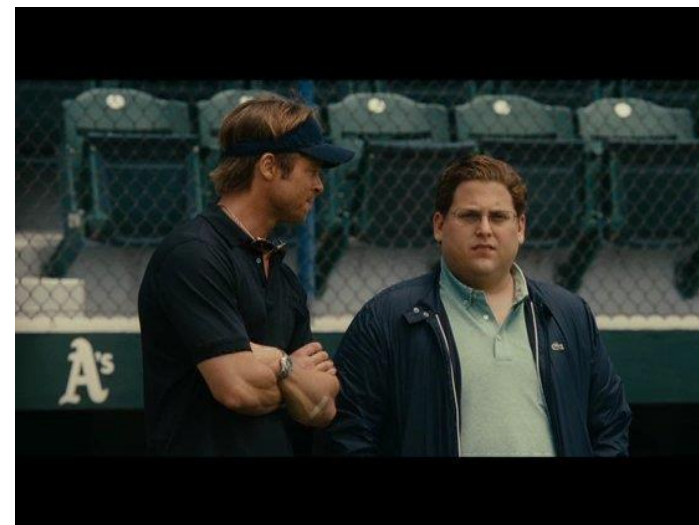
Erasmus

Sports Analytics

Our main interest:

Match prediction: Probabilistic models (Elo, past performance, likelihood models) to forecast win/draw/loss and simulate tables.

- **Management**
 - **Smarter decisions (Moneyball):**
Find undervalued skills, pick better lineups, and price contracts/options using statistics.
 - **Manage workload & prevent injuries:**
Use GPS tracking + inertial sensors (IMU) to measure movement and impacts, and session RPE (Rated Perceived Exertion) to record how hard players felt they worked. Combine these to flag risk and plan rest/rotations.
 - **Buy, sell, and value players:**
Blend scouting with models for fit and upside; value players with WAR (Wins Above Replacement), xG (expected goals) added and age curves.



Erasmus

Sports Analytics

On-field

- **In-game strategy (subs, 4th-down, PK)**
 - Live win-probability + fatigue models to time substitutions and tactic changes
 - 4th-down (American football): go-for-it vs. punt/field goal
 - PK (penalty kicks, soccer): shootout strategy, kicker/keeper matchups
- **Opponent scouting**
 - Map patterns & weaknesses: press triggers, set-piece tells, preferred zones
 - Tailor game plans and matchups accordingly



Erasmus

Research in sports

Two sub-streams of Sports Economics:

1. Economics of Sport → use economic methods to improve sport.
2. Use sports (data) to answer economic and managerial questions.

Benefits:

- High frequency data
- Clear rules, contracts and structure.
- Rich data

A stylized, handwritten-style logo of the word "Erasmus" in a dark blue or black color, located in the bottom right corner of the slide.

Economics of Sport

Economics of Sport (*Using economic theory to understand or improve sport itself*)

a) Contest Theory & Effort Allocation

- How incentives, rewards, and tournament structures affect player effort and performance.
 - **Example topics:**
 - Prize spread & effort: How prize distribution (top-heavy vs flat) changes effort and entry.
 - Tanking & draft incentives: When losing is optimal; effects of lottery reforms.

b) Industrial Organization of Sports Leagues

- Market structure, competitive balance, and profit vs. win maximization.
 - **Example topics:**
 - Salary caps and revenue sharing.
 - Open vs Closed leagues
 - Franchise fees

A stylized, handwritten-style logo for Erasmus, featuring a large, flowing 'E' followed by the word 'Erasmus' in a cursive script.

Economics of Sport

c) Labor Economics of Sport

- Labor market structure and wage determination in sports.
 - **Example topics:**
 - Free agency and player mobility.
 - Superstar effects and wage dispersion.
 - Discrimination in hiring and pay.

d) Sports Finance and Club Economics

- Financial management and sustainability of sports clubs.
 - **Example topics:**
 - Transfer market efficiency.
 - Financial Fair Play and regulation.
 - Profit vs Win maximization



Erasmus

Using Sports *as a Laboratory* for Economics.

“Ask not what economics can do for sports – Ask what sports can do for economics.” - Bar-Eli et al. (2020)

a) Labour economics

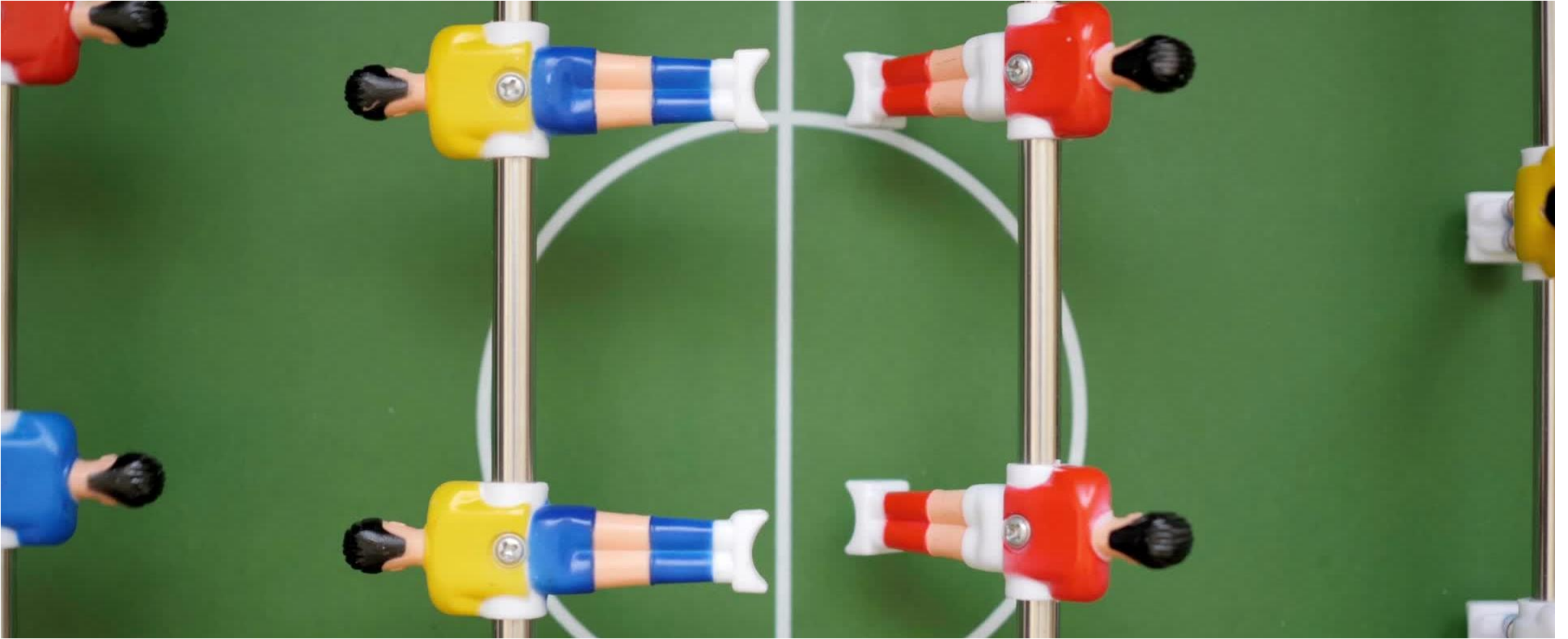
- Leverage data frequency and random shocks for research
 - Example topics:
 - Injuries as a source of random variation in team composition
 - Discrimination in sports, discrimination elsewhere?
 - Does climate change (heat/air quality) affect (cognitive) performance/productivity?

b) Behavioural economics

- Behavioural economics often done in lab → sports can provide real high-stakes lab environment
 - Example topics:
 - Choking under pressure (darts example)
 - Social pressure on referees (referee bias)

A stylized, handwritten-style logo of the word "Erasmus" in a dark grey or black color, located in the bottom right corner of the slide.

Sports Analytics – Match Prediction



Erasmus

Menti time

Join at menti.com | use code 5605 4055

What can we use to predict match outcomes?

fast bold
creative
inspiration leader focus
transpiration

Mentimeter



SH



Menti

New presentation



Choose a slide to present

What can we use to predict match outcomes?

One concept still fuzzy + one you can explain to a friend.



Ezra

What is there to predict?

Bookmakers generally provide odds for:

Match Outcome (1X2 Market)

- **1:** Home Win
- **X:** Draw
- **2:** Away Win

Proposition bets:

- Goals scored by each team
- Total goals (Over/Under)
- First goal scorer
- Correct score
- Half-time/full-time results
- Shots, corners, penalties, or player-specific events

These markets reflect expectations about team strength and can serve as a useful benchmark for model evaluation.

The screenshot shows the TOTO website interface. The top navigation bar includes links for SPORT, CASINO, LIVE CASINO, and BINGO. The main header features the TOTO logo and navigation links like SUPER WEEK, TOTO SCORE 6, TOTO EXTRA, KLANTENSERVICE, ACTIES, VERANTWOORD SPELEN, INLOGGEN, and AANMELDEN.

The main content area displays a live match page for the Netherlands Eredivisie, Feyenoord vs PSV. The score is 0-1, and the time is 46:32. The page shows various betting odds for match outcomes (1X2) and proposition bets (TOTO SPECIALS LIVE).

Match Outcome (1X2 Market) Odds:

Team	Score	Odds
Feyenoord	1	5,50
Draw	X	3,70
PSV	2	1,64

TOTO SPECIALS LIVE Odds:

Proposition	Odds
Ueda 1 of meer schoten op doel tussen minuut 45:00 en 54:59	5,25
Ueda 1 of meer schoten op doel tussen minuut 55:00 en 64:59	5,25
Hadj Moussa 1 of meer schoten op doel tussen minuut 45:00 en 54:59	7,00
Hadj Moussa 1 of meer schoten op doel tussen minuut 55:00 en 64:59	7,00
Sauer 1 of meer schoten op doel tussen minuut 45:00 en 54:59	9,50

The right sidebar shows a live match feed with a video player and a list of other matches to watch, including AZ vs FC Utrecht and Go Ahead Eagles vs Excelsior.

Erasmus

Match prediction methods

Generally, we use econometric/statistical methods to predict match outcomes

Three Key Stages of Match Prediction Modelling:

1. Training Stage

Choose model and estimate model parameters using historical data.

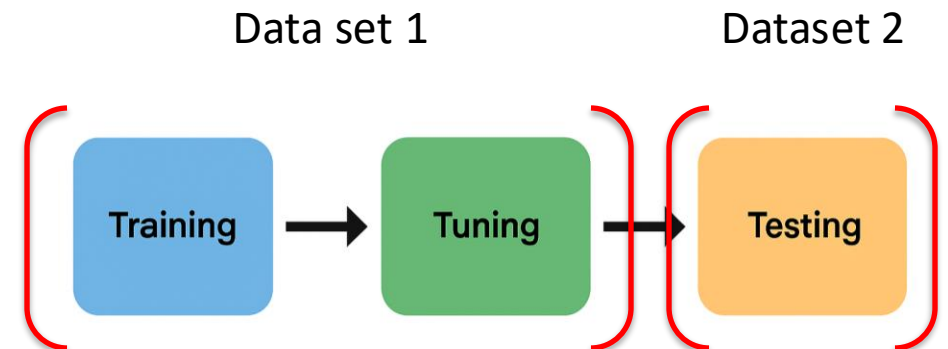
2. Tuning Stage

Adjust **hyperparameters** (e.g., window size, time decay rate, other features) to optimize predictive performance.

3. Testing Stage (next lecture)

Evaluate model performance on unseen data (test sample) to ensure robustness and avoid overfitting.

→ **Generally**, data is split into a training & tuning subset and a testing subset. **Often**, training and tuning is also split for model tuning.



Training Stage

Goal:

Build the *initial* model by estimating its parameters using **historical data**.

1. Prepare data

- Clean and structure match data (remove errors, ensure time order).
- Compute relevant features (e.g., team strength, form, ELO ratings)

2. Estimate model parameters

- Fit the chosen econometric/statistical model (e.g., OLS, ordered probit, logistic, Poisson).
- Learn relationships between covariates and match outcomes.

3. Assess in-sample fit

- Evaluate model fit (BIC, AIC, likelihood, pseudo- R^2 , residuals).
- Identify over-fitting patterns.

4. Save model outputs

- Store estimated parameters, fitted values, and diagnostics.
- Use these as inputs for tuning and validation.

Leakage: what NOT to do

- Use **end-of-season** stats to predict mid-season games.
- Build features with **future matches** (e.g., last-5 form that includes the target game).

A stylized, handwritten-style logo for Erasmus, featuring a large, flowing 'E' followed by the word 'Erasmus' in a cursive script.

Menti time

Join at menti.com | use code 5605 4055

 Mentimeter

0

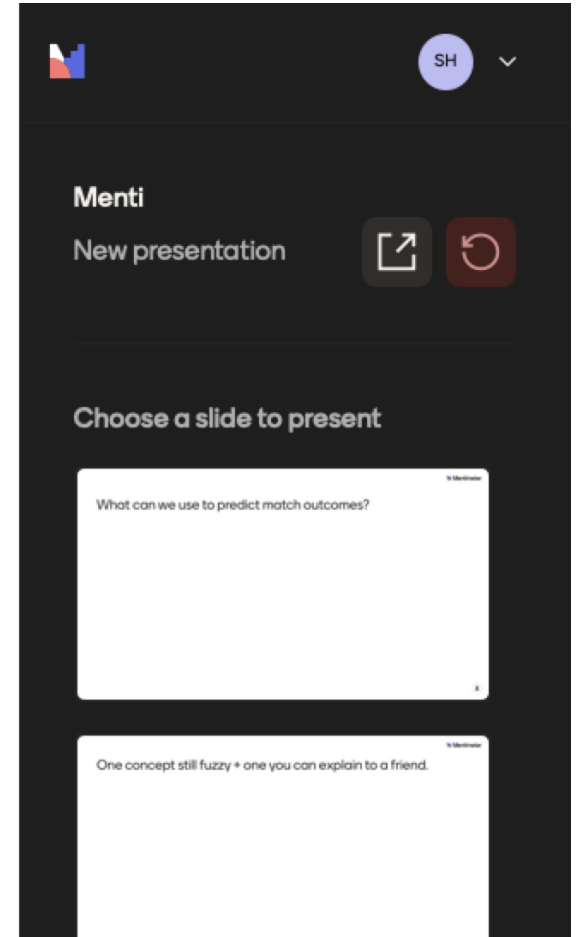
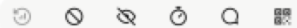
Rolling averages using performance data up until $t-1$.

0

Total goals scored and assists given at the end of current season.

0

Player salaries as stated by active contract.



Erasmus

Dixon & Coles (1997)

What they do:

1. Model the **number of goals** each team scores and concedes (poisson distribution).
2. In doing so: estimate team **attack** and **defence** strength parameters.
3. Convert goal distribution model under (1) into chances of **home win / draw / away win**.

In case of home win:

$$P_{home} = P(goals_{home} > goals_{away})$$

TABLE 4

Maximum likelihood estimates and standard errors for the attack and defence rate parameters, on August 5th, 1995, for Premiership teams (in the 1995–96 season)

<i>Team</i>	$\hat{\alpha}$	$se(\hat{\alpha})$	$\hat{\beta}$	$se(\hat{\beta})$
Arsenal	1.235	0.151	0.527	0.078
Aston Villa	1.278	0.178	0.649	0.086
Blackburn	1.730	0.209	0.534	0.082
Bolton	1.183	0.141	0.760	0.100
Chelsea	1.238	0.169	0.658	0.089
Coventry	1.115	0.164	0.699	0.094
Everton	1.177	0.169	0.667	0.091
Leeds	1.510	0.186	0.583	0.088
Liverpool	1.448	0.180	0.561	0.082
Manchester City	1.232	0.170	0.728	0.091
Manchester United	1.869	0.208	0.402	0.067
Middlesbrough	1.244	0.152	0.750	0.109
Newcastle	1.659	0.195	0.578	0.081
Nottingham Forest	1.460	0.170	0.658	0.095
Queen's Park Rangers	1.497	0.195	0.717	0.095
Sheffield Wednesday	1.387	0.179	0.698	0.091
Southampton	1.446	0.183	0.772	0.098
Tottenham	1.622	0.201	0.775	0.100
West Ham	1.192	0.169	0.649	0.087
Wimbledon	1.281	0.174	0.732	0.094

Erasmus

Dixon & Coles (1997)

What's plotted: time paths of the **maximum-likelihood parameter estimates** under **exponentially weighted likelihood** (recent games count more).

(a) Attack parameters (3 clubs):

Higher line \Rightarrow **stronger attacking rate** (team expected to score more).

(b) Defence parameters (3 clubs):

Higher line \Rightarrow **weaker defence** (opponents' expected scoring rate is higher).

Downward moves indicate tightening defence; upward moves indicate "leakier" periods.

(c) Common home-advantage parameter:

Stays **fairly stable** over time \rightarrow home effect is roughly constant in this sample.

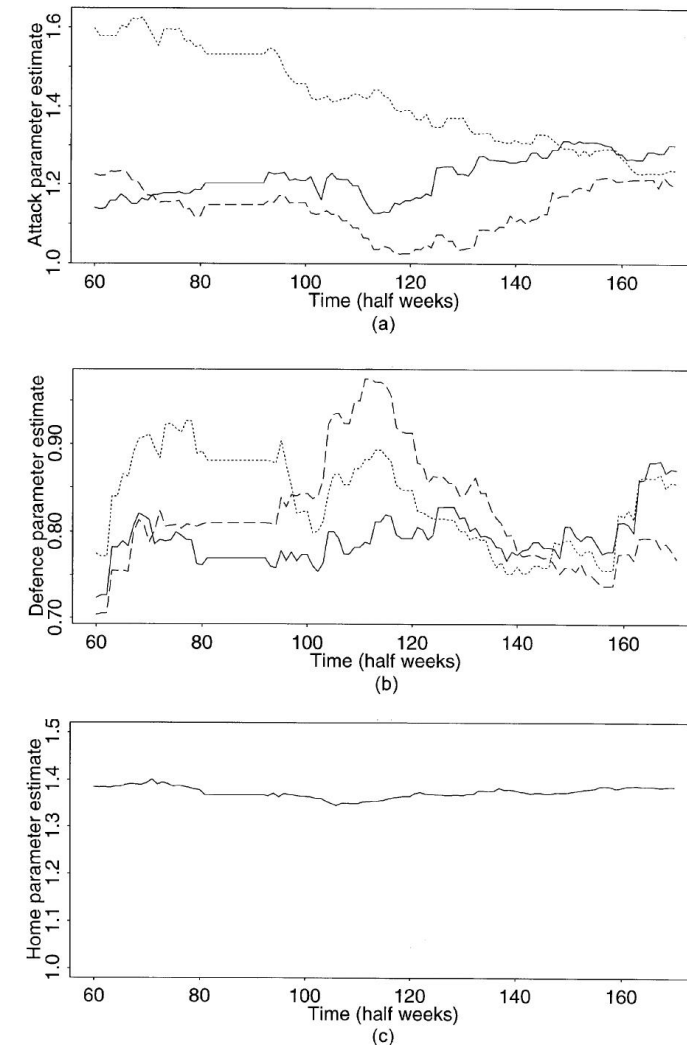


Fig. 2. (a), (b) Time series of the maximum likelihood estimates of attack and defence rate parameters for Sheffield United (—), Norwich (.....) and Everton (---); (c) variation of the common home effect parameter with time

Dixon & Coles (1997)

The results for *home win, draw, away win*.

TABLE 6

Maximum likelihood estimates for match outcome probabilities[†]

<i>Match</i>	<i>Maximum likelihood estimates for the following outcomes:</i>		
	<i>Home win</i>	<i>Draw</i>	<i>Away win</i>
Arsenal <i>versus</i> Middlesbrough	0.535 (0.069)	0.280 (0.030)	0.184 (0.046)
Aston Villa <i>versus</i> Manchester United	0.214 (0.054)	0.291 (0.029)	0.495 (0.072)
Blackburn <i>versus</i> Queen's Park Rangers	0.615 (0.078)	0.221 (0.033)	0.164 (0.049)
Chelsea <i>versus</i> Everton	0.457 (0.075)	0.298 (0.030)	0.245 (0.057)
Liverpool <i>versus</i> Sheffield Wednesday	0.535 (0.076)	0.262 (0.031)	0.203 (0.052)
Blackpool <i>versus</i> Wrexham	0.428 (0.077)	0.240 (0.018)	0.332 (0.070)
Stockport <i>versus</i> Burnley	0.480 (0.077)	0.259 (0.024)	0.261 (0.062)
Newcastle <i>versus</i> Stoke	0.705 (0.073)	0.198 (0.042)	0.097 (0.034)

[†]The matches are a subset of the fixtures from August 19th, 1995, plus one other across-division match. Approximate standard errors are calculated using the delta method. Standard errors are given in parentheses.

Note the standard errors reflect precision of underlying parameters used for prediction

A stylized, handwritten-style logo for Erasmus, featuring a large, flowing 'E' followed by the word 'Erasmus' in a cursive script.

Arntzen & Hvattum (2010)

What they do:

Keep a running **ELO rating** for each team and update it after every match. Then use a simple **ordered logit model** to turn the **rating gap** (home – away) into **home/draw/away** probabilities.

What Elo is (goal):

- Keep a running **team strength rating** that updates after every match based on match outcome or goal difference.
- Later, use **rating differences** to predict win/draw/loss.

Before the match — expected score between 0 and 1 (home team):

$$\gamma^H = \frac{1}{1 + c^{(\ell_0^A - \ell_0^H)/d}} \quad \text{and}$$

$$\gamma^A = 1 - \gamma^H = \frac{1}{1 + c^{(\ell_0^H - \ell_0^A)/d}}$$

Where ℓ_0^H and ℓ_0^A are initial elo ratings.

γ^H = expected home “score” (win=1, draw=0.5, loss=0)

γ^A = expected away “score”

Paper uses $c = 10$ and $d = 400$.

After the match — update rule:

$$\alpha^H = \begin{cases} 1 & \text{if the home team won,} \\ 0.5 & \text{if the match was drawn, or} \\ 0 & \text{otherwise.} \end{cases}$$

$$\longrightarrow \ell_1^H = \ell_0^H + k(\alpha^H - \gamma^H).$$

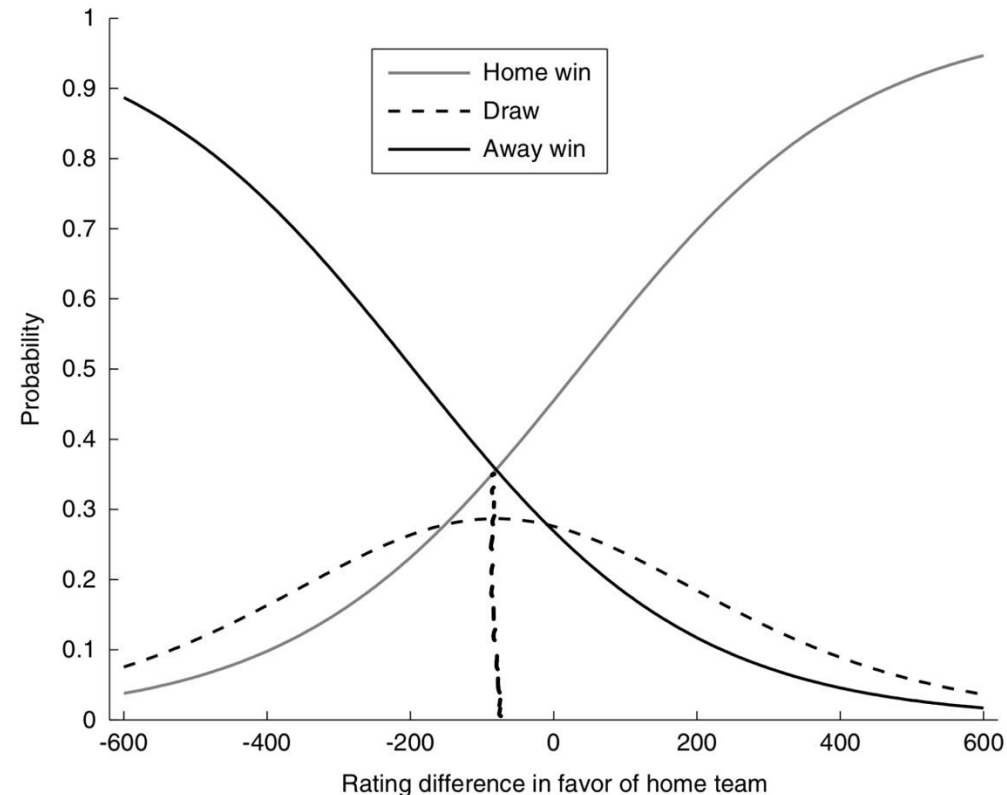
More on k later.

For home team above, do same for away team.

Arntzen & Hvattum (2010)

Next: use difference in elo ratings $x = l_0^H - l_0^A$ in an ordered logit model predicting away win, draw and home win (in that order) and derive $P(\text{home win})$, $P(\text{draw})$, $P(\text{away win})$

Fig. 2. Rating development for four selected teams in English association football.



Arntzen

Peeters (2018)

Framed more as an economics paper, how does wisdom of the crowds (woc) do in match prediction?

What they do:

Use the average Transfermarkt market value (woc) of each national team's pre-match squad (plus home and squad size) as predictors of match outcomes. Higher-valued squads should win more.

Output:

A simple ordered link turns the value difference into **home/draw/away** probabilities; performance is compared to **bookmakers, ELO, and FIFA**.

- **Takeaway:**

A public, crowd-sourced signal carries real predictive information; recency comes from **re-estimation and refreshed values**, not an explicit decay term.



Peeters (2018)

Best performing model
based on AIC and BIC

Training:

Table 8

Estimation results: regression model with wishful thinking controls.

Game result	Game result model				
Average value	0.450*** (0.024)	0.447*** (0.036)	0.422*** (0.026)	0.433*** (0.031)	0.479*** (0.047)
Number of fans		0.003 (0.023)			
Rel. number of fans			10.012*** (3.765)		
Players support teams				0.044 (0.049)	
Support weighted value					-0.013 (0.018)
Number of players	1.775** (0.880)	1.778** (0.880)	2.076** (0.898)	1.789** (0.879)	1.735** (0.881)
Home advantage	0.281*** (0.042)	0.281*** (0.042)	0.280*** (0.042)	0.281*** (0.042)	0.281*** (0.042)
Model	Ordered probit				
Observations	1,020	1,020	1,020	1,020	1,020
(Pseudo-)R ²	0.247	0.247	0.250	0.247	0.247
AIC	1648.4	1650.3	1642.9	1649.6	1649.9
BIC	1668.1	1675.0	1667.5	1674.2	1674.5

Notes: Standard errors are given in parentheses. Significance levels are denoted by: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Within sample
goodness-of-fit
measures.

Erasmus

Holmes and McHale (2024)

By far the most sophisticated model.

What they do:

Build **player ratings** (with positional/interaction effects) from event data, then use the **announced lineup** to construct a **team strength difference(s) for this match**. Convert those **differences** to **home/draw/away** probabilities with a multinomial model.

Table 2

Top ten Adjusted WhoScored (AWS) ratings achieved by players. There is no minimum number of games required for a player to appear in the table.

Player	Date	Team	League	AWS
Lionel Messi	19/05/2021	Barcelona	Spain LaLiga	1.799
Neymar	20/02/2018	PSG	France Ligue 1	1.603
Cristiano Ronaldo	07/08/2015	Real Madrid	Spain LaLiga	1.397
Robert Lewandowski	23/05/2021	Bayern	Germany Bundesliga	1.225
Kylian Mbappé	01/11/2020	PSG	France Ligue 1	1.166
Kevin De Bruyne	20/01/2021	Man City	England Premier League	1.106
Eden Hazard	10/05/2019	Chelsea	England Premier League	1.100
Zlatan Ibrahimovic	20/08/2016	Man Utd	England Premier League	1.091
Hakim Ziyech	11/03/2019	Ajax	Netherlands Eredivisie	1.075
Harry Kane	21/01/2018	Tottenham	England Premier League	1.053

- **Takeaway:**

Explains *why* a team is strong *today* (based on which players are on the pitch), not just *that* it's strong; recency is controlled by a decay function.



Team formation: 442
Opponent formation: 442
Player position: LST

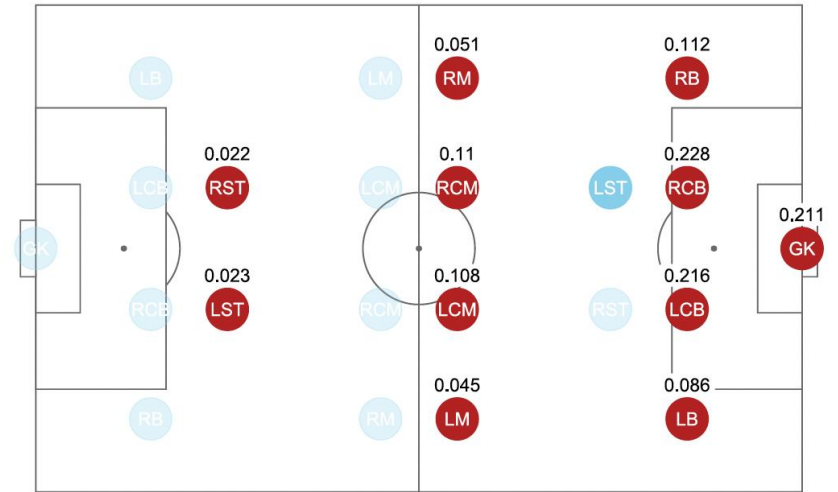
Holmes and McHale (2024)

Accounting for interaction effects with opposing team.

Team strength difference (Δ) is the main regressor in match outcome modelling:

$$\Delta = \sum_i \sum_j p_{ij} (AWS_i - AWS_j)$$

- Where p_{ij} represents the level of interaction between player i and opponents j . Weights sum to 1.
- AWS represents player rating of i and j
- Sum all weighted differences to get team strength difference.
- Weights are based on event-based interaction rates between players.



Team formation: 433
Opponent formation: 442
Player position: LW

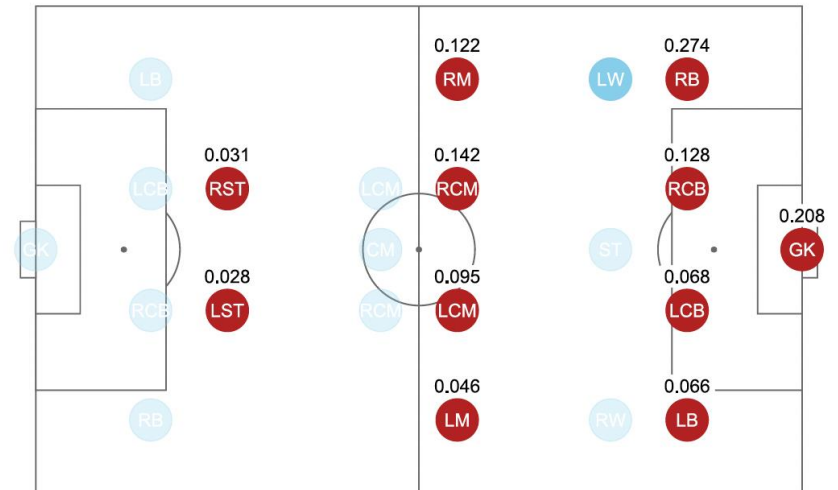


Fig. 3. Examples of the player weights for two different scenarios. The defending team is coloured in red.

Summary of papers

Core pipeline (all four papers):

- Build a **team strength rating** (goals model / ELO / crowd market values / player-based).
- Form match features (e.g., **rating gap** = Home – Away, plus **home** dummy).
- Convert features → **P(Home, Draw, Away)** with a **multinomial model** such as **ordered probit/logit**.

Methods used for building team strength variable:

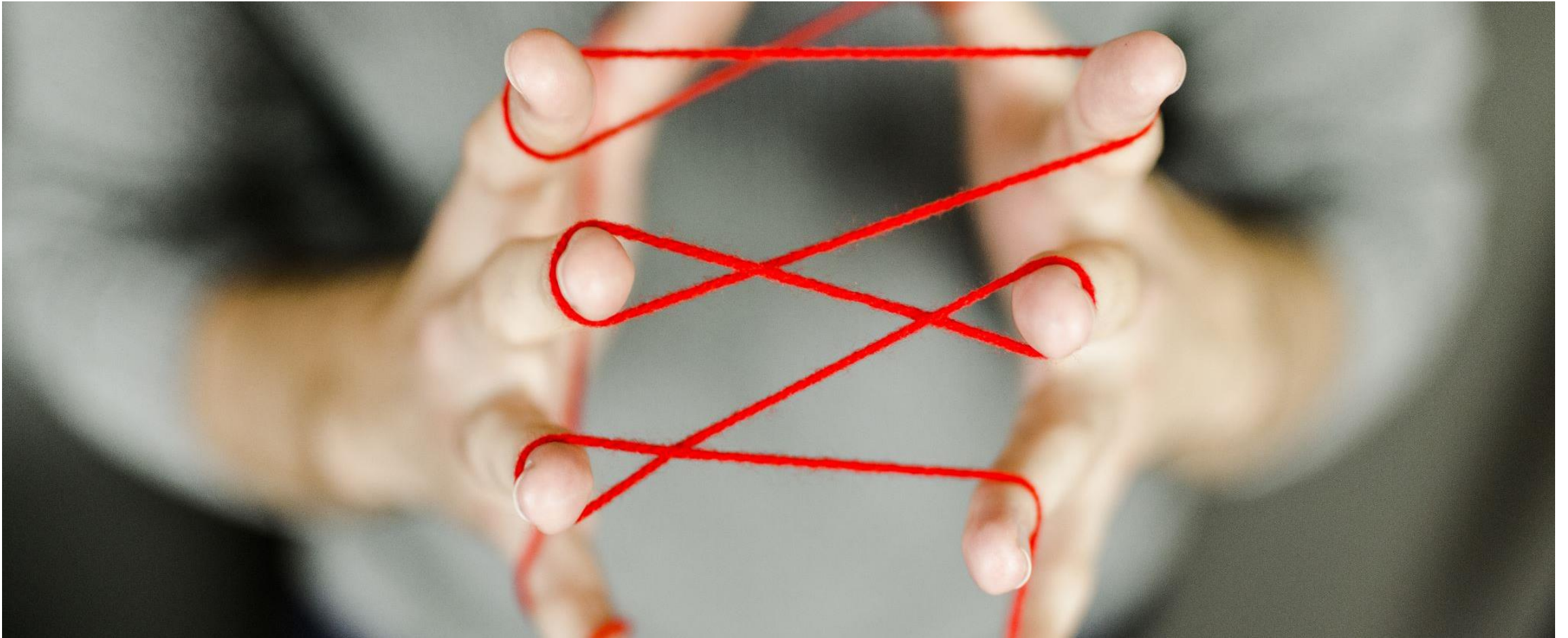
- Dixon and Coles (1997): home and away strength based on goals scored and conceded
- Artnzen & Hvattum (2010): rolling ELO ratings
- Peeters (2018): wisdom of the crowds measured by TM market values of players
- Holmes and McHale (2024): player-based model with fielded team taken as given.

Methods for within sample goodness-of-fit (Peeters, 2018):

- Psuedo R-squared
- AIC and BIC
- Log-likelihood



Tuning your model



Erasmus

Tuning and Training

Parameters vs. hyper parameters vs. features.

- Parameters are estimated during training or **learned from data**.
 - *Dixon–Coles (1997)*: team **attack, defence and home advantage**,.
 - *Arntzen & Hvattum (2010)*: **ordered-logit β** , relating elo difference to outcomes, and **cutpoints τ_1 and τ_2** .
 - Can be updated in test data set using **fixed update rule (we do not do this in assignment)**.
- Hyper parameters are **chosen** by the researcher
 - **Window length**: how many past matches enter estimation.
 - **Number of seasons included for player ratings** (Holmes & McHale, 2024).
 - **Recency/decay**: weight on past games (Dixon & Coles, 1997; Holmes & McHale, 2024).
 - **ELO update size**: learning rate $k \rightarrow$ how much ratings move after each game (Arntzen & Hvattum, 2010).
 - Not updated in test data set generally.
- Features/variables are variables that feed into models
 - **Examples**: ELO score, player ratings, Transfermarkt values
 - Can be updated in the test data set using **fixed update rule**

→ We *estimate* parameters and *tune* hyperparameters

→ To tune hyper parameters, we use *benchmarks*.



Benchmarks (needed for tuning)

Accuracy or Success ratio (*higher is better*)

- Evaluates how often the model predicts the right outcome by looking at highest probability outcome and comparing to actual outcome.
- Calculated as the percentage of cases that the model predicts right.
- **Pros:** Simple, intuitive headline metric.
- **Cons:** Ignores probability quality (51% = 99%); sensitive to outcome (im)balance.

Brier score (Quadratic Loss, MSE, or Square of RMSE, *lower is better*)

- Evaluates Quality of **probability forecasts** — penalizes being over/under-confident relative to what actually happened.
- **How it's computed (multi-class 1X2):**

$$Brier = \frac{1}{N} \sum \sum_{k \in \{H,D,A\}}^N (p_{ik} - o_{ik})^2$$

Where p_{ik} = predicted prob for outcome k (*home, draw, away*), $o_{ik} = 1$ if k occurred, else 0.

- Some papers divide by $K = 3$. **Do not do this in report (we want results to be comparable)**
- **Interpretation:** 0.67 \rightarrow model as good as random guess, 0 \rightarrow model predicts perfectly, good score is context dependent
- **Why use it?**
 - Uses full probability distribution (not just top pick)
 - Rewards good well-tuned models

Examples – Tuning decay Parameter

Dixon and Coles (1997) and Holmes and McHale (2024) **tune their time decay parameter.**

- **Dixon and Coles (1997)** less weight on observations further in the past in likelihood function
- **Holmes and McHale (2024)** less weight on player ratings/games (AWS) further in past when calculating player ratings
- Formula: $w_{it} = \exp(-\phi t)$, where t is number of days prior to calculation day.

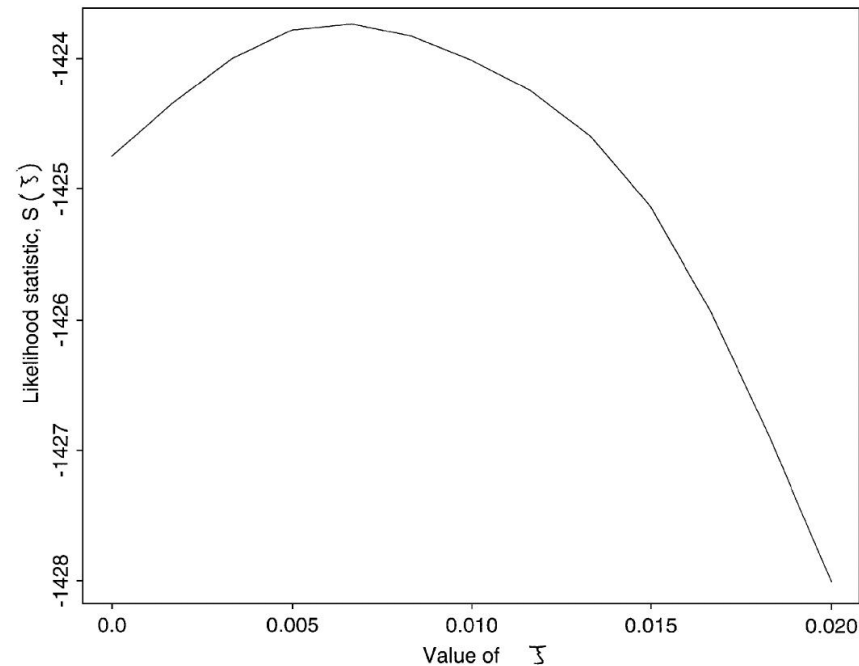
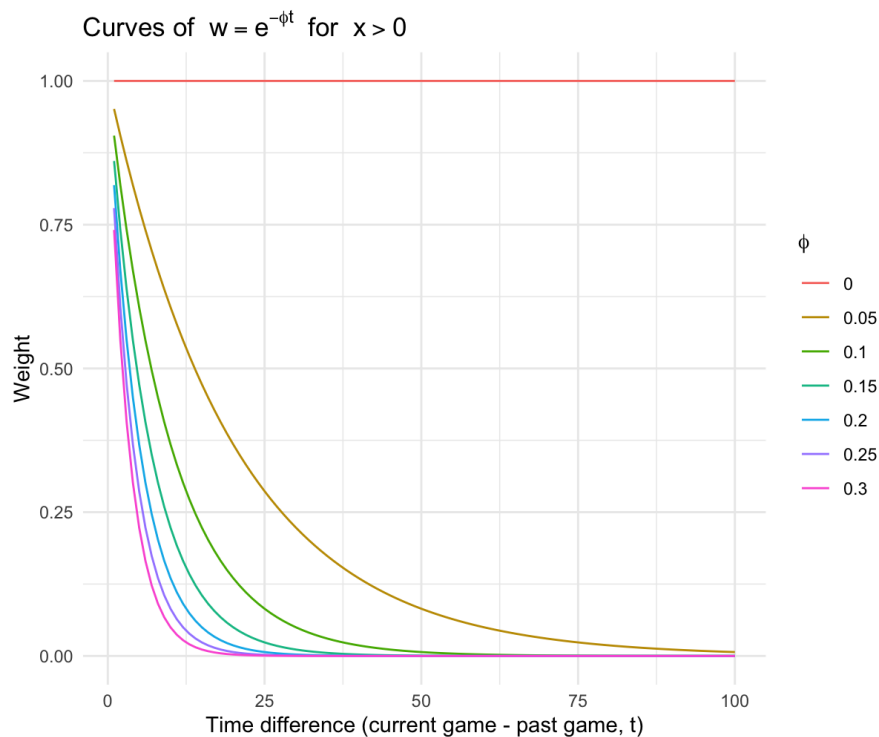


Fig. 1. $S(\xi)$ versus ξ : the maximum occurs at $\xi = 0.0065$

Dixon and Coles (1997)

Erasmus

Examples – Tuning updating Parameters

Arntzen and Hvattum (2010) tune two parameters at once

Elo formula:

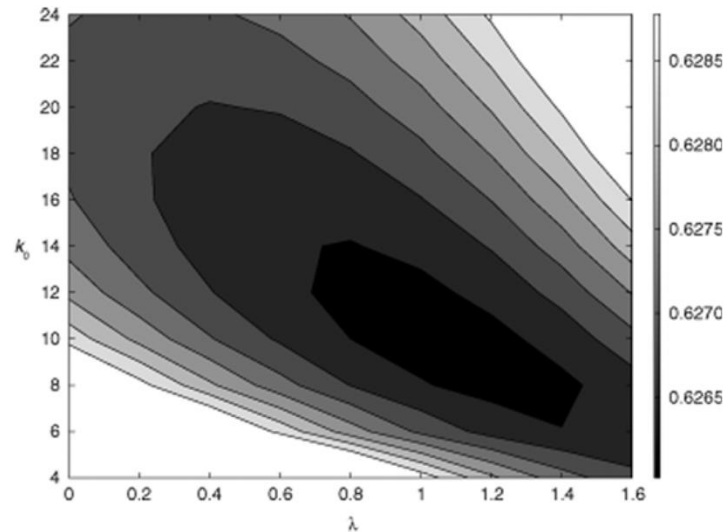
$$\ell_1^H = \ell_0^H + k(\alpha^H - \gamma^H).$$

k governs the amount of updating;
high k = a lot of weight on most recent game;
low k = less weight on most recent game

Goal adjusted k-factor:

$$k = k_0(1 + \delta)^\lambda,$$

k_0 and λ govern updating taking into account
goal difference δ



They end up using:

$$k_0 = 10$$

$$\lambda = 1$$

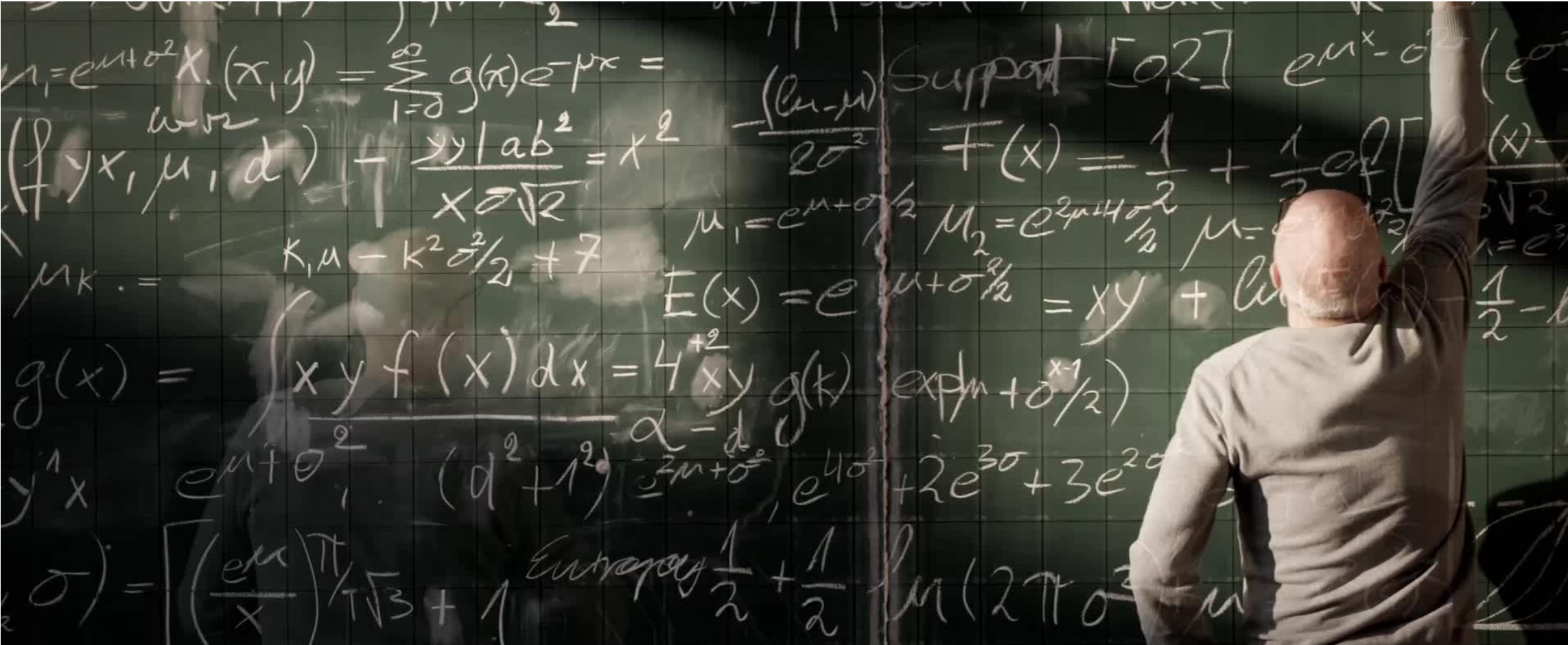
Arntzen

Fig. 1. Observed average quadratic loss when calibrating the two free parameters of the **ELO_g** method, k_0 and λ .

Training and Tuning Stage Summary

Study	Model Type	Training (data & method)	Validation (tuning)	Test (holdout)
Dixon & Coles (1997)	Bivariate Poisson for home/away goals	1992–1995 English league; estimate team attack/defence + home advantage ; exponential time weighting; derive 1X2 from goal model	Within 1992–1995 : choose decay parameter (and low-score tweak) by maximizing predictive score on past matches (time-ordered)	1995–1996 season; out-of-sample probabilities vs bookmaker odds
Hvattum & Arntzen (2010)	ELO ratings + ordered logit (rating gap → 1X2)	Initialize ELO on early seasons; estimate ordered-logit using ELO on 1995/96–1999/00	Tune k_0 (learning rate) and λ (margin amplifier) on pre-2000 seasons using Quadratic Loss (brier)	2000/01–2007/08 ; evaluate predictive scores and betting simulations
Holmes & McHale (2024)	Multinomial with player-based team strength (incl. pair/position effects)	Pre-train positional/interaction models 2013/14–2014/15 ; forecasting window 2015/16–2020/21	Inside 2015/16–2020/21 : first 80% train , last 20% of that block validate ; tune decay parameter , shrinkage, interaction set, EV threshold	Final 20% of 2015/16–2020/21; fixed parameters; report Brier/accuracy and betting vs bookmakers
Peeters (2018)	Ordered probit/logit using Transfermarkt squad values (crowd signal)	International matches (UEFA/CONMEBOL) 2008–2014 ; predictors: pre-match log squad value, home, squad size; rolling estimation	Peeters (2018) does limited tuning: comparing link (probit vs logit), feature sets, and daily vs monthly updating	Out-of-sample windows (held-out periods, e.g., later seasons/tournaments); compare to bookmakers, ELO, FIFA

Your teachers and their research



Ezra

Coordinator and Teacher

Sam Hoey:

- Web: www.samhoey.com
- Research using professional sports for labour/personnel economics
 - Managers
 - Racial discrimination
 - Career progression
- Lecture 1, 2, 3, 5, 6, 7, 8, 10



Erasmus

Teacher

Thomas Peeters

- Web: <https://sites.google.com/site/thpeeter/>
- Research on Professional sports:
 - TV rights and revenue sharing
 - Financial Fair Play
 - Managers
 - Mega events
 - Transfer system
- Other: inventors and innovation
- Lecture 4, 9



Erasmus

Erasmus Center for Applied Sports Economics

Research group for sports economics

- Web: <https://www.eur.nl/en/ease/ecase>
- ECASE football financial database
 - Financial accounts from 7 leagues
 - Canvas extra resources module
- In general we have many **sports data sets** which can be useful for bachelor thesis.
- Activities:
 - ECASE workshop: 27 February 2025



Practical issues

Erasmus

How to contact us?

YES:

- Personal questions: mail sportecon@ese.eur.nl
- Content questions: canvas discussion boards

NO:

- E-mail to teachers' personal accounts
- Canvas messages
- Phone calls
- Unannounced office visits

=> These will not be seen/answered

The Erasmus logo, featuring the word "Erasmus" in a stylized, handwritten script font.

Lectures (subject to change)

- Lectures on campus
 - Planning on Canvas
 - Watch out for room changes!
 - Lectures are **not** recorded or streamed
 - Slides will be posted
- Q&A session in final lecture
 - Tuesday 9 December 2022 13h00 – 15h00
 - Questions posted on Canvas discussion forum get priority
 - Additional questions in session are possible

The Erasmus logo, featuring a stylized, handwritten-style script of the word "Erasmus" in black.

Menti time



SH



Menti

New presentation



Choose a slide to present

What can we use to predict match outcomes?

past results
form
toto xg
injuries
home advantage

What is true about staking?

0 0 0
Setting a higher EV threshold leads to higher profits. Higher accuracy (the odds) means higher profits. Sharper odds quantifies risk adjusted returns.

Erasmus