

Separating the Crowds: Understanding determinants of home and away fan attendance at football matches

J. James Reade
Department of Economics, University of Reading
Oldham Athletic AFC

Work with Brad Humphreys, Dominik Schreyer and Carl
Singleton

ROSES Seminar

12 September 2025

Introduction

- Football matches are significant events:
 - ▶ Many thousands of people taking actions.
- Consequences for football clubs, authorities:
 - ▶ Preparing team for nature of match.
 - ▶ Preparing stadium for anticipated crowd.
 - ▶ Ensuring adequate capacity of local public infrastructure.

Introduction

- Strong past literature on modelling attendance demand:
 - ▶ Generally unsupportive of uncertainty as driver.
 - ▶ Quality matters, local economic conditions important.
- Most¹ past literature has treated crowd as homogenous:
 - ▶ Away fans travelling is a historic feature.
 - ▶ They've generally been segregated since the 1970s.
 - ▶ They often make up 10% or more of attendances.
 - ▶ Associated with 'better atmosphere'... also with trouble.
 - ▶ They represent a fan culture which differs across countries.

¹But not all... Pérez Hernández et al. (2023), Barajas and Gasparetto (2023), Aguiar Noury et al. (2025) and Varela-Quintana and Patuelli (2025).

Introduction

- In this presentation we:
 - ▶ Present a dataset of over 15,000 matches with away fan numbers.
 - ▶ Present a model of home and away fan attendance.
 - ▶ Create a model for forecasting home/away fan numbers.
 - ▶ Evaluate impact of expected/unexpected fans on outcomes.

Methodology

- Attendance demand a function of:
 - ▶ Quality of the product:
 - ▶ The teams participating but also their relative quality.
 - ▶ The attachment (brand loyalty) formed to team(s).
 - ▶ The broader event experience.
 - ▶ Price:
 - ▶ Ticket prices surprisingly difficult to collect.
 - ▶ But tickets not the only cost: travel, time, merch...
 - ▶ Income:
 - ▶ Hard to collect other than at local/regional level.
 - ▶ Football traditionally has working class appeal.
- Many of these aspects affect home and away fans differently.

Methodology

- Total attendance split: $att_{ijt} = Hatt_{ijt} + Aatt_{ijt}$.
- Estimate fairly standard attendance equation (Coates et al., 2014; Humphreys and Zhou, 2015):

$$\begin{aligned} \log(att_{ijt}) = & \gamma_0 + \gamma_1 probH_{ijt} + \gamma_2 probH_{ijt}^2 + \gamma_3 LS_t + \\ & \gamma_4 Q_{it} + \gamma_5 Q_{jt} + \gamma_6 distance_{ij} + \gamma_7 \log(population_i) \\ & + \gamma_8 \log(population_j) + u_{ijt}, \end{aligned} \tag{1}$$

- ▶ $probH_{ijt}$ is probability of home win. Details
- ▶ $\gamma_2 < 0$ indicates UOH, $\gamma_2 > 0$ implies loss aversion. concept
- Put $Hatt_{ijt}$ and $Aatt_{ijt}$ on LHS on (2).
- We add in many more variables to (2).

Methodology

- Construct forecast model:
 - ▶ Useful for planning considerations for clubs and authorities.
 - ▶ Away fan numbers in particular important.
- Presumption: factors affect away fans differently to home fans:
 - ▶ May not be appropriate to assume away fans 10% of total.
 - ▶ What is total attendance expected to be?
- Planning horizon likely longer than a few days.
 - ▶ Remove some variables, move to club-half-decade fixed effects.
- Reduced version of (2) produces \widehat{att}_{ijt} , \widehat{Hatt}_{ijt} , \widehat{Aatt}_{ijt} .
 - ▶ Of (some) interest: $\tilde{att}_{ijt} = \widehat{Hatt}_{ijt} + \widehat{Aatt}_{ijt}$ better than \widehat{att}_{ijt} ?
 - ▶ Evaluation based on forecast errors.

Methodology

- Finally: Consider the impact of crowds on outcomes.
- Regression models, for team i vs team j at time t :

$$outcome_{ijt} = \gamma_0 + \gamma_1 \log(Hatt_{ijt}) + \gamma_2 \log(Aatt_{ijt}) + \gamma_3 elo_{ijt} + u_{ijt}, \quad (2)$$

- Elo ratings to control for relative team strengths. Elo ratings
- Endogeneity issue: Fans want to watch better team?
- With our model we can evaluate expected fans as well as unexpected fans.
- Today: just match outcomes and goal difference as outcomes.

Data

- Multiple sources:
 - ▶ www.football-data.co.uk:
 - ▶ League matches in top 5 English leagues: Bet365 prices.
 - ▶ www.footballwebpages.co.uk:
 - ▶ Total attendances down to non-league Step 4 in England.
 - ▶ www.worldfootball.net:
 - ▶ Stadium information (location, capacity).
 - ▶ Websites: fanbanter.co.uk, the72.co.uk:
 - ▶ Numbers of away fans since 2013.
 - ▶ Club tweets on away fan attendance collected manually.
 - ▶ Some club fan sites: Wycombe, Barnet, Torquay.
 - ▶ Football clubs:
 - ▶ Reading FC, Oldham Athletic FC, Preston North End FC.

Team Twitter feed

Fan website

All data

Ratio

By club

Coverage of league matches

Variation by club

Travel network

Cross plot distance vs away fans

Data: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Attendance	137,265	8,493.730	12,022.710	0	89,874
Home attendance	15,382	8,158.513	7,908.720	0	80,018
Away attendance	15,513	873.069	948.513	0	9,500
Population estimate (home team)	159,017	243,820.100	162,328.600	53,232.000	1,195,447.000
Population estimate (away team)	159,115	243,447.500	161,790.700	53,232.000	1,195,447.000
Gross Value Added (home team Local Authority)	138,103	26,656.680	12,022.410	7,850.000	140,067.800
Gross Value Added (away team Local Authority)	138,202	26,644.300	11,973.990	7,850.000	140,067.800
Claimant Count (home team Local Authority)	156,427	5,997.528	7,184.490	205	83,920
Claimant Count (away team Local Authority)	156,524	5,968.457	7,109.722	205	83,920
Leisure price inflation	168,648	0.045	0.021	0.0004	0.132

Data: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Distance (miles)	159,083	95.696	60.834	0.000	334.141
Distance less than 10 miles	159,083	0.025	0.158	0	1
Public holidays	171,323	0.059	0.235	0	1
Days since last home match	170,520	35.508	386.489	0	46,053
Days since last away match	170,504	36.461	380.975	0	43,806
Years since last match-up between teams	171,323	12.629	20.897	0.003	140.052
First home match of season	171,323	0.068	0.252	0	1
First away match of season	171,323	0.070	0.256	0	1
Last away match of season	171,323	0.070	0.255	0	1
England play that day	171,323	0.023	0.149	0	1
England play previous day	171,323	0.018	0.134	0	1
England play next day	171,323	0.020	0.141	0	1
Champions League matches that day	171,323	0.098	0.297	0	1
Premier League matches that day	171,323	0.599	0.490	0	1
Win in last home match	171,323	0.435	0.496	0	1
Win in last away match	171,323	0.303	0.459	0	1
League position (home team)	127,968	12.331	6.701	1	28
League position (away team)	127,349	12.128	6.700	1	28
Predicted probability of home win	126,070	0.448	0.112	0.106	0.842

Data: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Goals scored (home)	105,762	27.710	18.569	0	116
Goals scored (away)	105,762	27.869	18.623	0	115
Points from 6th place (home team)	105,762	5.970	9.529	-40	63
Points from 6th place (away team)	105,762	5.756	9.522	-40	62
Home team promoted last season	116,135	0.120	0.325	0	1
Away team promoted last season	115,670	0.120	0.325	0	1
Home team promoted two seasons ago	110,135	0.123	0.328	0	1
Away team promoted two seasons ago	109,681	0.123	0.329	0	1
Home team relegated last season	116,135	0.107	0.310	0	1
Away team relegated last season	115,670	0.107	0.309	0	1
Home team relegated two seasons ago	110,135	0.107	0.309	0	1
Away team relegated two seasons ago	109,681	0.107	0.309	0	1
Standard deviation of league points	171,323	5.059	5.176	0.000	20.895
Total changes in league standings since previous league matches	171,323	7.418	7.447	0	26
Maximum air temperature (degrees)	150,303	12.867	5.542	-11.400	36.400
Mean wind speed (mph)	148,787	7.850	4.585	0.000	50.000
Precipitation (mm)	141,846	2.500	15.556	-99.900	4,302.800

Results: Basic

	Dependent variable:				
	lattendance		lhomeattt	lawayatt	
	(1)	(2)	(3)	(4)	(5)
Position (home)	-0.013*** (0.001)	-0.007*** (0.001)	-0.008*** (0.001)	-0.001 (0.001)	-0.001 (0.001)
Position (away)	-0.006*** (0.0004)	-0.004*** (0.001)	-0.003*** (0.001)	-0.017*** (0.001)	-0.017*** (0.001)
Prob of home win	-1.000*** (0.180)	-1.048*** (0.350)	-0.369 (0.326)	-2.542*** (0.738)	
Prob home win squared	1.499*** (0.184)	1.520*** (0.335)	0.937*** (0.312)	2.198*** (0.809)	
Prob of away win					-0.142 (0.682)
Prob away win squared					1.312 (1.038)
Observations	105,011	15,238	15,238	15,235	15,235
R ²	0.934	0.933	0.939	0.735	0.734
Adjusted R ²	0.934	0.931	0.938	0.727	0.727
Residual Std. Error	0.319 (df = 104476)	0.239 (df = 14835)	0.236 (df = 14835)	0.594 (df = 14832)	0.595 (df = 14832)

Note:

*p<0.1; **p<0.05; ***p<0.01

Coefficient plot

Quadratic plot

Results: Attendance demand — football

	Dependent variable:			
	lattendance		lhomeatt	lawayatt
	(1)	(2)	(3)	(4)
League position (home team)	-0.006*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)	-0.0001 (0.001)
League position (away team)	-0.004*** (0.0004)	-0.002*** (0.001)	-0.002*** (0.001)	-0.006*** (0.001)
Forecast probability of winning (home team)	-0.497** (0.210)	-0.808** (0.318)	-0.201 (0.293)	
Forecast probability of winning squared (home team)	1.047*** (0.222)	1.261*** (0.300)	0.709** (0.278)	
Forecast probability of winning (away team)				0.078 (0.590)
Forecast probability of winning squared (away team)				0.686 (0.927)
Won last match (home team)	0.029*** (0.003)	0.024*** (0.004)	0.024*** (0.004)	0.015 (0.010)
Won last match (away team)	0.009*** (0.002)	0.011*** (0.004)	0.009** (0.004)	0.033*** (0.012)
Observations	82,897	14,189	14,189	14,288
R ²	0.952	0.946	0.954	0.752
Adjusted R ²	0.952	0.944	0.953	0.745
Residual Std. Error	0.261 (df = 82425)	0.210 (df = 13802)	0.201 (df = 13802)	0.566 (df = 13900)

Note:

*p<0.1; **p<0.05; ***p<0.01

Coefficient plot

Results: Attendance demand — football

	<i>Dependent variable:</i>			
	lattendance	lhomeatt	lawayatt	
	(1)	(2)	(3)	(4)
Goals scored (home team)	0.003*** (0.001)	0.002* (0.001)	0.002* (0.001)	-0.0001 (0.001)
Goals scored (away team)	0.002*** (0.0003)	0.002*** (0.001)	0.002** (0.001)	0.005*** (0.001)
Points from top 6 positions (home team)	-0.003*** (0.001)	-0.002** (0.001)	-0.003*** (0.001)	-0.0003 (0.001)
Points from top 6 positions (away team)	-0.001*** (0.0004)	-0.001 (0.001)	0.0001 (0.001)	-0.007*** (0.002)
Promotion last season (home team)	0.088*** (0.013)	0.102*** (0.021)	0.093*** (0.023)	0.183*** (0.026)
Promotion last season (away team)	0.031*** (0.005)	0.042*** (0.009)	0.015** (0.007)	0.276*** (0.025)
Promotion two season ago (home team)	0.039*** (0.011)	0.038** (0.017)	0.038** (0.019)	0.021 (0.020)
Promotion two season ago (away team)	0.012*** (0.004)	0.013* (0.008)	0.008 (0.006)	0.062*** (0.020)
Observations	82,897	14,189	14,189	14,288
R ²	0.952	0.946	0.954	0.752
Adjusted R ²	0.952	0.944	0.953	0.745
Residual Std. Error	0.261 (df = 82425)	0.210 (df = 13802)	0.201 (df = 13802)	0.566 (df = 13900)

Note:

*p<0.1; **p<0.05; ***p<0.01

Coefficient plot

Results: Attendance demand — football

	<i>Dependent variable:</i>			
	lattendance	lhomeatt	lawayatt	
	(1)	(2)	(3)	(4)
Relegation last season (home team)	0.049*** (0.012)	0.045*** (0.016)	0.046*** (0.016)	0.084*** (0.021)
Relegation last season (away team)	0.024*** (0.005)	0.025*** (0.007)	0.017*** (0.006)	0.088*** (0.023)
Relegation two season ago (home team)	0.001 (0.012)	-0.016 (0.022)	-0.021 (0.024)	0.013 (0.017)
Relegation two season ago (away team)	0.002 (0.004)	0.008 (0.008)	0.010 (0.007)	-0.008 (0.022)
Standard deviation of league points	-0.002 (0.002)	-0.003 (0.003)	-0.002 (0.003)	-0.010** (0.005)
Total league standings changes	0.0004 (0.0003)	-0.0002 (0.0005)	-0.0001 (0.001)	-0.0004 (0.001)
Observations	82,897	14,189	14,189	14,288
R ²	0.952	0.946	0.954	0.752
Adjusted R ²	0.952	0.944	0.953	0.745
Residual Std. Error	0.261 (df = 82425)	0.210 (df = 13802)	0.201 (df = 13802)	0.566 (df = 13900)

Note:

*p<0.1; **p<0.05; ***p<0.01

Coefficient plot

Results: Attendance demand — scheduling

	Dependent variable:			
	lattendance		lhomeatt	lawayatt
	(1)	(2)	(3)	(4)
League position (home team)	-0.013*** (0.001)	-0.007*** (0.001)	-0.008*** (0.001)	-0.001 (0.001)
League position (away team)	-0.006*** (0.0004)	-0.004*** (0.001)	-0.003*** (0.001)	-0.018*** (0.001)
Forecast probability of winning (home team)	-0.845*** (0.173)	-0.922*** (0.338)	-0.305 (0.324)	
Forecast probability of winning squared (home team)	1.325*** (0.177)	1.392*** (0.321)	0.873*** (0.308)	
Forecast probability of winning (away team)				0.202 (0.607)
Forecast probability of winning squared (away team)				0.630 (0.918)
Distance (miles)	-0.001*** (0.0001)	-0.001*** (0.0001)	-0.0003*** (0.0001)	-0.007*** (0.0003)
Distance less than 10 miles	0.117*** (0.034)	0.194*** (0.040)	0.141*** (0.026)	0.359*** (0.101)
Observations	82,897	14,189	14,189	14,288
R ²	0.952	0.946	0.954	0.752
Adjusted R ²	0.952	0.944	0.953	0.745
Residual Std. Error	0.261 (df = 82425)	0.210 (df = 13802)	0.201 (df = 13802)	0.566 (df = 13900)

Note:

*p<0.1; **p<0.05; ***p<0.01

Coefficient plot

Results: Attendance demand — scheduling

	Dependent variable:			
	lattendance		lhomeatt	lawayatt
	(1)	(2)	(3)	(4)
Weekend early kick-off	0.063*** (0.021)	0.010 (0.015)	0.019 (0.014)	-0.001 (0.043)
Weekend late kick-off	0.023 (0.037)	-0.012 (0.047)	-0.015 (0.057)	-0.016 (0.065)
Public Holidays	0.097*** (0.006)	0.098*** (0.010)	0.088*** (0.009)	0.231*** (0.024)
Days since last match (home team)	0.0003*** (0.0001)	0.001 (0.0003)	0.001** (0.0003)	-0.001 (0.001)
Days since last match (away team)	-0.00004 (0.00004)	0.0002 (0.0002)	0.00001 (0.0003)	0.002** (0.001)
Years since both teams played eachother	0.002*** (0.0002)	0.002*** (0.0004)	0.001*** (0.0003)	0.008*** (0.001)
First home match	0.052*** (0.014)	0.017 (0.039)	-0.011 (0.037)	0.133* (0.078)
First away match	0.023* (0.012)	0.008 (0.028)	0.011 (0.031)	0.005 (0.062)
Last away match	0.127*** (0.008)	0.089*** (0.016)	0.082*** (0.017)	0.161*** (0.032)
Observations	82,897	14,189	14,189	14,288
R ²	0.952	0.946	0.954	0.752
Adjusted R ²	0.952	0.944	0.953	0.745
Residual Std. Error	0.261 (df = 82425)	0.210 (df = 13802)	0.201 (df = 13802)	0.566 (df = 13900)

Note:

*p<0.1; **p<0.05; ***p<0.01

Coefficient plot

Results: Attendance demand — scheduling

	<i>Dependent variable:</i>			
	lattendance	lhomeatt	lawayatt	
	(1)	(2)	(3)	(4)
England play that day	-0.053*** (0.009)	-0.018 (0.016)	-0.005 (0.016)	-0.098*** (0.033)
England play previous day	0.025** (0.011)	0.046** (0.019)	0.051*** (0.018)	0.005 (0.030)
England play next day	-0.002 (0.010)	0.026* (0.015)	0.028* (0.014)	-0.013 (0.039)
Champions League matches that day	-0.038*** (0.006)	-0.010 (0.011)	-0.005 (0.012)	-0.051** (0.021)
Premier League matches that day	-0.021*** (0.004)	-0.007 (0.007)	-0.008 (0.006)	-0.039*** (0.012)
Observations	104,701	15,238	15,238	15,364
R ²	0.937	0.939	0.941	0.855
Adjusted R ²	0.937	0.937	0.939	0.851
Residual Std. Error	0.311 (df = 104154)	0.228 (df = 14819)	0.233 (df = 14819)	0.439 (df = 14944)

Note:

*p<0.1; **p<0.05; ***p<0.01

Coefficient plot

Results: Attendance demand — economic/demographic/weather

	Dependent variable:			
	lattendance	lhomeatt	lawayatt	
	(1)	(2)	(3)	(4)
League position (home team)	-0.011*** (0.0002)	-0.008*** (0.0004)	-0.008*** (0.0004)	-0.001 (0.001)
League position (away team)	-0.005*** (0.0002)	-0.004*** (0.0004)	-0.003*** (0.0004)	-0.017*** (0.001)
Forecast probability of winning (home team)	-0.818*** (0.076)	-0.981*** (0.177)	-0.314* (0.175)	
Forecast probability of winning squared (home team)	1.261*** (0.081)	1.455*** (0.188)	0.887*** (0.187)	
Forecast probability of winning (away team)			0.005 (0.420)	
Forecast probability of winning squared (away team)			1.069* (0.636)	
Log population estimate (home team)	0.025 (0.027)	-0.053 (0.053)	0.005 (0.052)	-0.532*** (0.133)
Log population estimate (away team)	-0.222*** (0.028)	-0.580*** (0.059)	-0.614*** (0.059)	-0.410*** (0.149)
Log Gross Value Added (home team)	-0.137*** (0.021)	-0.107** (0.049)	-0.090* (0.048)	-0.165 (0.122)
Log Gross Value Added (away team)	-0.138*** (0.020)	-0.186*** (0.050)	-0.185*** (0.049)	-0.005 (0.125)
Log claimant count (home team)	-0.035*** (0.007)	0.001 (0.015)	-0.001 (0.015)	0.021 (0.038)
Log claimant count (away team)	-0.109*** (0.008)	0.001 (0.015)	-0.002 (0.015)	0.021 (0.039)
Leisure.price.inflation	-0.061 (0.129)	0.759*** (0.260)	0.870*** (0.258)	0.022 (0.656)
Maximum air temperature (degrees celcius)	0.002*** (0.0004)	0.001* (0.001)	0.001 (0.001)	0.002 (0.002)
Mean wind speed (mph)	0.001*** (0.0003)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.002)
Precipitation amount (mm)	-0.0001	0.00004	-0.00001	0.0004

Results: Forecasting

- Forecasting model a pared down version of combination of models reported.
- Forecasts online at
<https://home-and-away-fan-forecasts.streamlit.app/>.

Statistic	N	Mean	St. Dev.	Min	Max
Aggregate mean forecast errors	1,908	404.093	1,704.424	-12,236.470	9,835.673
Home mean forecast errors	1,852	579.750	1,716.542	-6,306.876	11,325.580
Away mean forecast errors	1,852	35.274	590.615	-5,149.200	3,149.171
Disaggregate mean forecast errors	1,852	615.024	1,895.882	-10,567.790	11,922.270
Aggregate root mean squared errors	1,908	1,088.675	1,372.080	0.437	12,236.470
Home root mean squared errors	1,852	1,126.465	1,418.872	0.012	11,325.580
Away root mean squared errors	1,852	323.530	495.321	0.022	5,149.200
Disaggregate root mean squared errors	1,852	1,266.629	1,538.704	0.153	11,922.270
Aggregate mean absolute percentage errors	1,852	0.127	0.114	0.0001	1.527
Home mean absolute percentage errors	1,852	0.124	0.113	0.00000	1.810
Away mean absolute percentage errors	1,852	0.320	0.341	0.00003	3.820
Disaggregate mean absolute percentage errors	1,852	0.124	0.110	0.00002	1.521

Basic Results: Outcomes

	Dependent variable											
	goal difference			home win (0/1)			draw (0/1)			away win (0/1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Elo prediction	3.207*** (0.058)	2.944*** (0.107)	2.069*** (0.154)	0.729*** (0.012)	0.702*** (0.037)	0.488*** (0.047)	-0.075*** (0.010)	-0.093*** (0.034)	-0.077** (0.042)	-0.054*** (0.010)	-0.059*** (0.022)	-0.410*** (0.036)
Log total attendance	-0.155*** (0.018)	-0.066 (0.032)		-0.042*** (0.004)	-0.035*** (0.013)		0.027*** (0.003)	0.037*** (0.010)		0.016*** (0.004)		-0.002 (0.012)
Log home attendance		0.264*** (0.057)			0.061*** (0.012)			0.027** (0.012)				-0.078*** (0.014)
Log away attendance			-0.012*** (0.024)		-0.020*** (0.006)			0.007 (0.005)				0.022*** (0.005)
Observations	132,606	14,648	14,638	132,606	14,648	14,638	132,606	14,648	14,638	132,606	14,648	14,638
R-squared	0.038	0.057	0.021	0.061	0.060	0.065	0.013	0.017	0.031	0.061	0.056	0.077
Adjusted R ²	0.091	0.074	0.094	0.054	0.047	0.057	0.003	0.003	0.001	0.055	0.042	0.049
Residual Std. Error	1.693 (df = 131643)	1.615 (df = 14436)	1.597 (df = 14202)	0.484 (df = 131843)	0.484 (df = 14436)	0.481 (df = 14202)	0.431 (df = 131643)	0.440 (df = 14436)	0.440 (df = 14202)	0.447 (df = 131643)	0.451 (df = 14436)	0.449 (df = 14202)
Final effects	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	
Clustering	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	

Note: *p<0.1; **p<0.05; ***p<0.01

Basic Results: Yellow cards

	Dependent variable:								
	HY - AY			HY			AY		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elo prediction	-1.304*** (0.060)	-1.145*** (0.161)	-0.839*** (0.197)	-0.918*** (0.047)	-0.858*** (0.111)	-0.519*** (0.133)	0.386*** (0.063)	0.287** (0.117)	0.321** (0.145)
Log total attendance	0.033 (0.026)	0.067 (0.059)		0.139*** (0.023)	0.184*** (0.062)		0.106*** (0.021)	0.117** (0.055)	
Log home attendance			-0.140** (0.065)			-0.053 (0.053)			0.087 (0.064)
Log away attendance			0.107*** (0.025)			0.124*** (0.012)			0.017 (0.018)
Observations	55,611	12,075	12,074	55,611	12,075	12,074	55,611	12,075	12,074
R ²	0.027	0.032	0.055	0.042	0.060	0.085	0.027	0.043	0.065
Adjusted R ²	0.024	0.018	0.030	0.039	0.047	0.061	0.024	0.029	0.040
Residual Std. Error	1.539 (df = 55427)	1.594 (df = 11906)	1.584 (df = 11763)	1.143 (df = 55427)	1.169 (df = 11906)	1.160 (df = 11763)	1.264 (df = 55427)	1.280 (df = 11906)	1.272 (df = 11763)
Fixed effects	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S
Clustering	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S

Note:

*p<0.1; **p<0.05; ***p<0.01

Basic Results: Fouls

	Dependent variable:								
	HF - AF			HF			AF		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elo prediction	-3.107*** (0.283)	-3.277*** (0.580)	-2.820** (0.655)	-2.043*** (0.197)	-2.219*** (0.375)	-1.477*** (0.452)	1.064*** (0.267)	1.058*** (0.349)	1.343*** (0.456)
Log total attendance	-0.445*** (0.118)	-0.337* (0.177)		-0.076 (0.109)	0.024 (0.203)		0.369*** (0.128)	0.362** (0.162)	
Log home attendance			-0.812** (0.305)			-0.482** (0.216)			0.330 (0.250)
Log away attendance			0.283*** (0.089)			0.283*** (0.016)			0.001 (0.075)
Observations	51,940	10,382	10,381	51,940	10,382	10,381	51,940	10,382	10,381
R ²	0.030	0.051	0.096	0.086	0.061	0.097	0.075	0.063	0.099
Adjusted R ²	0.026	0.037	0.070	0.082	0.047	0.072	0.072	0.049	0.073
Residual Std. Error	4.760 (df = 51766)	4.694 (df = 10223)	4.611 (df = 10091)	3.611 (df = 51766)	3.448 (df = 10223)	3.403 (df = 10091)	3.762 (df = 51766)	3.544 (df = 10223)	3.499 (df = 10091)
Fixed effects	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S
Clustering	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S

Note:

*p<0.1; **p<0.05; ***p<0.01

Basic Results: Fouls per yellow card

	Dependent variable:								
	(HF + AF)/(HY + AY)				HF/HY		AF/AY		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Elo prediction	1.269*** (0.236)	1.432** (0.554)	0.957 (0.682)	0.931*** (0.145)	0.780** (0.286)	0.369 (0.396)	-0.277 (0.183)	-0.110 (0.324)	0.131 (0.449)
Log total attendance	-0.728*** (0.101)	-0.579** (0.276)		-0.461*** (0.050)	-0.440*** (0.124)		-0.063 (0.078)	-0.132 (0.165)	
Log home attendance			-0.150 (0.214)			-0.168 (0.136)			-0.175 (0.184)
Log away attendance				-0.269*** (0.081)		-0.121** (0.044)			-0.056 (0.047)
Observations	51,940	10,382	10,381	51,932	10,382	10,381	51,938	10,382	10,381
R ²	0.048	0.063	0.081	0.060	0.048	0.064	0.042	0.050	0.070
Adjusted R ²	0.044	0.049	0.055	0.057	0.033	0.037	0.039	0.035	0.043
Residual Std. Error	6.409 (df = 51766)	5.826 (df = 10223)	5.808 (df = 10091)	3.862 (df = 51758)	3.679 (df = 10223)	3.672 (df = 10091)	3.948 (df = 51764)	3.776 (df = 10223)	3.760 (df = 10091)
Fixed effects	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S
Clustering	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S	HT, S	HT, S	HT, AT, S

Note:

*p<0.1; **p<0.05; ***p<0.01

Expected/unexpected fans

	Dependent variable:			
	outcomeH	outcomeD	outcomeA	Gdiff
	(1)	(2)	(3)	(4)
Constant	−0.068** (0.027)	0.075 (0.054)	0.010 (0.019)	−1.827*** (0.089)
eloPH	0.992*** (0.054)			4.191*** (0.179)
eloPD		0.804*** (0.214)		
eloPA			1.130*** (0.064)	
home.forecasts	0.00001*** (0.00000)	−0.00000 (0.00000)	−0.00000*** (0.00000)	0.00002*** (0.00000)
home.forecast.errors	0.00001** (0.00000)	−0.00000 (0.00000)	−0.00001* (0.00000)	0.00003*** (0.00001)
away.forecasts	−0.00004*** (0.00001)	0.00000 (0.00001)	0.00003*** (0.00001)	−0.0001*** (0.00002)
away.forecast.errors	−0.00004*** (0.00001)	0.00001* (0.00001)	0.00003*** (0.00001)	−0.0002*** (0.00003)
Observations	12,592	12,592	12,592	12,592
R ²	0.043	0.002	0.040	0.067
Adjusted R ²	0.043	0.001	0.039	0.067
Residual Std. Error (df = 12586)	0.484	0.443 26	0.451	1.591

Conclusions

- Present dataset of 15,000+ away attendances at English/Welsh football matches.
- We find home (non-home) win preference dominates for home (away) fans.
- Find different drivers for away fans relative to home fans.
- Indicative evidence that away fans matter for outcomes.
 - ▶ Both those expected and those unexpected.
- Dataset will be made available on Github.

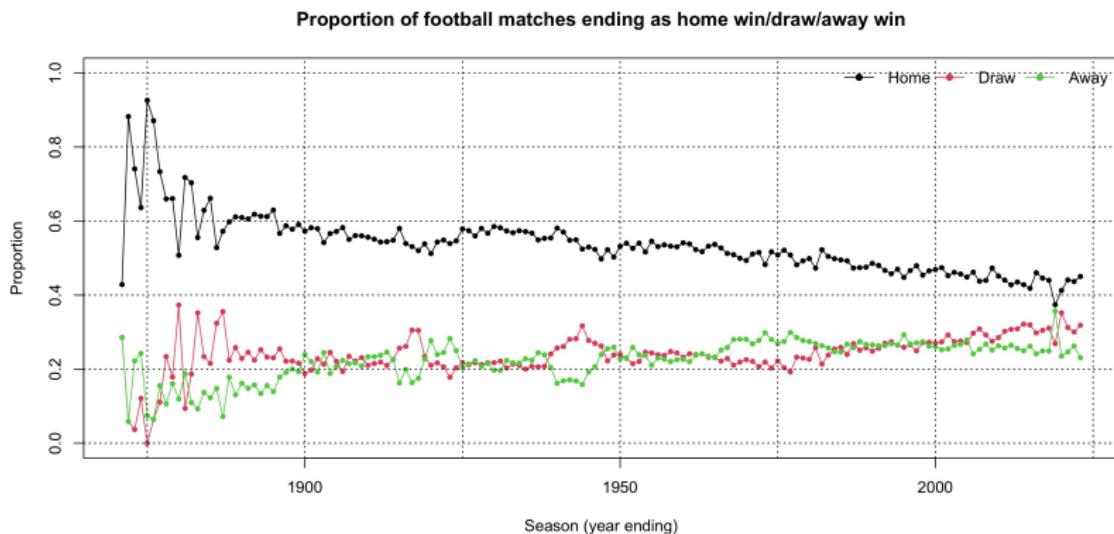
Bibliography I

- A. Aguiar Noury, D. Coates, and A. Feddersen. Determinants of Visitor Attendance. Presentation, European Association of Sports Economics Annual Conference, August 2025.
- A. Barajas and T. Gasparetto. They are not all the same: Determinants of attendance across different sectors in a stadium. *Plos one*, 18(7):e0289331, 2023.
- D. Coates, B.R. Humphreys, and L. Zhou. Reference-dependent preferences, loss aversion, and live game attendance. *Economic Inquiry*, 52(3):959–973, 2014.
- G. Elaad, J.J. Reade, and C. Singleton. Information, prices and efficiency in an online betting market. *Finance Research Letters*, 2020. ISSN 1544-6123. doi:
<https://doi.org/10.1016/j.frl.2019.09.006>. URL
<http://www.sciencedirect.com/science/article/pii/S1544612319306440>.

Bibliography II

- J. Goddard. Regression Models for Forecasting Goals and Match Results in Association Football. *International Journal of Forecasting*, 21:331–340, 2005.
- B.R. Humphreys and L. Zhou. The Louis–Schmelling paradox and the league standing effect reconsidered. *Journal of Sports Economics*, 16(8):835–852, 2015.
- I. Pérez Hernández, P. de Frutos Madrazo, B. Rodriguez-Prado, and P.A. Martin-Cervantes. A gravity model to explain the flows away of football fans: A panel data analysis. *The Social Science Journal*, pages 1–19, 2023.
- C. Varela-Quintana and R. Patuelli. Following my Team on the Road: The Impact of Top-Division Football Clubs on Tourism. Presentation, European Association of Sports Economics Annual Conference, August 2025.

Historical Match Outcomes



Example I

- Crawley Town 2-2 Oldham Athletic, March 1 2022.
 - ▶ Attendance: 1,927 (248 away fans).



Example II

- Oldham Athletic vs Port Vale, Nov 20 2021.
 - ▶ Attendance: 4,993 (1,913 away fans).



Data



Tweet



Carlisle United FC @officialcufc

...



600 of you here today - thank you for your support ❤️
#cufc



5:37 PM · Apr 23, 2022 · Twitter for iPhone

4 Retweets

2 Quote Tweets

61 Likes

33

Data



FAN banter



HOME

ENGLISH PREMIER LEAGUE

CHAMPIONSHIP

LEAGUE ONE

LEAGUE

These were the top 5 away attendances for League One at the weekend...

Seal

1) Sheffield Wednesday fans at Wycombe Wanderers –
1,841

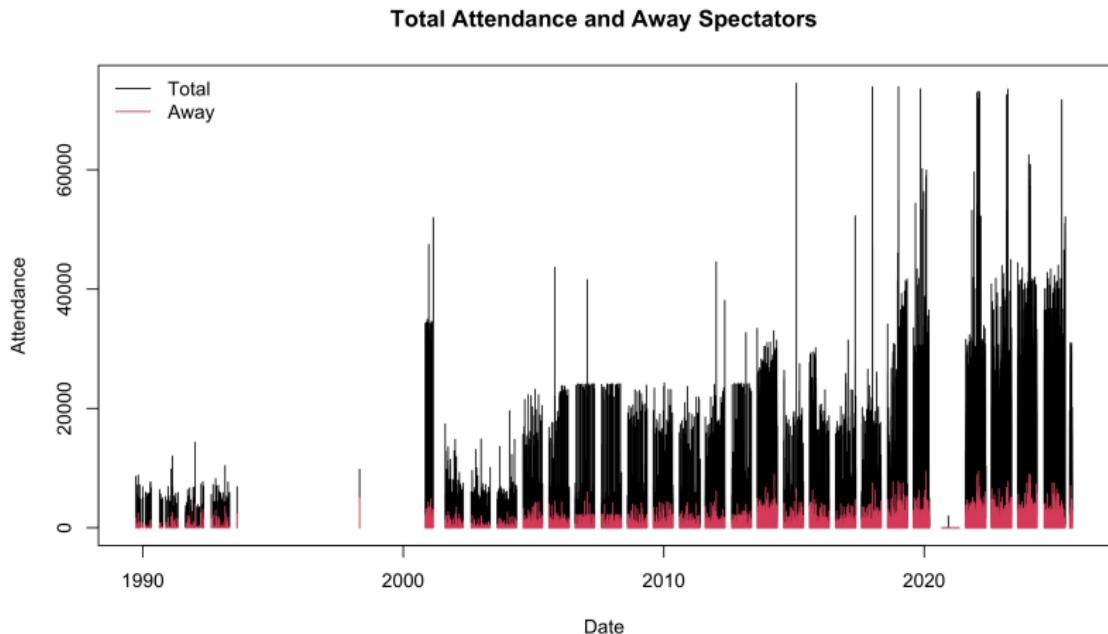
The away end at Adams Park this
afternoon... #SWFC
pic.twitter.com/HGp0cdESPD

– Rob Staton (@robstaton) [April 23, 2022](#)

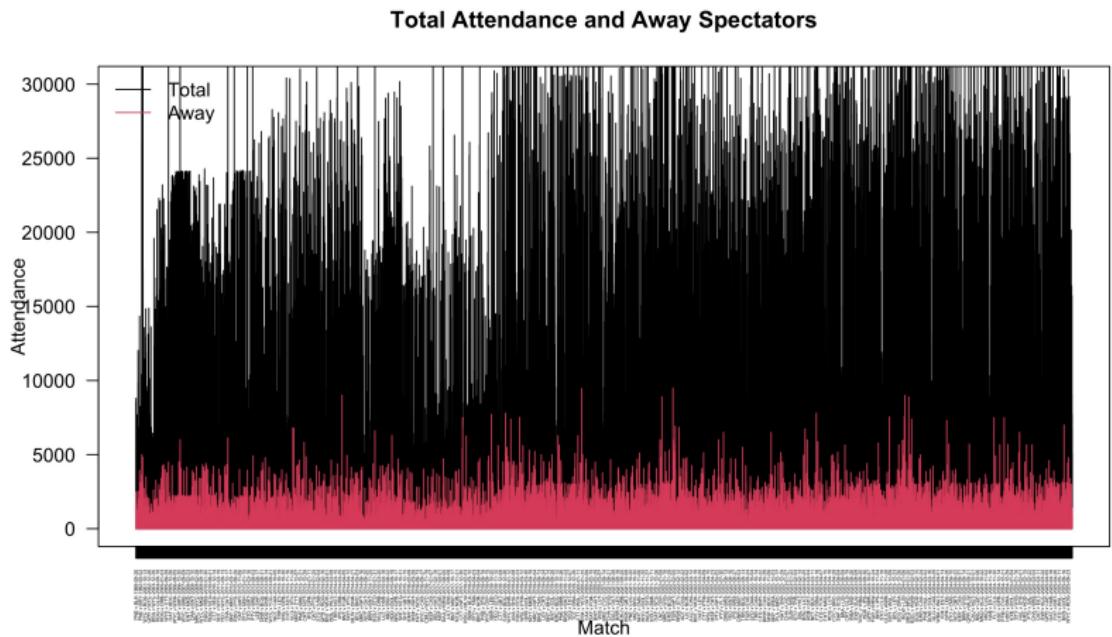
2) Plymouth Argyle fans at Wigan Athletic – 1,700

Wigan Athletic – PLYMOUTH ARGYLE 📸🔥
23/4/22@MrJordanWard 📸🔥#pafc
pic.twitter.com/JcBVtYvHNx

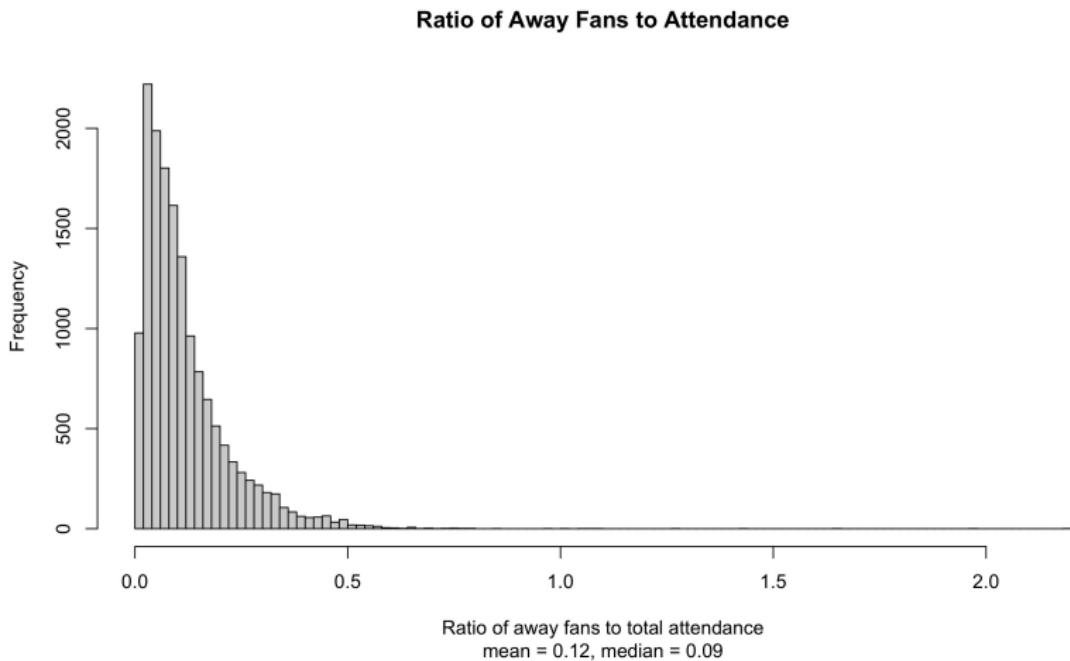
Separating the Crowds



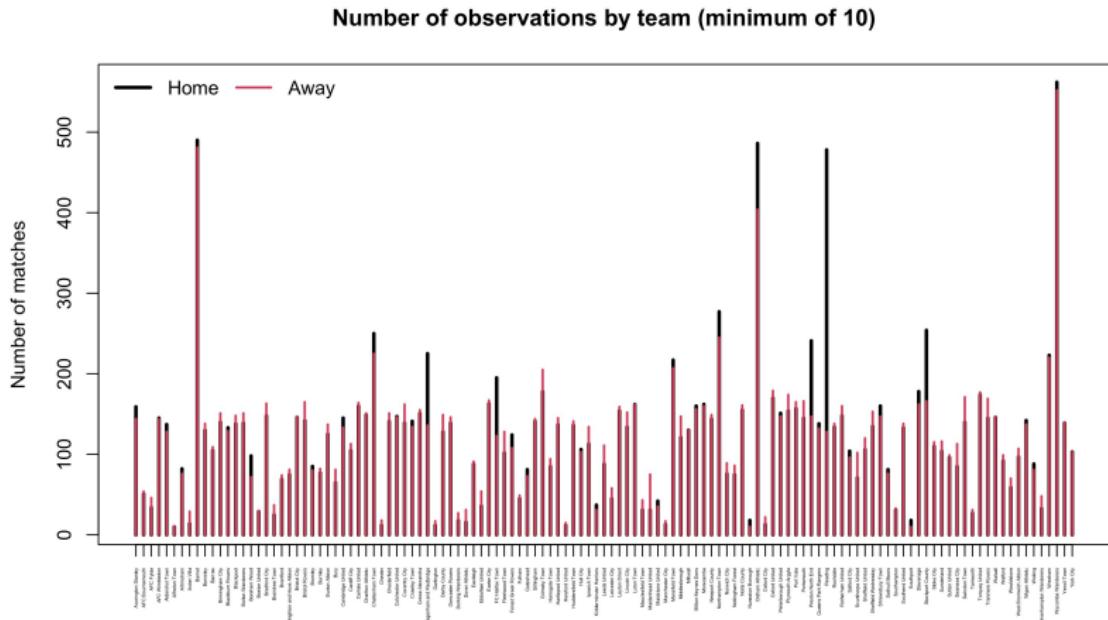
Data



Data II

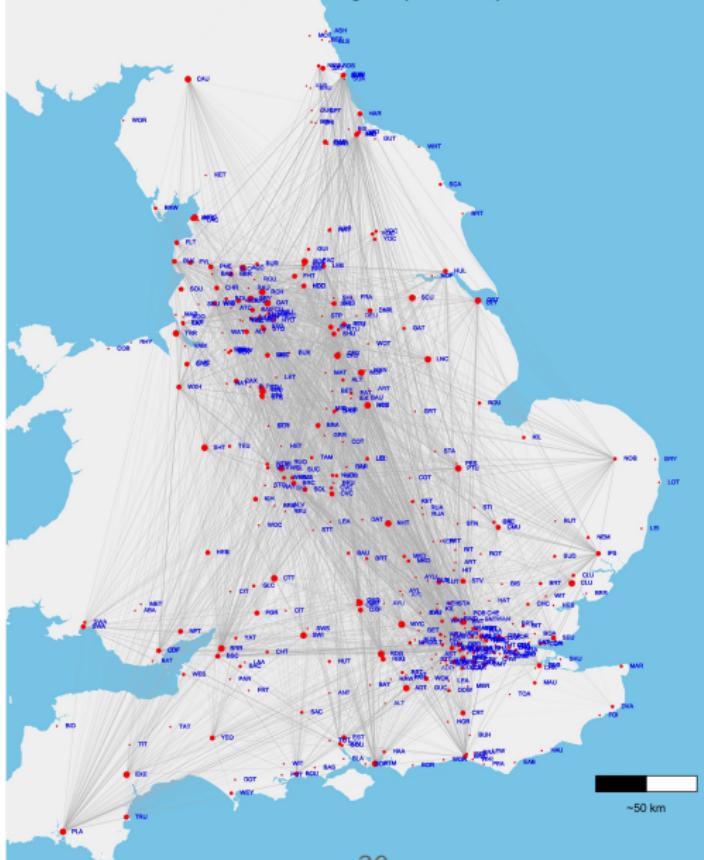


Separating the Crowds

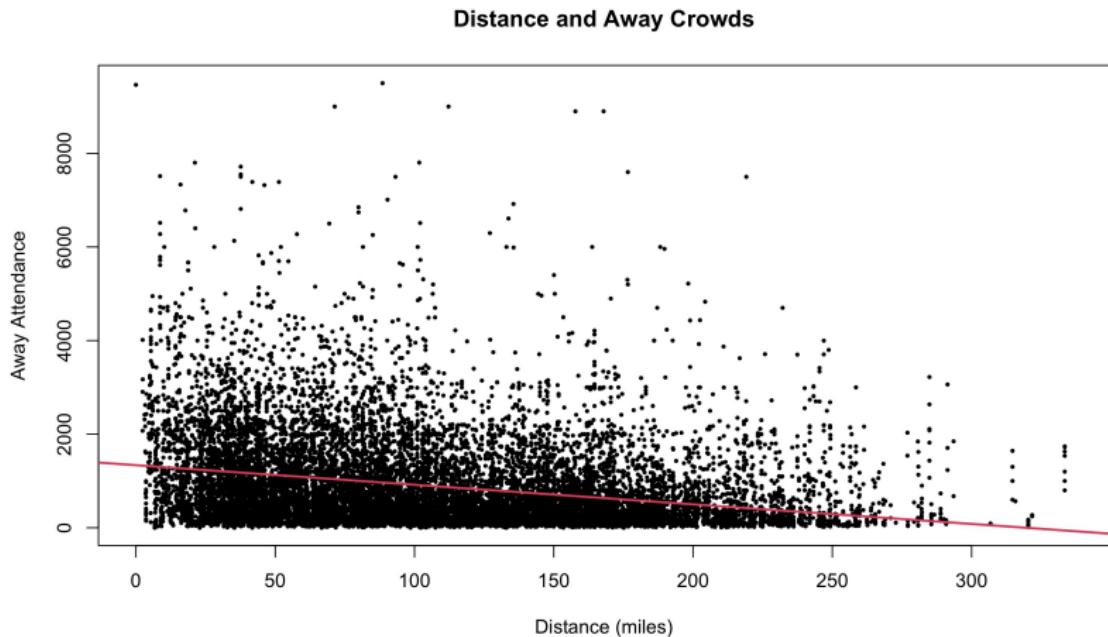


Separating the Crowds

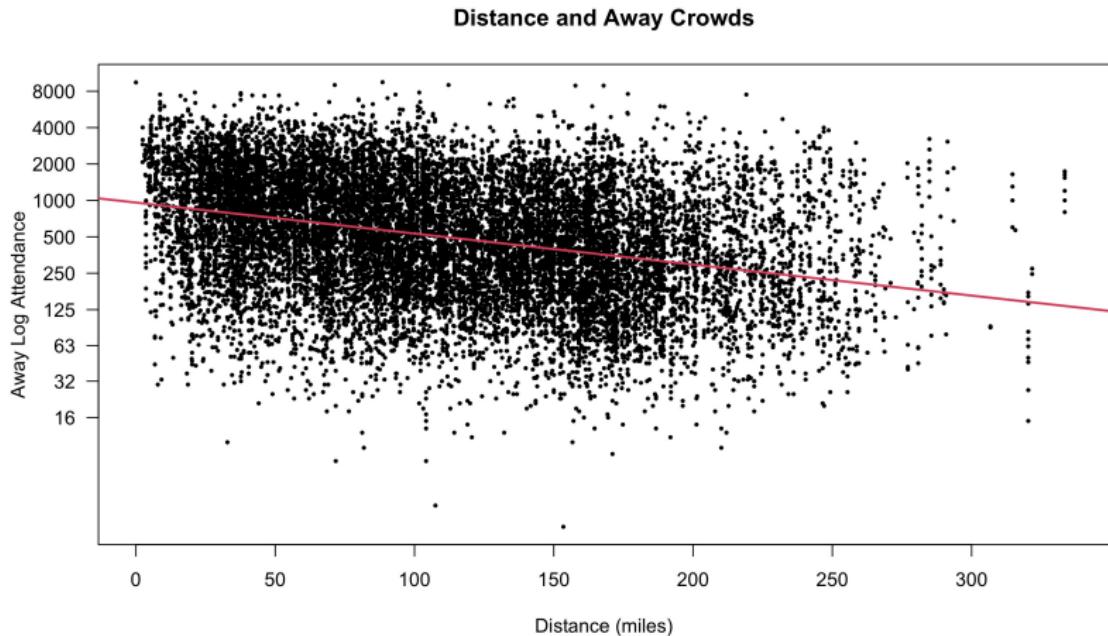
Fan Travel in England (and Wales)



Separating the Crowds



Separating the Crowds



Elo ratings and predictions

- Elo ratings assign strength rating to each participant.
- For two participants at time t , R_{it} , R_{jt} .
- R_{it} , R_{jt} used to create prediction:

$$elo_{j|t} = \frac{1}{1 + 10^{(R_{jt} - R_{it})/400}}. \quad (3)$$

- By design, $0 \leq elo_{j|t} \leq 1$, for participant j $elo_{j|t} = 1 - elo_{j|t}$.
- Outcome Y_1 , ratings adjust: $R_{t+1,1} = R_{t,1} + K(Y_1 - X_1)$.

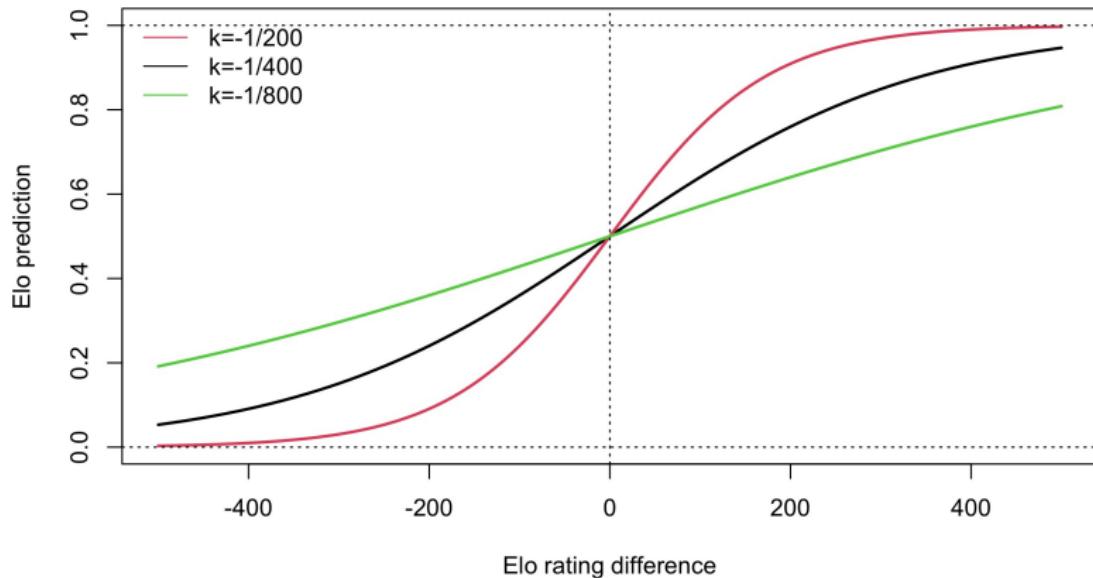
Concept

Example

Back

Elo ratings and predictions

Logarithmic Curves in Elo Ratings



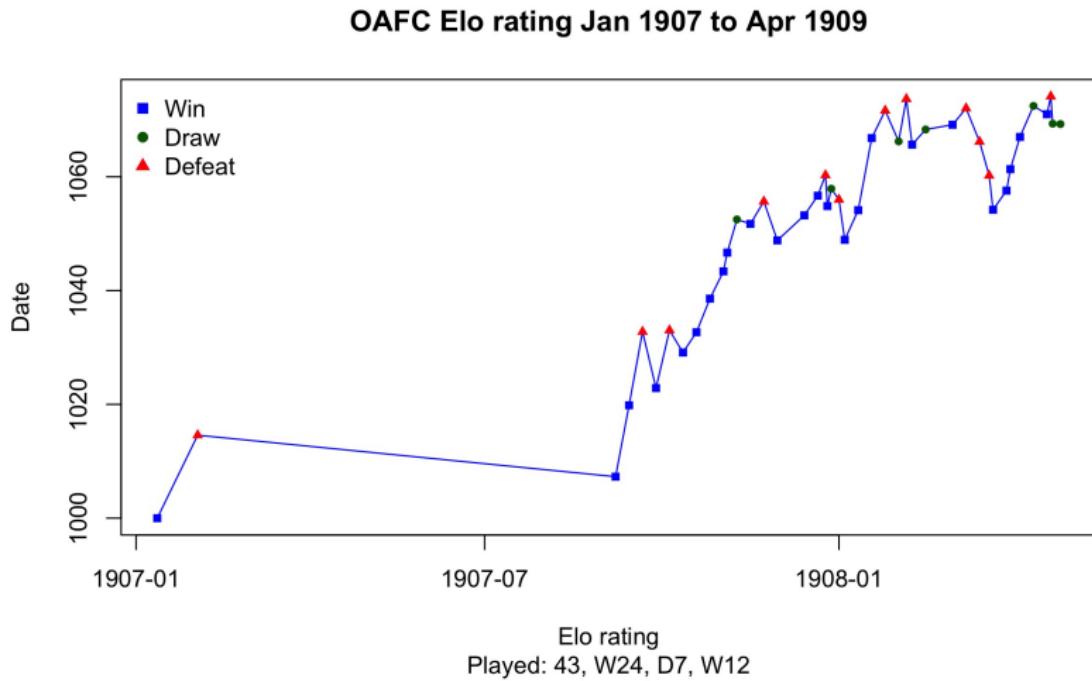
Back

Elo ratings example: Oldham Athletic

Date	Home		Goals		Away team	Home	Elo ratings		Probabilities		
	Date	team	Home	Away			Prediction	Away	Home	Draw	Away
1907-01-12		Oldham	5	0	Kidderminster	1000.0	0.51	990.5	50.2	24.6	25.3
1907-02-02		Oldham	0	1	Liverpool	1014.6	0.24	1211.8	24.3	24.3	51.4
1907-09-07		Stoke	1	3	Oldham	1097.7	0.63	1007.3	61.9	20.8	17.3
1907-09-14		Oldham	2	1	West Brom	1019.8	0.35	1124.7	33.9	26.2	39.9
1907-09-21		Bradford	1	0	Oldham	1036.1	0.50	1032.8	49.2	24.8	26.0
1907-09-28		Oldham	3	0	Hull	1022.9	0.49	1028.1	47.9	25.1	27.0
1907-10-05		Derby	1	0	Oldham	1110.8	0.61	1033.0	60.2	21.4	18.4
1907-10-12		Oldham	4	0	Lincoln	1029.1	0.64	926.3	63.5	20.1	16.3
1907-10-19		Fulham	1	2	Oldham	1096.2	0.59	1032.7	58.2	22.2	19.7
1907-10-26		Oldham	1	0	Barnsley	1038.6	0.52	1024.5	50.9	24.4	24.7

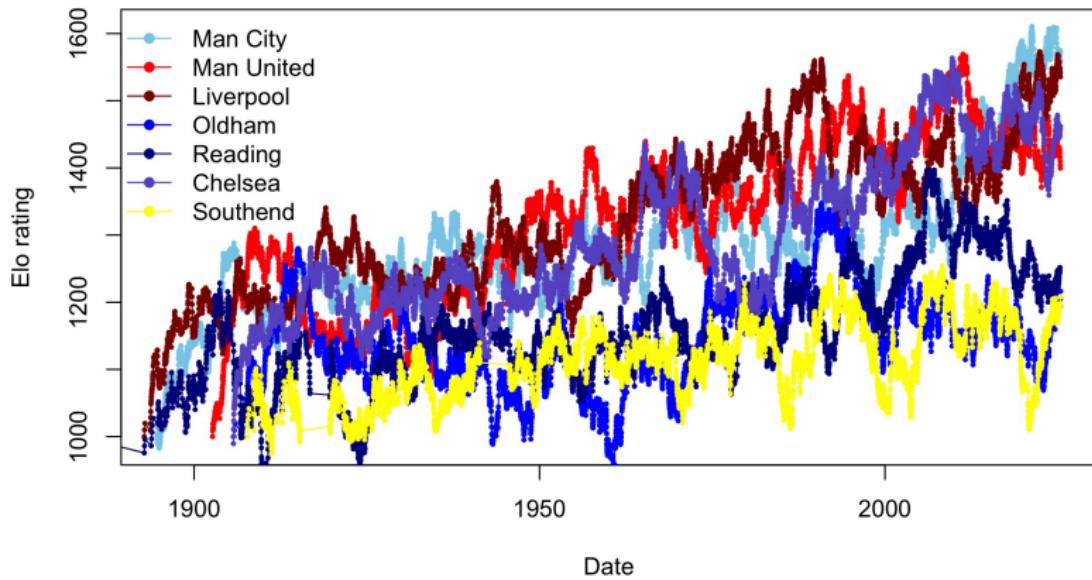
Back

Elo ratings example: Oldham Athletic



[Back](#)

Elo ratings example: Oldham Athletic



Back

Elo top 10

Rank	team	elo strength	division	league position
1	Manchester City	1554	English Premier	5
2	Liverpool	1546	English Premier	1
3	Arsenal	1515	English Premier	2
4	Newcastle United	1481	English Premier	6
5	Chelsea	1452	English Premier	4
6	Brighton and Hove Albion	1429	English Premier	10
7	Manchester United	1424	English Premier	13
8	Tottenham Hotspur	1422	English Premier	14
9	Aston Villa	1412	English Premier	8
10	Fulham	1408	English Premier	9

Back

Elo divisional mix

Rank	Team	elostrength	division	league pos
61	Exeter City	1215	Football League One	18
62	Walsall	1215	Football League Two	1
63	Hull City	1214	Football League Championship	20
64	Chorley	1213	Football Conference North	4
65	Doncaster Rovers	1213	Football League Two	5
66	Huddersfield Town	1209	Football League One	6
67	Oldham Athletic	1208	Football Conference	4
68	Cardiff City	1208	Football League Championship	19
69	AFC Wimbledon	1207	Football League Two	4
70	Halifax Town	1207	Football Conference	6

Probability model

Example

- MK Dons vs Oldham Athletic, attendance: 9,384. Away fans: 1,891, home fans: 7,493.

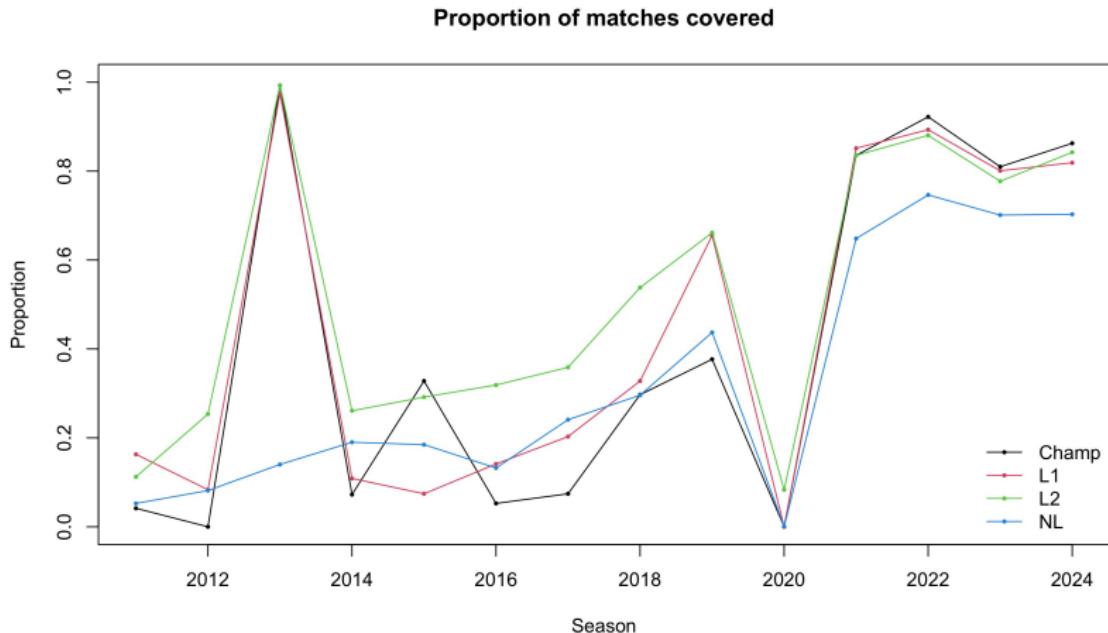


Example

- FC Halifax Town vs Oldham Athletic, attendance: 4,202.
Away fans: 2,404, home fans: 1,798.

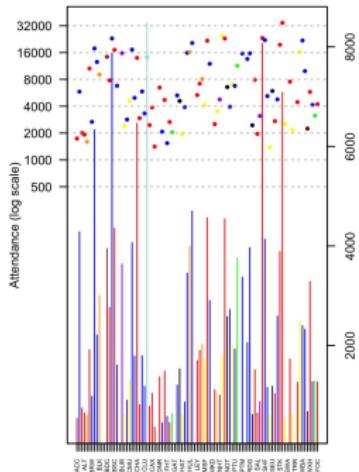


Data coverage

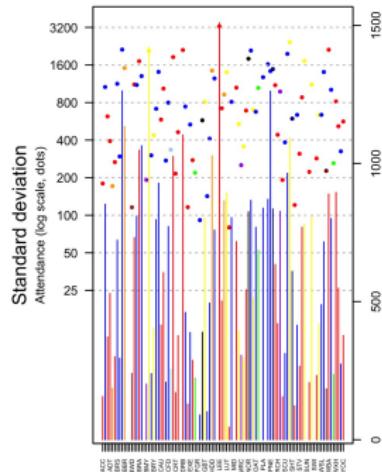


Data variation

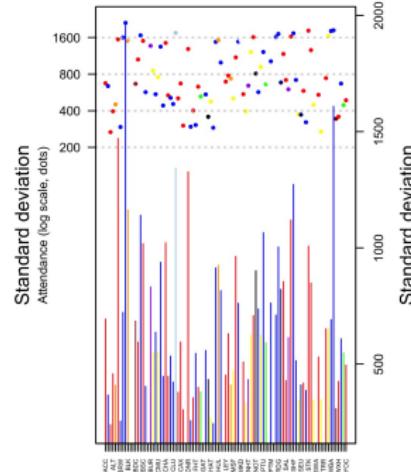
Home attendances



Away by club



Away section by home club

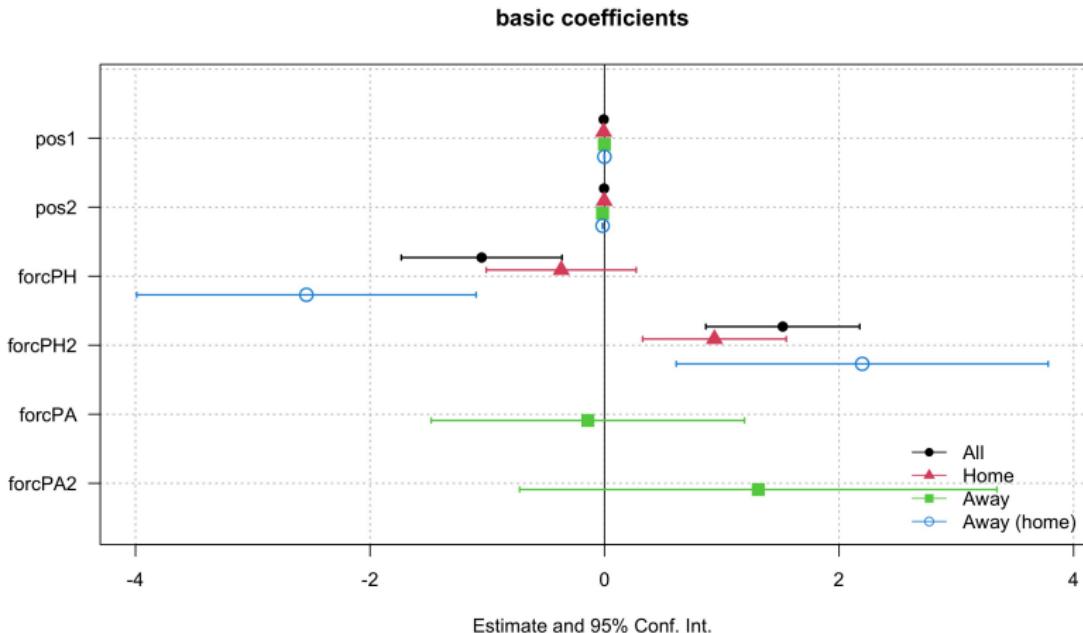


Dots = mean attendance, lines = standard deviation

Dots = mean attendance, lines = standard deviation

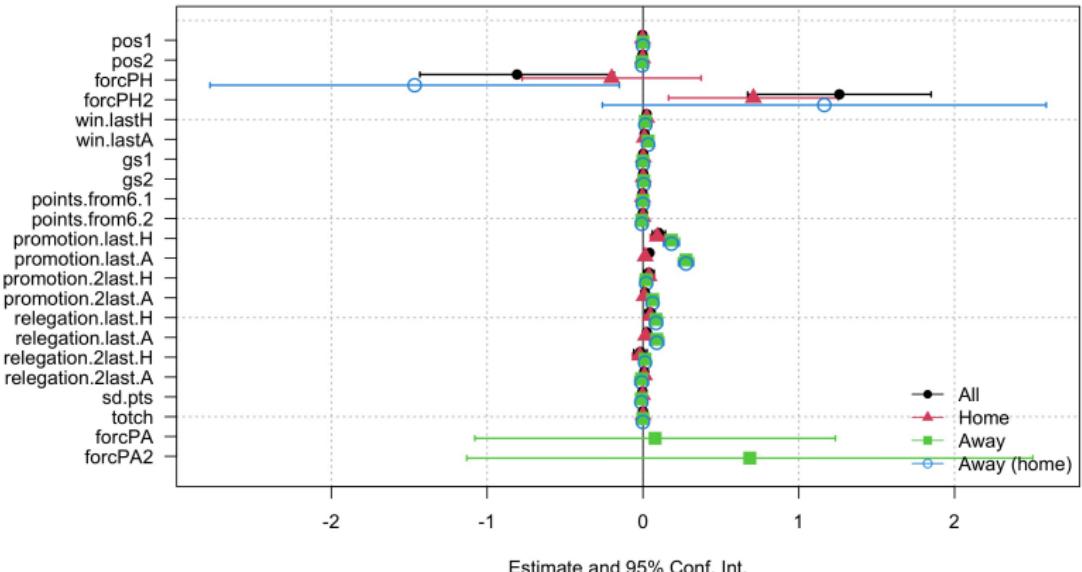
Dots = mean attendance, lines = standard deviation

Coefficient plot: Basic

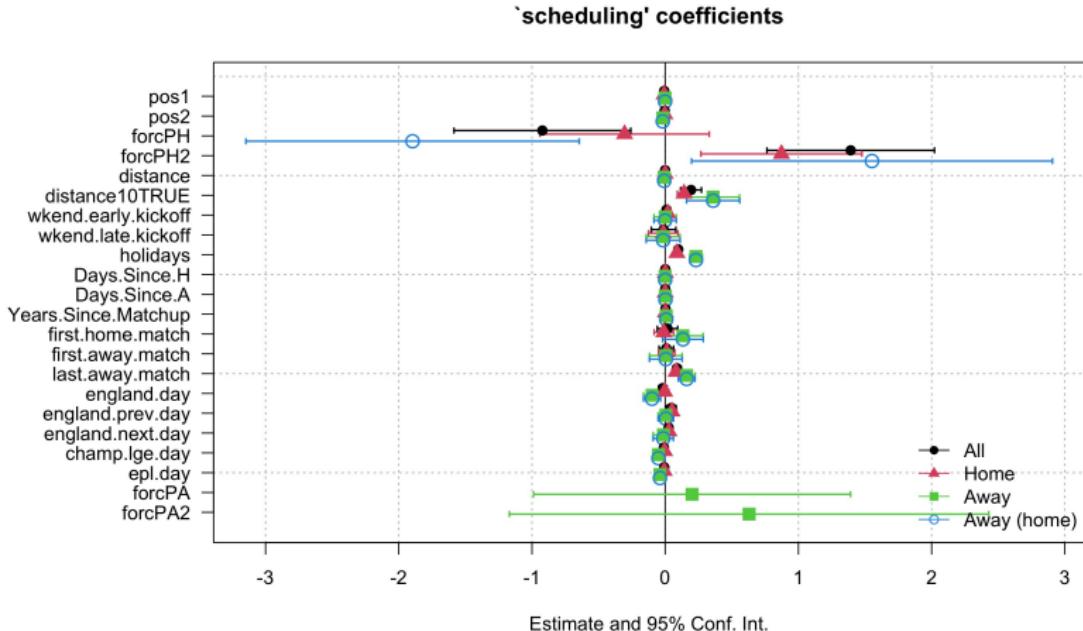


Coefficient plot: Football

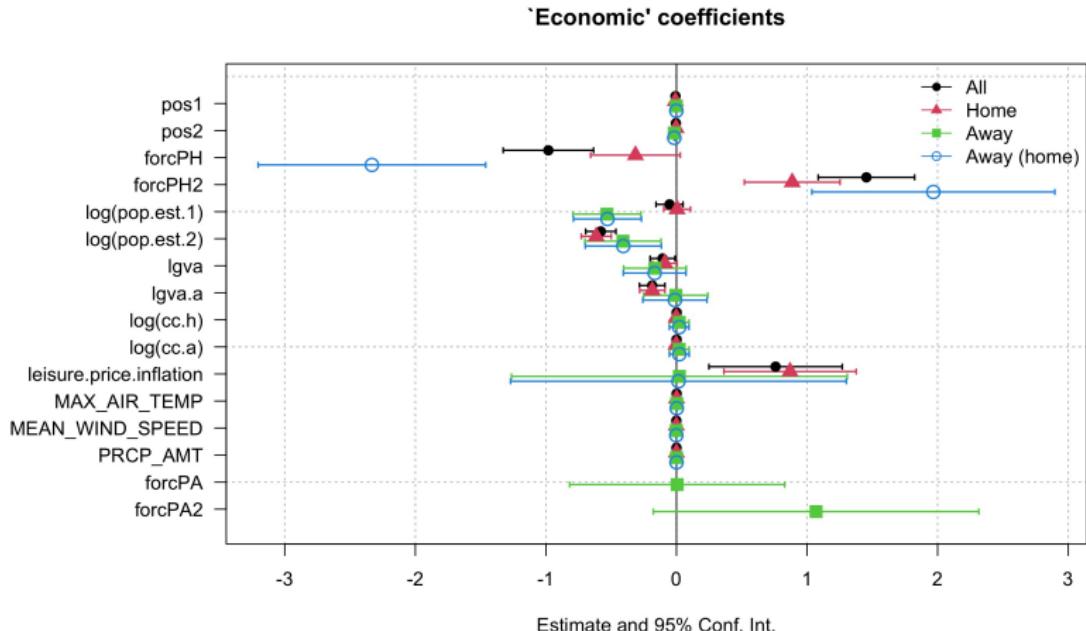
'football' coefficients



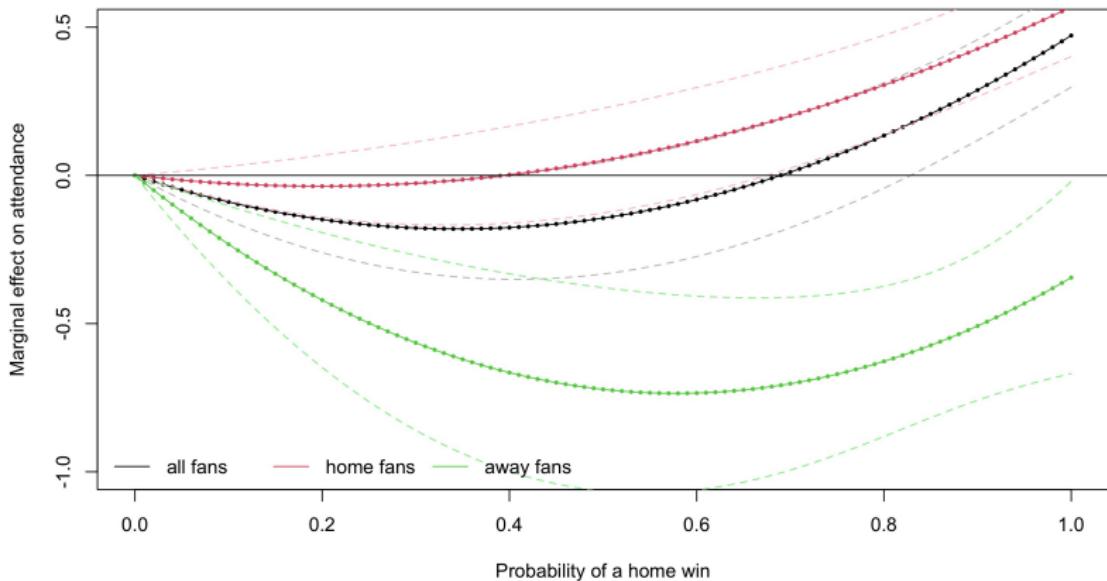
Coefficient plot: Scheduling



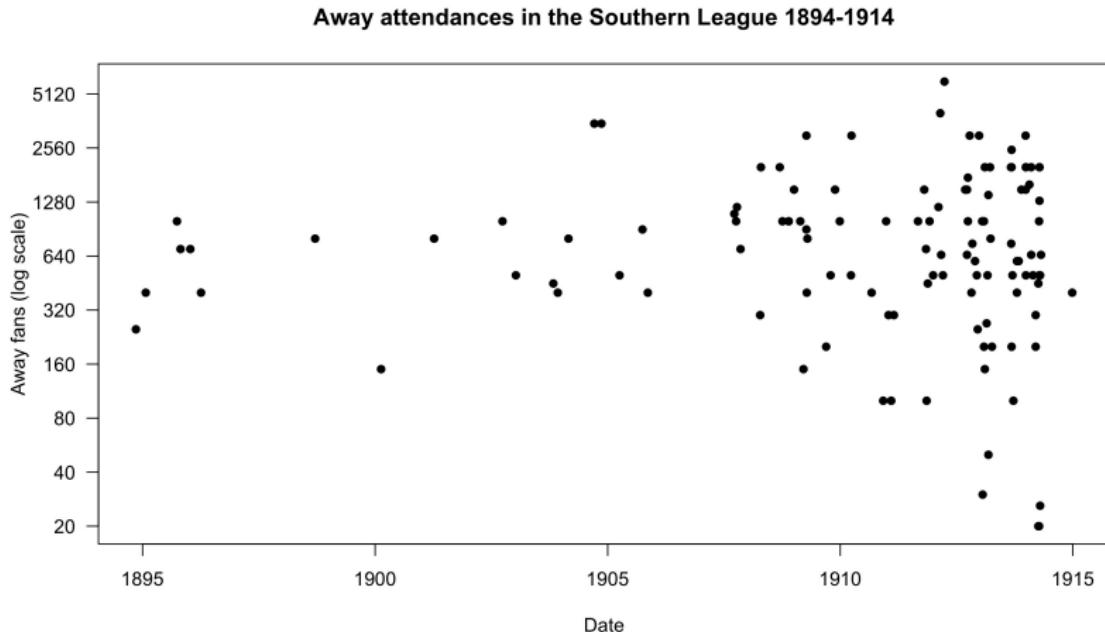
Coefficient plot: Economic/Demographic



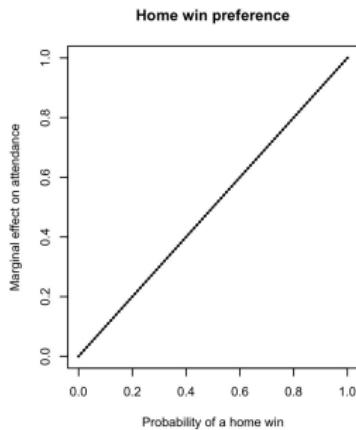
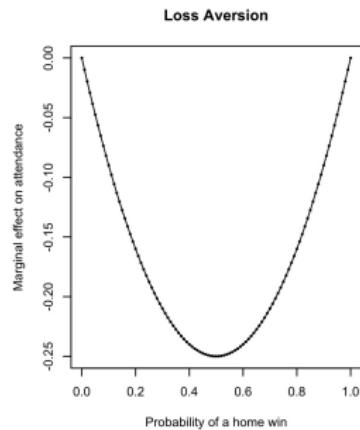
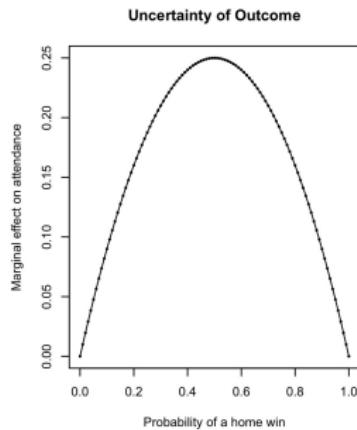
Coefficient plot: Quadratic relationship



Historic numbers (source: Roger Titford)



Fan demand concepts

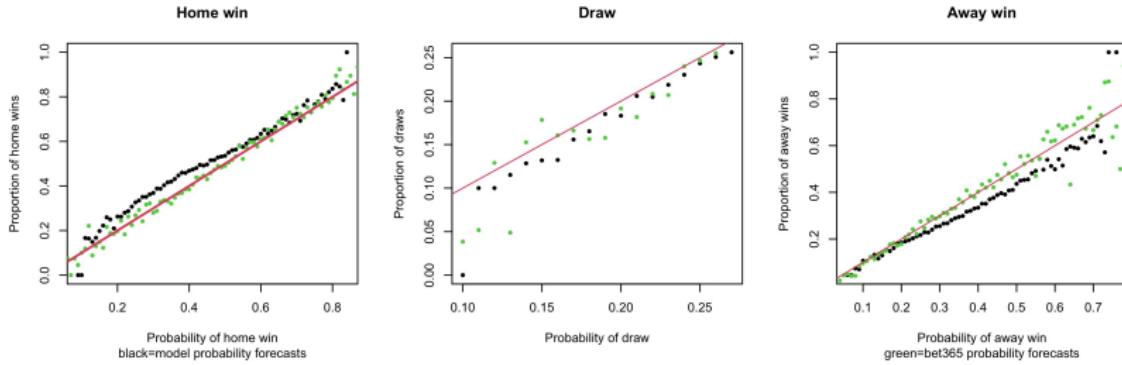


Probability model

- Probability of match outcomes needed.
 - ▶ Could use bookmaker prices (e.g. Elaad et al., 2020).
- We create match-outcome probability (Goddard, 2005):
 - ▶ Based on Elo ratings.
 - ▶ Add in league information: goal difference, points per game, form.
- Ordered logit model: delivers probabilities for all three outcomes.
- Resulting model forecasts online here:
<https://match-and-season-predictions.streamlit.app/>.

Any good?

Probability model



Oldham Athletic AFC — Boundary Park



Oldham Athletic AFC — Wembley



Oldham Athletic AFC — Wembley

