Research Report

Department of Statistics

Stockholms universitet

No. 2018:3

Bayesian Change-point Modelling of the Effects of 3-points-for-a-win Rule in Football

Gebrenegus Ghilagaber
Parfait Munezero

Department of Statistics, Stockholm University, SE-106 91 Stockholm, Sweden

Bayesian Change-point Modelling of the Effects of 3-points-for-a-win Rule in Football

Gebrenegus Ghilagaber* and Parfait Munezero†

June 6, 2018

Abstract

We examine the effects of the 3-point-for-a-win (3pfaw) rule in the football (soccer) world. Data on mean goals and proportions of decided matches from seven leagues around the world form the basis of our analyses. Bayesian change-point analysis shows that the rule had no effects on the mean goals in any of the leagues but, indeed, had increasing effects on the proportions of decided matches in most of the leagues studied. This, in turn, implies that while the rule has given teams the incentive to aim at winning matches, such aim was achieved not by scoring excess goals. Instead, it was achieved by scoring enough goals in order to win and, at the same, defending enough in order not to lose.

Key words: Bayesian inference; Change-point models; Football/Soccer; Three-points-for-a-win rule; Albania, Brazil, England, Germany, Poland, Romania, Scotland.

*Department of Statistics, Stockholm University, Sweden. Gebre@stat.su.se
†Department of Statistics, Stockholm University, Sweden. Parfait.Munezero@stat.su.se

1 Introduction

1.1 Background

Prior to 1981, the point-awarding system in football was 2 points for a win, 1 point for a draw, and 0 point for a loss. A 3-points for a win (3pfaw) system was introduced first in England in 1981 with the aim to provide teams with incentive to aim at winning games. This, in turn, was expected to lead to more offensive and, hence, more exciting football that would increase public attendance. The reform which started in England was recognized by the international governing body of football (FIFA) in 1994 when it was adopted in the Football World Cup and other leagues worldwide.

The main objective of the 3pfaw system, encouraging teams to aim at winning matches, was expected to be reflected in either increased number of goals per match, increased number of decided matches (decreased draws), or both.

Some researchers have attempted to measure the effect of the reform in various ways and with diverging results. For instance, Brocas and Carrillo (2004) analyze the dynamics of the game strategies of teams using game theory and observe that, under the 3pfaw rule, teams tend to play more defensively rather than playing offensive football. Dilger (2009), on the other hand, applies regression model on data from the German league and concludes that the introduction of the 3pfaw rule has significantly increased the mean goals as well as the proportion of decided matches. Such observation is supported by Moschini (2010) who uses a game-theoretic model to investigate the effects of the 3pfaw rules in 35 different countries and concludes that the rule has led to statistically significant increase in the number of expected goals and decrease in the number of drawn matches. In contrast, Hon et al. (2014) use regression discontinuity design on data from the German league and find no evidence that 3pfaw makes the games more decisive, increases the number of goals, or decreases goal differences.

1.2 Excitement Index Approach

Aylott and Aylott (2007) define an Excitement Index (EI) by combining points allocated to decided matches and to each goal scored. A decided match is allocated 1 point while a match that ends in a draw is allocated 0 point. Further, each goal scored in a match is allocated 0.2 points and the EI for each year and country is defined as a sum of these points. The allocation of points to decided matches and scored goals is in accordance with the intended aim of the 3pfaw rule (increase the proportion of decided matches and the number of goals). They then plot the EI-values for each of the seven leagues and years (from about 6 years before and about 7 years after the 3pfaw rule). Based on the trends in the EI for each league (and another plot of aggregated effect of the 3pfaw on EI in all seven leagues), Aylott and Aylott (2007) conclude that there was an increase in excitement after the introduction of 3pfaw rule but that it took some years for the rule to take full effect.

A drawback in the above approach is that no formal test is made to see if the differences between the Excitement Indexes before and after the introduction of the 3pfaw are statistically significant. Moreover, the definition of Excitement Index is in itself questionable and subjective. For instance, a match that ends with $3-0$ win for the home-team will have an Excitement Index of 1.6 but a match that ends with $0-3$ win for the away-team or even in a $4-4$ draw will also have the same value ($EI = 1.6$). But, this doesn't imply that these three matches are equally exciting as this is subjective and depends on the observer's 'taste', whether she or he is a neutral observer or supports the home- or away-team, etc.

1.3 The current study

The drawback outlined above calls for a statistically well grounded, theoretically appropriate, and empirically evident alternative to measure the effects of the 3pfaw rule. In the present paper, we use Bayesian change-point modelling (Raftery and Akman, 1986) to re-analyze the data in Aylott and Aylott (2007). We use Bayes Factors and posterior probabilities as evidence in support for or against a no-change model. Our empirical results do not lend support to changes in the mean goals after the introduction of the 3pfaw rule in any of the leagues studied. On the other hand, we find strong evidence for increase in the proportions of decided matches after the 3pfaw rule in at least four of the seven leagues studied. Further, our results show that the effects of the 3pfaw rule on the proportion of decided matches was immediate in some of the leagues and didn't take long time to take full effect. This, in turn, implies that the 3pfaw rule has encouraged teams to aim more at winning matches (even marginally) than at scoring excess goals. In other words, the 3pfaw rule which was expected to make football more offensive has, in fact, made it both offensive and defensive.

In Section 2, we introduce the data sets and provide results from preliminary analyses using standard statistical methods. In Section 3, we describe the Bayesian change-point model in general and how it can be applied to the data sets at hand. In Section 4, we present empirical findings from fitting the Bayesian change-point model to mean goals and proportions of decided matches. Section 5 summarizes the findings of the paper by way of concluding remarks and suggestions for future work.

2 The Data Sets

The data sets that form the basis for this study come from seven leagues around the world - Albania, Brazil, England, Germany, Poland, Romania, and Scotland where the 3pfaw rule was introduced in 1996, 1994, 1981, 1995, 1995, 1994, and 1994, respectively (Aylott and Aylott, 2007). The data for each league contains information on the number of games, mean goals per game, and proportion of decided matches in each of the six years before introduction of the 3pfaw rule and seven years after the introduction of the rule. According to Aylott and Aylott (2007), these seven leagues were maximally heterogeneous in different aspects like geographic location, history, wealth, country size, population size, and the time of introduction of the 3pfaw rule and were selected in order to avoid or minimize any kind of correlations across the selected cases.

We begin our analyses by conducting t-tests for equality between mean goals before and after the 3pfaw and, separately, tests of association between period (before or after the 3pfaw rule) and match-outcomes (drawn or decided). The relevant data are obtained by aggregating data from the 6 years before the 3pfaw rule and those from seven years after the 3pfaw rule.

2.1 Mean goals per match

Table 1 contains mean goals from the seven leagues under study over a period of about 6 years prior to the introduction of the 3pfaw (indicated by '-' in the first column) and about 7 years after (indicated by '+' in the first column). See also Figure 1 in the Appendix for a plot of the mean goals.

In Table 2 we present results from t-tests for equality of population means across the seven leagues.

Table 1: Mean goals in 7 football leagues within about 6 years before and about 7 years after the introduction of the 3pfaw rule.

Years	**Albania**	**Brazil**	**England**	**Germany**	**Poland**	**Romania**	**Scotland**
−5.5	2.42	2.01	2.57	2.79	2.02	3.21	2.69
−4.5	2.63	1.89	2.64	2.62	2.41	2.92	2.73
−3.5	2.42	2.21	2.60	2.58	2.22	2.76	2.67
−2.5	2.27	2.29	2.56	2.62	2.61	2.49	2.71
−1.5	2.26	2.53	2.57	2.54	2.34	2.63	2.80
−0.5	2.19	2.40	2.47	2.63	2.40	2.44	2.51
+0.5	2.23	2.50	2.60	2.70	2.53	2.94	2.65
+1.5	2.28	2.72	2.77	2.64	2.50	2.63	2.69
+2.5	2.56	2.78	2.75	2.61	2.39	2.85	2.71
+3.5	2.58	2.88	2.72	2.61	2.53	2.91	2.80
+4.5	2.80	2.78	2.80	2.76	2.81	2.90	2.74
+5.5	2.81	2.89	2.60	2.86	2.56	2.98	2.81
+6.5	3.04	2.86	2.62	3.02	2.51	2.61	2.71

Source: Aylott and Aylott (2007)

Table 2: Results from tests of equality between mean goals before and after the 3pfaw

	Albania	**Brazil**	**England**	**Germany**	**Poland**	**Romania**	**Scotland**
Mean (before)	2.37	2.22	2.58	2.62	2.32	2.73	2.68
Mean (after)	2.61	2.79	2.69	2.75	2.54	2.83	2.73
Var (before)	0.02	0.05	0.002	0.01	0.04	0.09	0.01
Var (after)	0.08	0.02	0.008	0.026	0.016	0.026	0.002
T-values	**1.94**	**3.28**	**2.35**	1.72	**2.28**	0.84	1.19
$T_{(0.025,11)}$	2.20	2.20	2.20	2.20	2.20	2.20	2.20
$T_{(0.05,11)}$	1.80	1.80	1.80	1.80	1.80	1.80	1.80

The results indicate that a two-sided test at 5 % level shows significant differences between the mean goals for Brazil, England, and Poland. In a one-sided test at 5 % level of significance (with the prior expectation that the rule increases mean goals), even Albania shows significant differences between mean goals (in addition to Brazil, England, and Poland). The above results provide some general picture of the phenomenon under investigation. However, it is to be recalled that they are obtained by aggregating the mean goals over the 6 years before the 3pfaw rule and comparing it with another mean that is obtained by aggregating the mean goals over the seven years after the 3pfaw rule.

2.2 Proportions of decided matches

Table 3 contains proportions of decided matches from the seven leagues under study over the same period as for the mean goals in the previous sub-section (see also Figure 2 in the Appendix).

Table 3: Proportions of decided matches in 7 football leagues within about 6 years before and 7 years after the 3pfaw rule.

Years	**Albania**	**Brazil**	**England**	**Germany**	**Poland**	**Romania**	**Scotland**
−5.5	0.749	0.662	0.723	0.697	0.625	0.863	0.712
−4.5	0.777	0.637	0.721	0.676	0.679	0.830	0.707
−3.5	0.774	0.648	0.691	0.656	0.644	0.797	0.734
−2.5	0.725	0.690	0.704	0.700	0.724	0.778	0.737
−1.5	0.723	0.705	0.721	0.704	0.654	0.784	0.740
−0.5	0.788	0.719	0.720	0.683	0.654	0.791	0.689
+0.5	0.761	0.732	0.726	0.703	0.745	0.817	0.739
+1.5	0.834	0.786	0.737	0.717	0.742	0.850	0.744
+2.5	0.850	0.695	0.740	0.686	0.742	0.824	0.744
+3.5	0.821	0.743	0.757	0.699	0.750	0.837	0.738
+4.5	0.794	0.775	0.754	0.706	0.738	0.843	0.761
+5.5	0.862	0.743	0.726	0.719	0.771	0.853	0.751
+6.5	0.809	0.770	0.727	0.757	0.714	0.754	0.747

Source: Aylott and Aylott (2007)

In Table 4 we present results from tests for association between period (before or after the rule) and match-outcomes (drawn or decided). The frequencies are obtained by adding corresponding values in Table 2 of Aylott and Aylott (2007) over the years before the 3pfaw rule and those after the rule. The results in Table 4 show that there are strong associations between the periods and the match-outcomes. This is true for all the leagues except Romania where the p-values associated with the χ^2 is 0.10 though even the p-value for the χ^2 of Germany is also not as small as for the other countries.

Table 4: Results from tests of associations between the 3pfaw rule and match outcomes

	Albania				
	Drawn	**Decided**	**Total**	**% decided**	χ^2
Before	580	18837	2417	76	
After	432	1965	2397	82	25.87 ($p = 0.000$)
Total	1012	3802	4814	79	

	Brazil				
	Drawn	**Decided**	**Total**	**% decided**	χ^2
Before	438	937	1375	68	
After	519	1543	2062	75	18.35 ($p = 0.000$)
Total	957	2480	3437	72	

	England				
	Drawn	**Decided**	**Total**	**% decided**	χ^2
Before	3488	8680	12168	71	
After	3718	10480	14198	74	20.27 ($p = 0.000$)
Total	6720	19160	26366	73	

	Germany				
	Drawn	**Decided**	**Total**	**% decided**	χ^2
Before	1367	3011	4378	69	
After	1232	3052	4284	71	6.27 ($p = 0.012$)
Total	2599	6063	8662	70	

	Poland				
	Drawn	**Decided**	**Total**	**% decided**	χ^2
Before	572	1130	1702	66	
After	482	1396	1878	74	27.11 ($p = 0.000$)
Total	1054	2526	3580	71	

	Romania				
	Drawn	**Decided**	**Total**	**% decided**	χ^2
Before	354	1482	1836	81	
After	558	1718	2076	83	2.70 ($p = 0.100$)
Total	712	3200	3912	82	

	Scotland				
	Drawn	**Decided**	**Total**	**% decided**	χ^2
Before	1283	3298	4581	72	
After	1291	3797	5088	75	8.56 ($p = 0.003$)
Total	2574	7095	9669	73	

3 Bayesian change-point models

A change-point model assumes that a sequence of random variables $y_1, \cdots, y_T$ follows a distribution F before a change-point, say k, and follows another distribution G after the change-point (Smith, 1975). The distributions F and G need not be the same and they may be known or unknown. The change-point parameter, k, is unknown and is assumed to take on values in the set $\kappa = \{1, 2, ..., T\}$ where T is the length of the observation period. Evidence about whether a change occurred during the period of study is provided by testing a null hypothesis $H_0 : k = T$ (no change) against $H_1 : k < T$.

Change-point models have been applied in different areas like economics (Western and Kleykamp, 2004), wage growth (Dominicus et al., 2008), and cognitive science (Hall et al., 2003). Raftery and Akman (1986) derive the posterior for the change-point model in Poisson process and use Bayes Factors to compare this model with a model that assumes a constant rate. Carlin et al. (1992) develop Bayesian hierarchical models for change-point problems while Smith (1975) considers cases of the change-point in the binomial and normal distributions. Lee (1998) investigates, in a Bayesian framework, change-points in exponential family of distributions. Carlin et al. (1992) and Chin Choy and Broemeling (1980) study the problem of a single change-point in linear models using the Bayesian framework while Bauwens and Rombouts (2012) and Van den Hout et al. (2013) study a more general case of multiple change-points in linear models. Petrone and Raftery (1997) investigate change point problems using a non-parametric Bayesian approach.

Despite the abundance of literature in the area of change-point models no previous study (to the best of our knowledge) has applied change-point modelling to assess the effect of the 3pfaw rule. In the next two sub-sections we apply Bayesian change-point modelling to analyze the mean goals and the proportions of decided matches in the data previously analyzed by Aylott and Aylott (2007).

3.1 Bayesian change-point model for mean goals

We consider the case where F and G are known and our interest is in inference about the parameters in the models as well as testing whether a change-point model is adequate for the data at hand. We assume that the mean goals per match, Y, follow a normal distribution with mean μ_1 and variance σ_1^2 until the k^{th} year. After the k^{th} year, the mean goals per match follow a normal distribution with mean μ_2 and variance σ_2^2 :

$$\begin{array}{ll} \mathbf{Y_t} \sim N\left(\mu_1, \sigma_1^2\right) & t = 1, ..., k \\ \mathbf{Y}_t \sim N\left(\mu_2, \sigma_2^2\right) & t = k+1, .., T \end{array} \tag{1}$$

Further, following Gill (2008), we assume Inverse gamma (IG) conjugate prior for the variances and conjugate normal prior for the means (which, in turn, are conditional on the variances):

$$p\left(\sigma_j^2 | \alpha_j, \beta_j\right) \propto \left(\sigma_j^2\right)^{-(\alpha_j - 1)} \exp\left(-\frac{1}{\sigma_j^2}\beta_j\right), \ j = 1, 2 \tag{2}$$

$$p\left(\mu_j | \sigma_j^2 / s_j, m_j\right) \propto \left(\sigma_j^2\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma_j^2 / s_j}\left(\mu_j - m_j\right)^2\right), \ j = 1, 2 \tag{3}$$

where the prior for μ_j is explicitly conditional on σ_j^2 and m_j is prior mean. The parameter s_j is the so called 'confidence parameter' that measures strength of belief in the prior specification (Gill, 2008).

The change-point, k, is assumed to be discrete uniform over $\{1, 2, ..., 13\}$ a priori since there are $T = 13$ data points (6 years before the introduction of the rule and 7 years after its introduction). However, we set the density at the point $k = T$ to zero because it represents a model without a change-point:

$$f(k) = \begin{cases} \frac{1}{T-1} & k < T \\ 0 & k = T \end{cases}$$

Further, μ_1, μ_2, and k are assumed to be independent of each other. The resulting joint posterior is then

$$p\left(\mu_j, \sigma_j^2, k | y\right) = p\left(y | \mu_j, \sigma_j^2\right) p\left(\sigma_j^2 | \alpha_j, \beta_j\right) p\left(\mu_j | \sigma_j^2, s_j, m_j\right) p\left(k\right) \tag{4}$$

Following the procedure in Gill (2008, Section 3.4), it can be shown that the conditional posterior distributions for σ_j^2 , $j = 1, 2$ are given by

$$\sigma_1^2 | y, k \sim IG\left(\alpha_1 + \frac{k}{2} - \frac{1}{2}, \ \beta_1 + \frac{1}{2}\sum_{t=1}^{k}\left(y_t - \bar{y_1}\right)^2\right) \tag{5}$$

$$\sigma_2^2 | y, k \sim IG\left(\alpha_2 + \frac{T-k}{2} - \frac{1}{2}, \ \beta_2 + \frac{1}{2}\sum_{t=k+1}^{T}\left(y_t - \bar{y_2}\right)^2\right), \tag{6}$$

and the conditional posterior distributions for the means $\mu_j,\ j = 1,2$ are given by

$$\mu_1|y,\sigma_1,k \sim N\left(\frac{m_1 s_1 + k\bar{y_1}}{k+s_1},\ \frac{\sigma_1^2}{k+s_1}\right) \tag{7}$$

$$\mu_2|y,\sigma_2,k \sim N\left(\frac{(T-k)\bar{y_2} + m_2 s_2}{T-k+s_2},\ \frac{\sigma_2^2}{T-k+s_2}\right), \tag{8}$$

where $\bar{y_1} = \frac{\sum_{t=1}^{k} y_t}{k}$ and $\bar{y_2} = \frac{\sum_{t=k+1}^{T} y_t}{T-k}$.

Finally the conditional posterior distribution of the change-point parameter is

$$p\left(k|y,\mu_1,\mu_{2,}\sigma_1^2,\sigma_2^2\right) = \frac{L\left(y;\ k,\ \ \mu_1,\mu_2\right)}{\sum_{k=1}^{T} L\left(y;\ k,\ \ \mu_1,\mu_2\right)} \tag{9}$$

where

$$L\left(y;\ k,\ \ \mu_1,\mu_2\right) \propto \left(\sigma_1^2\right)^{-\frac{k}{2}} \left(\sigma_2^2\right)^{-\frac{T-k}{2}} \exp\left\{ \begin{array}{c} -\frac{1}{2\sigma_1^2}\sum_{t=1}^{k}\left(y_t-\mu_1\right)^2 - \\ \frac{1}{2\sigma_2^2}\sum_{t=k+1}^{T}\left(y_t-\mu_2\right)^2 \end{array} \right\} \tag{10}$$

Thus, the problem reduces to sampling from the joint posterior distributions above. We use the Gibbs sampling algorithm that samples iteratively from the above conditional posterior distributions. We simulate 12000 Gibbs sampling replications where the first 2000 are treated as burn-in and discarded from the analyses. The following values were used in the prior distributions while modelling mean goals: $\alpha_1 = \alpha_2 = 2, \beta_1 = \beta_2 = 1, s_1 = s_2 = 1, m_1 = m_2 = 2$.

3.2 Bayesian change-point model for proportions of decided matches

Let d_t denote the number of decided matches out of a total of N_t matches in year t ($t = 1, ...13$). Our model assumes that d_t follows a binomial distribution with parameters N_t and θ_t until the k^{th} year. After the k^{th} year, d_t follows a binomial distribution with parameters N_t and λ_t:

$$\begin{array}{ll} d_t \sim Bin\left(N_t,\theta_t\right) & t = 1,...,k \\ d_t \sim Bin\left(N_t,\lambda_t\right) & t = k+1,..,T, \end{array} \tag{11}$$

where θ_t and λ_t are the probabilities that a randomly selected match in year t is decided (not drawn).

Further, we assume conjugate beta priors for the parameters θ and λ which is a special case of the prior in Smith (1975) and that θ, λ, and k are independent of each other:

$$\begin{aligned} \theta &\sim Beta\left(a_1, b_1\right) \\ \lambda &\sim Beta\left(a_2, b_2\right) \\ f(k) &= \begin{cases} \frac{1}{T-1} & k < T \\ 0 & k = T \end{cases} \end{aligned} \tag{12}$$

where a_i and b_i are hyper-parameters.

The joint posterior of the parameters is given by

$$p\left(\theta, \lambda, k | d_1, ..., d_T\right) \propto p\left(d_1, ..., d_T | \theta, \lambda, k\right) * p\left(\theta | a_1, b_1\right) * p\left(\lambda | a_2, b_2\right) * p\left(k\right) \tag{13}$$

Following the procedure in Smith (1975), the conditional posterior densities for each parameter can easily be derived as follows:

$$\theta | \lambda, k, d_1, ..., d_T \sim Beta\left(a_1 + \sum_{t=1}^{k} d_t,\ b_1 + \sum_{t=1}^{k} (N_t - d_t)\right) \tag{14}$$

$$\lambda | \theta, k, d_1, ..., d_T \sim Beta\left(a_2 + \sum_{t=k+1}^{T} d_t,\ b_2 + \sum_{t=k+1}^{T} (N_t - d_t)\right) \tag{15}$$

$$p\left(k | d, \theta, \lambda\right) = \frac{L\left(d, k, \theta, \lambda\right)}{\sum_{k=1}^{k=T} L\left(d, k, \theta, \lambda\right)}, \tag{16}$$

where

$$L\left(d, k, \theta, \lambda\right) = \theta^{\sum_{t=1}^{k} d_t} \left(1 - \theta\right)^{\sum_{t=1}^{k} (n_t - d_t)} \lambda^{\sum_{t=k+1}^{T} d_t} \left(1 - \lambda\right)^{\sum_{t=k+1}^{T} (n_t - d_t)}. \tag{17}$$

Again, the problem reduces to sampling from the joint posterior distributions above and we use Gibbs sampling algorithm that samples iteratively from the above conditional posterior distributions. We simulate 12000 Gibbs sampling iterations where the first 2000 are treated as burn-in and discarded from the analyses. The following values are used in the prior distributions while modelling proportions of decided matches: $a_1 = a_2 = 1, b_1 = b_2 = 1$.

3.3 Model Comparisons using Bayes Factors

Bayes Factors provide a magnitude of the evidence contained in the data that is in favour of one model, say Model 2 (in our case a change-point model) over another model, say Model 1 (in our case no change-point model). The general form of Bayes Factor (with models 1 and 2 denoted by M_1 and M_2, respectively) is

$$BF_{21} = \frac{p\left(D|M_2\right)}{p\left(D|M_1\right)} = \frac{p\left(M_2|D\right)/p\left(M_2\right)}{p\left(M_1|D\right)/p\left(M_1\right)} = \frac{p\left(M_2|D\right)/p\left(M_1|D\right)}{p\left(M_2\right)/p\left(M_1\right)} = \frac{\text{Posterior odds}}{\text{Prior odds}}, \quad (18)$$

where D denotes data. From the above equation it is clear that smaller values of Bayes Factor indicate that the data support Model 1 (that no change has taken place) while larger values of Bayes Factor indicate that the data support Model 2 (that a change has taken place). Kass and Raftery (1995) provide the following guidelines in comparing models:

$2 * \ln\left(BF_{21}\right) < 0$	:	Evidence in favour of Model 1
$0 < 2 * \ln\left(BF_{21}\right) \leq 2$	:	Insignificant evidence against Model 1
$2 < 2 * \ln\left(BF_{21}\right) \leq 6$	:	Substantial evidence against Model 1
$6 < 2 * \ln\left(BF_{21}\right) \leq 10$	:	Strong evidence against Model 1
$2 * \ln\left(BF_{21}\right) > 10$	:	Decissive evidence against Model 1

For more on Bayes Factors, see for instance Gill (2008), Kass and Raftery (1995), and Raftery (1996).

4 Empirical Findings

4.1 Results from analyses of mean goals

Plots of posterior distributions of the mean goals for the seven leagues, the change-point k, and differences between the means are presented in Figure 3 (see Appendix). We observe that the plots of the posterior distributions of mean goals before and after the 3pfaw rule overlap in all the seven leagues (panels a-g in Figure 3). Further, the posterior distributions of the change-points k for all leagues (panel h) are almost flat indicating the posterior distributions of k resemble their corresponding priors (discrete uniform) - which is an indication of no change within the study period. Lastly, the differences between the posterior means after and before the 3pfaw (panel

i) are centered around zero for all leagues - again indicating no differences between the corresponding mean-goals before and after the 3pfaw rule. In Table 5, we summarize posterior estimates of the parameters of interest across the seven leagues (with D, again, denoting data).

Table 5: Posterior estimates from the analysis of mean goals.

	Albania	**Brazil**	**England**	**Germany**	**Poland**	**Romania**	**Scotland**
$E(k\|D)$	9	7.85	9.55	9.5	9.16	9.61	9.73
$E(k\|D) - 6.5$	2.5	1.35	3.15	3.00	2.66	3.11	3.23
$E(\mu_1\|D)$	2.24	2.16	2.46	2.48	2.25	2.60	2.53
$E(\mu_2\|D)$	2.60	2.66	2.47	2.61	2.42	2.55	3.52

The results in Figure 3 indicate lack of evidence in support of the change-point model. But, to avoid subjectivity we compute the corresponding Bayes Factors, as shown in Table 6, for comparing a change-point model in mean goals against a no-change model. The posterior probabilities are the proportions of times (out of the 10000 replications) where a mean goal after the 3pfaw rule, μ_2, is greater than its counterpart before the 3pfaw rule, μ_1. Note, however, these do not tell the magnitude of difference between the mean goals and even a difference in the 10^{th} decimal may contribute to the proportion.

Table 6: Posterior probabilities and Bayes Factors for changes in mean goals

Country	$\mathbf{p}(\mu_2 > \mu_1)$	$\mathbf{2 * \ln}(BF_{21})$	**Evidence**
Albania	0.83	−0.04	Model 1 favoured (no-change)
Brazil	0.90	0.55	Insignificant evidence (no-change)
England	0.54	−0.52	Model 1 favoured (no-change)
Germany	0.65	−0.44	Model 1 favoured (no-change)
Poland	0.69	−0.43	Model 1 favoured (no-change)
Romania	0.49	−0.43	Model 1 favoured (no-change)
Scotland	0.51	−0.57	Model 1 favoured (no-change)

Thus, on the basis of the guidelines for Bayes Factors presented in Section 3.3, we can argue that the 3pfaw rule didn't have any effect on the mean goals in any of the seven leagues studied. The results here differ from those in our preliminary analyses using t-tests to compare means and where we got significant difference in at least three countries. But, as already pointed out there, the change-point model utilizes the data more efficiently (as it doesn't require aggregation) and, hence, we will base our conclusions on the results in this section.

4.2 Results from analyses of proportions of decided matches

Plots of posterior distributions of the proportions of decided matches for the seven leagues, the change-point k, and differences between the proportions of decided matches are presented in Figure 4 in the Appendix. We observe that the plots of the posterior distributions of proportions of decided matches before and after the 3pfaw rule are far apart for some leagues (Albania, Brazil, England, Poland) while they overlap in some of the countries (Romania, for instance). Further, the posterior distributions of the change-points k (panel h) are flat for one-two leagues (resembling their corresponding priors) but have peaks in most of the other leagues (indicating they are well-updated by the data).

More importantly, the differences between the posterior proportions of decides matches after and before the 3pfaw (panel i) are centered to the right of zero for most of the leagues indicating higher proportions of decided matches after the 3pfaw rule. We summarize the posterior estimates of the parameters in Table 7:

Table 7: Posterior estimates from the analysis of proportions of decided matches.

	Albania	**Brazil**	**England**	**Germany**	**Poland**	**Romania**	**Scotland**
$E(k\|D)$	7	6.5	7	11	6	11	7.5
$E(k\|D) - 6.5$	0.5	0.0	0.5	4.5	−0.5	4.5	1
$E(\theta\|D)$	0.76	0.68	0.71	0.69	0.66	0.82	0.72
$E(\lambda\|D)$	0.83	0.75	0.74	0.74	0.74	0.77	0.75

Comparing models of no-change point with a change-point model with regard to the proportions of decided matches, we get the posterior probabilities and Bayes Factors presented in Table 8.

Again, on the basis of the guidelines for Bayes Factors, we can argue that the 3pfaw rule did, indeed, have a strong effect in increasing the proportions of decided matches in at least four of the leagues studied (Albania, Brazil, England, and Poland). There is also marginal (insignificant) evidence of its effect in Germany but it has not affected at all the outcomes in the Romanian and Scottish leagues. These results based on Bayes Factors also seem to be consistent with those obtained in our preliminary analyses using Chi-square tests of association using aggregated data.

Table 8: Posterior probabilities and Bayes Factors for changes in proportions of decided matches.

Country	$\mathbf{p}\left(\lambda > \theta\right)$	$\mathbf{2} * \mathbf{ln}\left(BF_{21}\right)$	**Evidence**
Albania	1	11.08	Model 2 favoured decisively (change)
Brazil	1	3.63	Model 2 favoured substantially (change)
England	1	4.26	Model 2 favoured substantially (change)
Germany	1	0.30	Insignificant evidence (no-change)
Poland	1	7.52	Model 2 favoured strongly (change)
Romania	0.17	−1.60	Model 1 favoured (no-change)
Scotand	0.99	−1.04	Model 1 favoured (no-change)

Finally, Figure 5 in the Appendix presents posterior distributions of the change-points, k, for the seven leagues, separately. One can easily see that while the curves corresponding to means goals are flat (indicating the data do not update the discrete uniform prior) the curves corresponding to proportions of decided matches have peaks at various points within the study period - thereby supporting the change-point model.

5 Summary and concluding remarks

We have proposed and applied a Bayesian change-point model to measure the effects of the 3-points-for-a-win rule in the football world. Our model is statistically well-grounded and, more importantly, much less subjective compared to previous approaches used to analyze the same data sets. One of the strengths in our approach is that we let the time for change, if any, to be part of the model and estimate it together with other parameters of interest.

Our results do not support a change-point model for mean goals in any of the leagues studied. On the other hand, the results provide strong evidence for a change-point model for the proportion of decided matches in most of the leagues studied.

This, in turn, implies that the 3pfaw rule has given teams the incentive to win matches (rather than draw). However, this was achieved not by scoring 'too many' goals but, rather, by scoring enough goals in order to win and, at the same time, defending enough in order not to lose.

For instance, a hypothetical match that could have ended in a 2-2 draw before the 3pfaw rule might end in a 1-0 or 2-1 win after the 3pfaw rule. Such a match contributes towards increasing the proportion of decided matches but also towards decreasing the mean goals scored per match.

The current study is limited to only seven leagues and over a shorter period before and after the introduction of the 3pfaw rule (only about 7 years). This, in turn, corresponds to the 1990s (late 1970s to early 1980s in the case of England). Much has happened in the football world over the past two-three decades and future works can examine if our results hold over longer periods or for other leagues. Our hope is that the current contribution serves as an inspiration for such future investigations.

Authors' contributions

A first draft of this work with preliminary results from analyzing the mean-goals has been on the first author's (GG) table for some time. But, it couldn't see the light of the day until the arrival of the co-author (PM). PM improved the analysis of the mean goals by extending it to the case where the variances are assumed unknown. Further, he was responsible for the analyses of the proportions of decided matches which form the second half of the paper. Both authors contributed in reviewing relevant literature and in the write-up and are listed in alphabetical order to reflect approximately equal contributions.

Acknowledgments

Previous versions of this report have been presented at two of the regular seminars at the Department of Statistics, Stockholm University. The authors would like to thank participants at those occasions and particularly to Daniel Thorburn and Frank Miller for constructive comments and suggestions.

References

Aylott, M. and Aylott, N. (2007) A Meeting of Social Science and Football: Measuring the Effects of Three Points for a Win. *Sport in Society* **10**, 205-222.

Bauwens, L. and Rombouts, J.V.K. (2012) On Marginal Likelihood Computation in Change-point Models. *Computational Statistics & Data Analysis* **56**, 3415-3429.

Brocas, I. and Carrillo, J. D. (2004) Do the "Three-point Victory" and "Golden Goal" Rules Make Soccer More Exciting? *Journal of Sports Economics* **5**, 169-185.

Carlin, B. P., Gelfand, A. E., and Smith, A. F. M. (1992) Hierarchical Bayesian Analysis of Change-point Problems. *Appl. Statist.* **41,** 389-405.

Chin Choy, J. H. and Broemeling, L. D. (1980) Some Bayesian Inferences for a Changing Linear Model. *Technometrics* **22**, 71-78.

Dilger, A. and Geyer, H. (2009) Are Three Points for a Win Really Better Than Two? A Comparison of German Soccer League and Cup Games *Journal of Sports Economics* **10**, 305-318

Dominicus, A., Ripatti, S., Pedersen, N. L. and Palmgren, J. (2008) A Random Change-point Model for Assessing Variability in Repeated Measures of Cognitive Function. *Statist. Med.* **27**, 5786-5798.

Gill, J. (2008) Bayesian Methods - A Social and Behavioral Sciences Approach. 2^{nd} Edition. Chapman and Hall/CRC

Hall, C. B., Ying, J., Kuo, L. and Lipton, R. B. (2003) Bayesian and Profile Likelihood Change-point Methods for Modeling Cognitive Function Over Time. *Computational Statistics & Data Analysis* **42**, 91-109.

Hon, L. Y. and Parinduri, R. A. (2016) Does the Three-point Rule Make Soccer More Exciting? Evidence from a Regression Discontinuity Design. *Journal of Sports Economics* **17**, 377-395.

Kass, R. E. and Raftery, A. E. (1995) Bayes Factors. *Journal of the American Statistical Association* **90**, 773-795.

Lee, C. -B. (1998) Bayesian Analysis of a Change-point in Exponential Families with Applications. *Computational Statistics & Data Analysis* **27**, 195-208.

Moschini, G. C. (2010) Incentives and Outcomes in a Strategic Setting: The 3-points-for-a-win System in Soccer. *Economic Inquiry* **48**, 65-79.

Petrone, S. and Raftery, A. E. (1997) A Note on the Dirichlet Process Prior in Bayesian Nonparametric Inference with Partial Exchangeability. *Statistics & Probability Letters* **36**, 69-83.

Raftery, A. E. (1996) Hypothesis Testing and Model Selection. Pp. 163-188 inn Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (Eds), Markov Chain Monte Carlo in Practice. Chapman & Hall. New York

Raftery, A. E. and Akman, V. E. (1986), Bayesian Analysis of a Poisson Process with a Change-point. *Biometrika* **73**, 85-89.

Smith, A. F. M. (1975) A Bayesian Approach to Inference About a Change-point in a Sequence of Random Variables. *Biometrika* **62**, 407-416.

Van Den Hout, A., Muniz-Terrera, G., and Matthews, F. E. (2013) Change-point Models for Cognitive Tests Using Semi-parametric Maximum Likelihood. *Computational Statistics & Data Analysis* **57**, 684-698.

Western, B. and Kleykamp, M. (2004) A Bayesian Change-point Model for Historical Time Series Analysis. *Political Analysis* **12**, 354-374.

Appendix

Figure 1: Trends in mean goals before and after introduction of 3pfaw rule across seven leagues

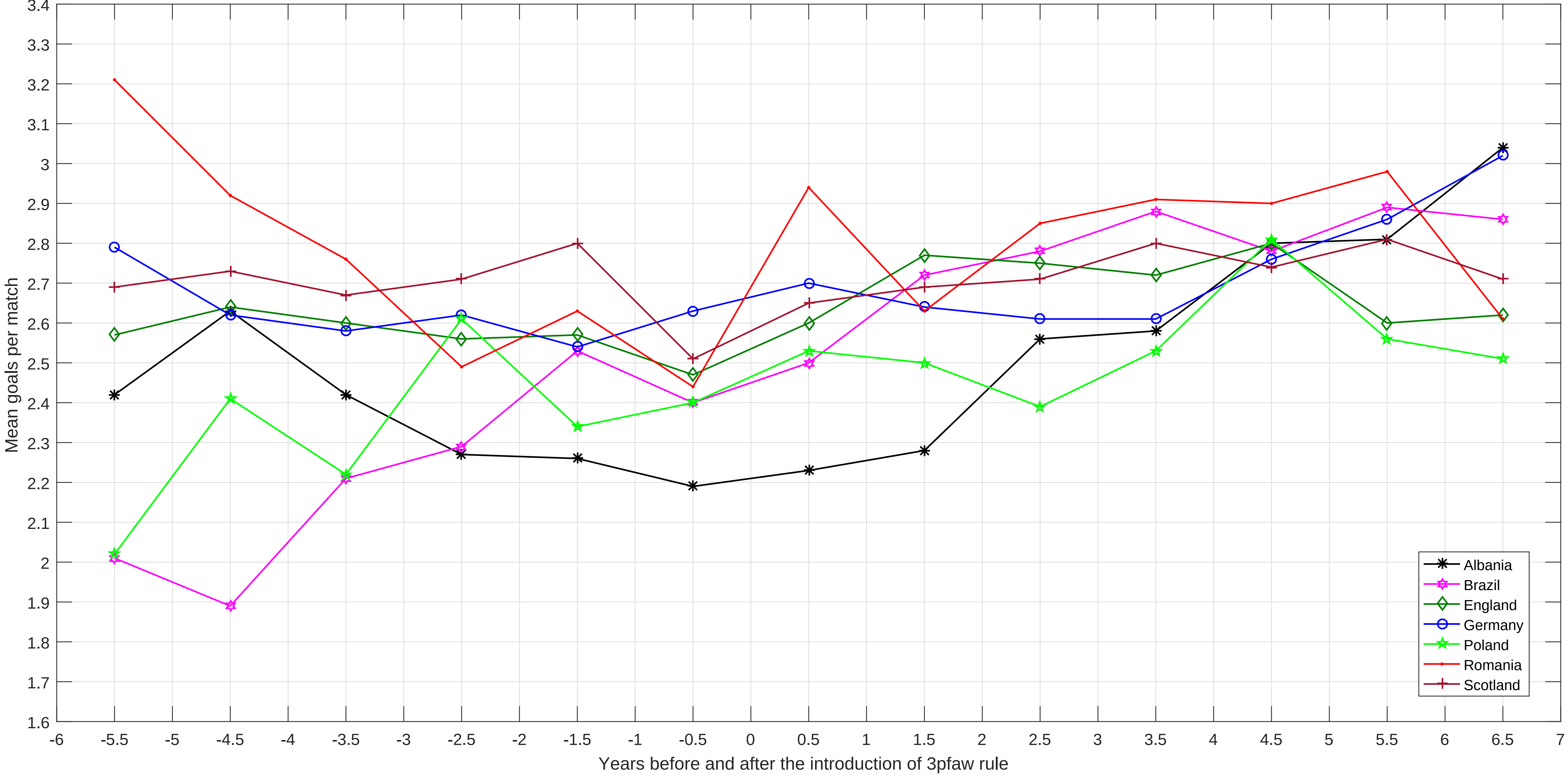

Figure 2: Trends in proportions of decided matches before and after introduction of the 3pfaw rule across seven leagues

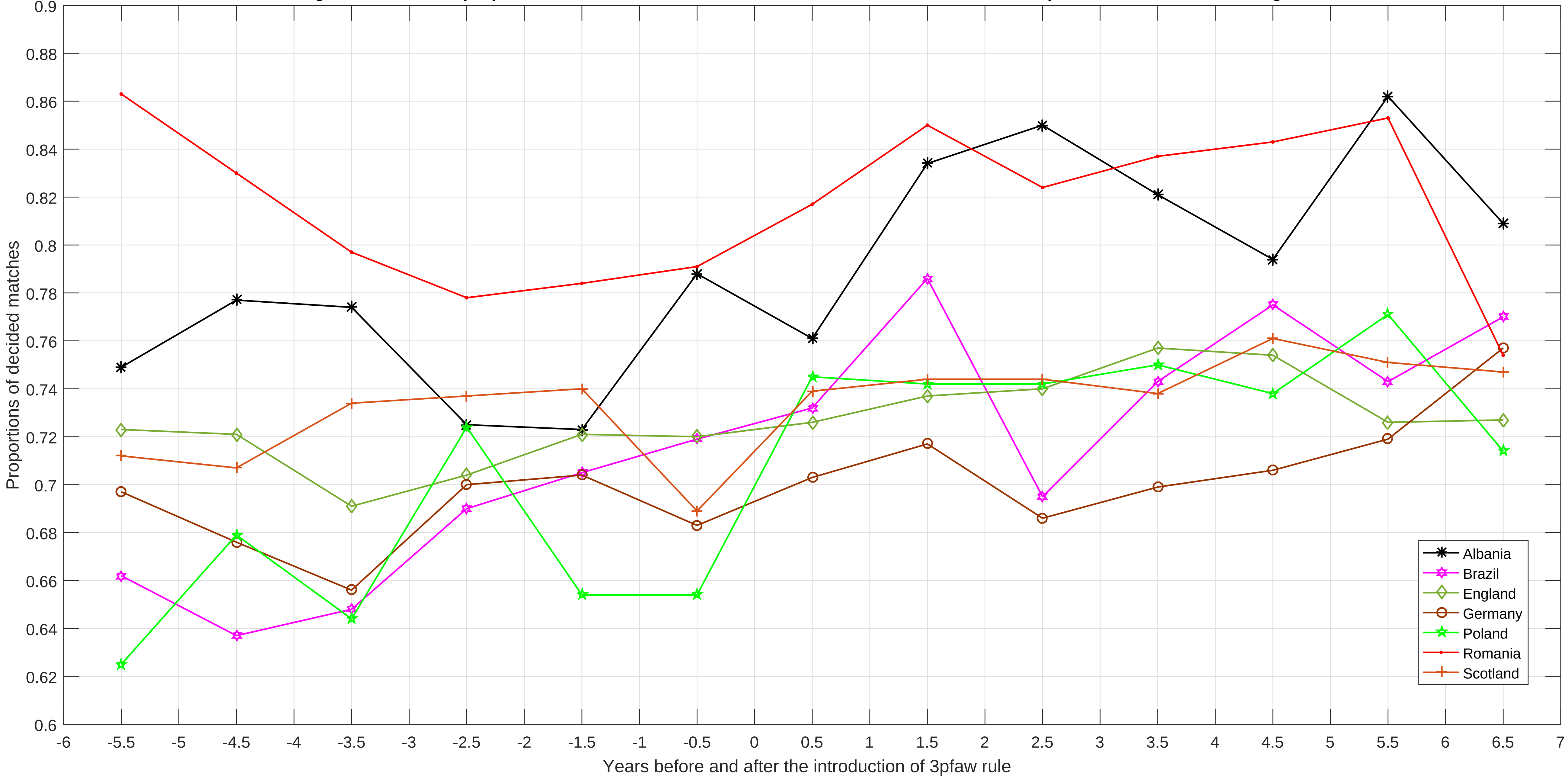

Figure 3: Posterior distributions of mean goals before and after the 3pfaw rule in seven leagues (a - g), change-points (h), and differences in mean goals (i)

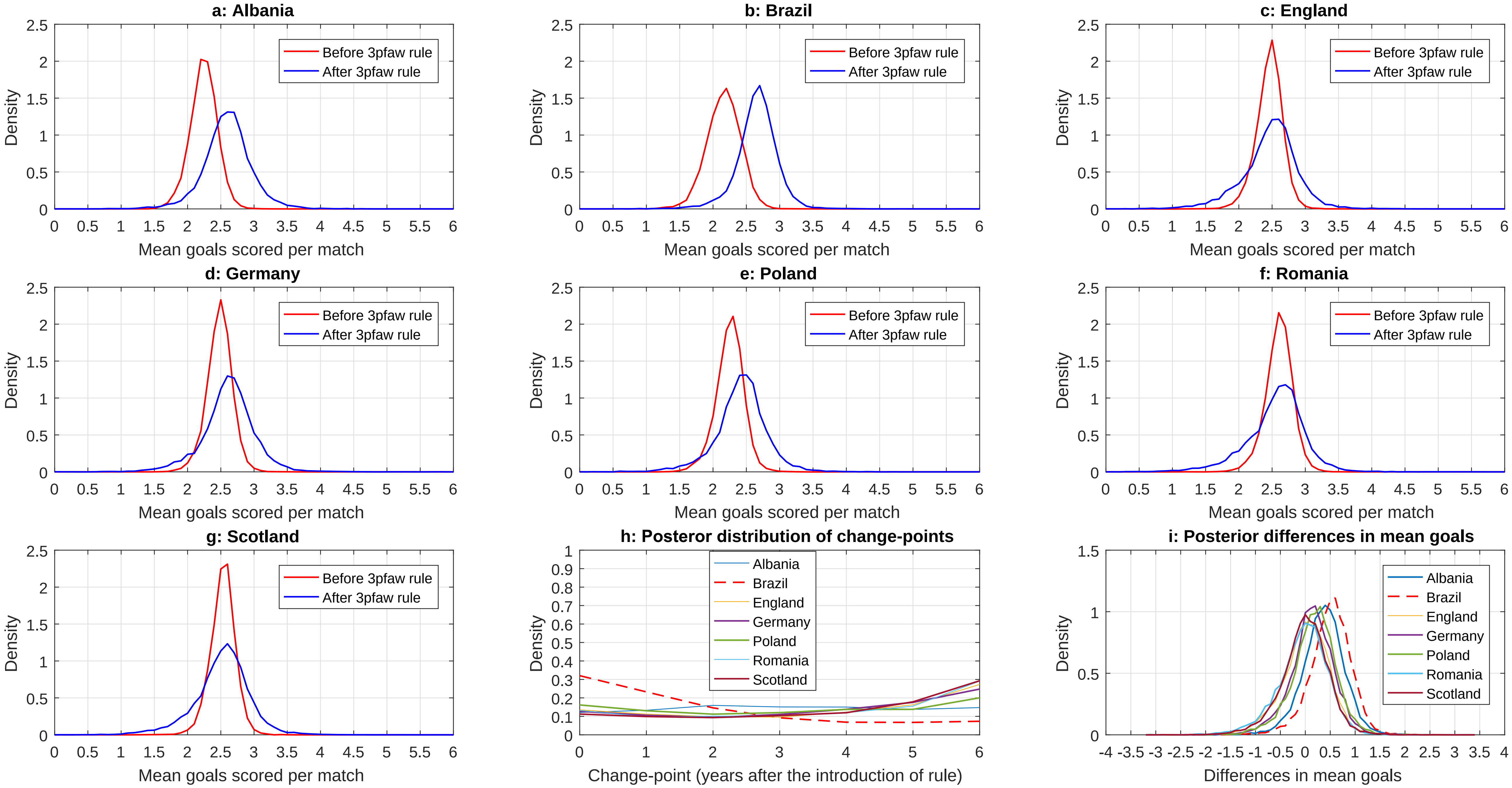

Figure 4: Posterior distributions of proportions of decided matches before and after the 3pfaw rule (a - g), change-points (h), and differences in proportions (i)

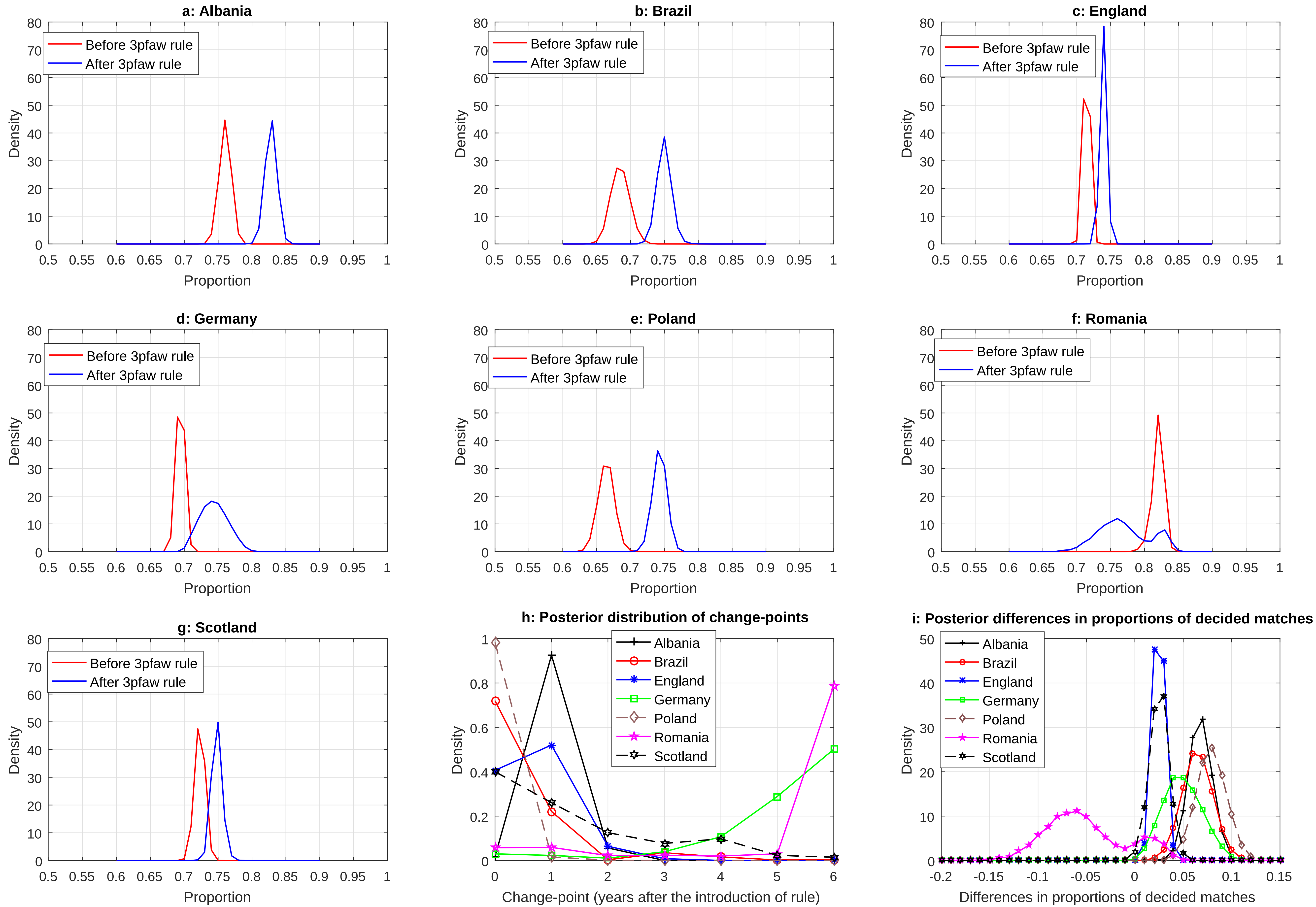

Figure 5: Posterior distributions of change-points (k) for mean-goals (red) and proportions of decided matches (blue) across the seven leagues studied