Is gender a competitive balance driver? Evidence from Scandinavian football

Morten Kringstad

Abstract: Research question: Women's football is more popular than ever, but it still has much smaller attendance than men's football. Rottenberg's uncertainty of outcome hypothesis suggests a relationship between competitive balance and attendance. This paper focuses on competitive balance. The research question is: What is the level of competitive balance in the women's football leagues in Scandinavia (Denmark, Sweden and Norway), and how is this level compared to their respective men's leagues? Research methods: The research methods are quantitative in two ways. The first part is related to measures for measuring competitive balance and the second part concerns statistical tests for analysing the differences in competitive balance. The data is from the end-of-season tables from the six Scandinavian football leagues for men and women over the period from 1995 to 2015 (up to 120 observations). Results and findings: A comparison with the men's leagues shows that competitive balance is much weaker in women's football in Scandinavia. This statement holds for all measures except championship winner concentration. Among the women's leagues, Denmark has the weakest competitive balance. Norway and Sweden are very similar, but two of the measures find the latter to be better balanced. Implications: Competitive balance in Scandinavian women's football is weak. If competitive balance in women's football matters, the results in this paper suggest more equal distribution of league level revenues than for the men's football leagues.

ABOUT THE AUTHOR
Dr Morten Kringstad is an associate professor at NTNU Business School, Norwegian University of Science and Technology, Norway. The main research area is business and economics in professional team sports, with a particular focus on competitive balance and related topics in European football. The results in this paper is the first step in looking further into whether competitive balance differs between the genders in football and other team sports.

PUBLIC INTEREST STATEMENT
The level of competitive balance in a sports league reflects whether the results (sporting outcome) among the teams are tight or not. For example, big point score differences in the final standing between the stronger and weaker teams indicate weak competitive balance. This paper focuses on whether competitive balance differs between men's and women's leagues in Scandinavian football. Results show that the women's leagues are less balanced than the men's. In women's leagues, the differences in sporting results (outcome) among the teams are larger, and sporting results (outcome) varies less across seasons. Overall, results indicate weak competitive balance in the Scandinavian football leagues for women. Among these leagues, Denmark has weakest competitive balance, while measures find the Swedish to either be better than or equally balanced to the Norwegian league.
1. Introduction

The Fédération Internationale de Football Association (FIFA)\(^1\) has issued the following statement: “Women’s football has flourished in recent years, and interest in football played by women is currently at an all-time high”. According to FIFA, there are about 30 million female football players worldwide,\(^2\) and the most watched football match on TV in the US was the 2015 World Cup final, in which the US women’s national team was playing.\(^3\) However, despite the all-time high popularity, attendance is in general low compared to the men’s leagues. For example in 2012, the top division in Sweden and Denmark had an average of 856 and 137 spectators, respectively (Kjaer & Agergaard, 2013). In Norway, this average was 197 for the same season.\(^4\) The comparative numbers for the men’s top division were 7,210 in Sweden, 7,103 in Denmark and 7,010 in Norway.\(^5\) The Scandinavian countries (Denmark, Sweden and Norway) are used as a case, because they have “pioneered the development of women’s soccer” (Kjaer & Agergaard, 2013, p. 823).\(^6\) This is reflected by Sweden as the first winner of the Union of European Football Association (UEFA) Women’s Euro in 1984,\(^7\) and by Norway, as the runners-up in the first FIFA Women’s World Cup final in 1991 and as the winner of the second in 1995.\(^8\)

While the study of drivers for the demand for men’s football has been popular in the sports economics and management literature, there is according to Meier, Konjer, and Leinwather (2016, p. 5) “... a general lack of econometric studies on attendance for women’s sport”. On this basis, it is interesting to return to the seminal paper on sports economics by Rottenberg (1956) and the uncertainty of outcome hypothesis. This hypothesis is about the relationship between competitive balance and spectator interest. Although this relationship has been more difficult to detect in demand studies (see, for example, Dobson & Goddard, 2011),\(^9\) perhaps because of the many other drivers behind the number of stadium attendees, one can still claim that weak competitive balance is problematic. This was also emphasised by Szymanski (2001, p. F69): “Without a degree of competitive balance, fans will lose interest in a competition. However, it is less clear that every decline in competitive balance will lead to a falling off of fan interest”.

Several studies have measured competitive balance in different sports leagues for men, among others, Kringstad and Gerrard (2007), who provided an overview of competitive balance in European football and the major leagues in North America from 1966/67–2005/06. With this overview in mind, it is also of interest to compare competitive balance with and within women’s football leagues. The main objectives of this paper are therefore to measure the level of competitive balance in the Scandinavian (Denmark, Sweden and Norway) women’s football leagues, and to compare the results both between these leagues and with the men’s leagues in their respective countries.

Competitive balance is a complex concept. Kringstad and Gerrard (2007) drew on the three dimensions of win dispersion, performance persistence and prize concentration. These dimensions are also in focus in this paper. The first is categorised as within season competitive balance, while the two others are related to across season (between season) competitive balance. In addition to the traditional ratio of standard deviation (RSD) for measuring win dispersion, a new measure for the average absolute goal differences is developed to provide information beyond the typical measures for seasonal competitive balance. There seems to be a wider spread in win dispersion in women’s football, and this measure is of particular interest as goal differences can give a further indication of how strong or weak teams are in relative terms. Moreover, this paper also measures the draw percentage, since ties may indicate sporting closeness in football games.

There seems to be more or less systematic differences in the level of competitive balance in different sports, such as the National Basketball Association (NBA) and the National Football League (NFL)
in North America (see, for example, Fort, 2011; Quirk & Fort, 1992). Contrary to Rottenberg’s invariance proposition, the explanations are often related to differences in market regulations. However, market regulations cannot be the sole determinant, because the highly regulated NFL does not have a better RSD than the far less regulated big five European football leagues (England, Spain, France, Germany and Italy) (see, for example, Kringstad & Gerrard, 2007). One example of another driver is the structure of a sport. Fort and Lee (2007, p. 528) used “substantial rule change” as one of the explanations for a “time series break” in competitive balance in the NFL, while Kringstad and Gerrard (2007) hypothesised that the financial and sporting prize structure can be a competitive balance driver. The latter has some differences between the genders in Scandinavian football, while many of the other sporting and market rules are similar. Therefore, the null hypothesis is that competitive balance should be at the same level within and between these women’s and men’s leagues.

The paper is further organised as follows. The next section provides a discussion of the theory on competitive balance. This is followed by the research methods applied in the analysis, the presentation of the empirical results, and the discussion and conclusions.

2. Theory

A league has a competitive balance problem when weak uncertainty of outcome leads to low interest, *ceteris paribus*. This is the core part of Rottenberg’s uncertainty of outcome hypothesis, according to that: “Attendance at the games of any given team is … a negative function … of the dispersion of percentages of games won by the teams in the league”, and it is emphasised in a related footnote that “… the ‘tighter’ the competition, the larger the attendance” (Rottenberg, 1956, p. 246). Therefore, Fort and Quirk (2011, p. 464) claim that: “Rottenberg (1956) was the first to detail the problems associated with a lack of competitive balance.” Competitive balance problems are hence hypothesised to cause reduced attendance and lower revenues for the involved parties in the strategic map in Gerrard (2000). In other words, a league with a competitive balance problem may struggle with what Gerrard (2004) described as sporting and financial viability in a sports league.

Competitive balance is a driver for uncertainty of outcome because improved competitive balance is expected to increase uncertainty of outcome in general terms. Kringstad and Gerrard (2007, p. 151) defined uncertainty of outcome as “the probability distribution of the alternative outcomes of a specified sporting contest”. As was explained in Cairns, Jennett, and Sloane (1986), uncertainty of outcome is a concept that can be divided into the three levels match uncertainty, seasonal uncertainty and long-run uncertainty, where seasonal uncertainty is related to both the overall dispersion of results and to what can be called match significance. The econometric demand study of the top division in women’s football in Germany (FBL) by Meier et al. (2016) found match uncertainty to be insignificant, while match significance was a significant demand driver. Further, they claimed that: “… given the rather small marginal effects the results do not suggest a strong need for policies to improve competitive balance within the FBL.… The finding that seasonal uncertainty matters for attendance is insofar relevant as the FBL represents a rather imbalanced league in which the top teams heavily concentrate national playing talent” (p. 15). The latter statement is of particular interest for this paper.

The concept of competitive balance can be defined both ex ante and ex post, and the definition is based on the distribution of sporting outcomes in a sporting contest (Kringstad & Gerrard, 2007). This paper will focus on “the ex post distribution of actual sporting outcomes” (p. 152). According to Kringstad and Gerrard (2007) this means that the dimension win dispersion is related to differences in sporting outcomes between the teams in the tournament for a given season. Win dispersion is weak when there are large differences in sporting outcomes between the teams. An example of weak win dispersion can be drawn from Norwegian women’s football over the seasons 2008–2012, where the bottom teams on aggregate had one win and eight draws out of 110 matches. Performance persistence is related to variation in league outcomes, such as end-of-season standings, across seasons. If the same teams are approximately at the same level season after season, it is a sign of weak competitive balance. An example here stems from the 2010/11 to the 2011/12 season among the
ten-team league of Danish women’s football, where only one team had a change in two places, and all the other teams had one or no change in their final league standing between these two seasons. Prize concentration is the third dimension of competitive balance. Weak competitive balance is reflected when, for example, one or only a few teams win the championship over many seasons. Among the women’s leagues in focus in this paper, Denmark seems to have a high ten-season championship winner concentration from 2006, with one team winning seven out of the ten championships, and the other three were won by one other team. The study by Treber, Levy, and Matheson (2013) finds the championship concentration among women’s intercollegiate basketball (NCAA) to be higher than for the men’s leagues. Moreover, prize concentration can also be related to teams qualifying for domestic post seasonal play-offs or for international tournaments arranged by, for example, the Union of European Football Associations (UEFA), such as the top four in the English Premier League. The first is shown for the Women’s Basketball Association (WNBA) in North America in Berri and Krautmann (2013). This paper focuses only on championship winner concentration in this dimension.

The literature has, over time, regarded the distribution of drawing power among the teams as the main driver for the level of competitive balance within a given team sport league. According to Vrooman (1995), Quirk and El-Hodiri (1974, p. 58) claimed, on the basis of a model by El-Hodiri and Quirk (1971), that “…imbalance is due to the differences in the drawing potentials ….” This was further emphasised by Dobson and Goddard (2004, p. 359), who claimed: “In the North American (NA) literature on resource allocation in professional team sports leagues, relative market sizes are the final arbiter of the allocation of playing talent between teams, and therefore competitive balance (El-Hodiri & Quirk, 1971; Fort & Quirk, 1995; Rottenberg, 1956; Vrooman, 1995).” Competitive balance has been a motivation for introducing market regulations (the opposite of Rottenberg’s invariance proposition) in the labour market (e.g. reserve rules and maximum wages) and the product market (e.g. gate revenue sharing and collectively sold media rights). Other regulations that may affect interest and competitive balance are related to prize structure (e.g. post seasonal play-offs) and playing rules (e.g. the back-pass rule in football). These regulations will not be discussed in this paper. Additional competitive balance level drivers analysed in the literature include the limitations of particular playing talent in the population, referring to basketball, where player height may also be relevant (Berri, Brook, Frick, Fenn, & Vicente-Mayoral, 2005), and the general level of the number of goals or points scored in each match of the sport (Groot, 2008). The latter is related to the notion that more goals in the average game implies weaker competitive balance, ceteris paribus.

3. Research methodology

The three dimensions of competitive balance will be measured by the ratio of standard deviation (RSD) for win dispersion, the Spearman’s rank correlation coefficient (SRCC) for performance persistence and the Herfindal’s index (HI) for championship concentration. These measures are all used and explained in Kringstad and Gerrard (2007). In addition, two other measures are applied to further reflect the level of seasonal competitive balance. They are the draw percentages and an introduced measure of average absolute goal difference.

Win dispersion is usually measured by the ratio of standard deviation (RSD) (Fort, 2007). This measure is applied in Quirk and Fort (1992), who attributed it to Noll (1988) and Scully (1989). The RSD measure is defined as in Equation (1) (see, for example, Fort, 2007):

$$ RSD = \frac{ASD}{ISD} $$

where ASD is the actual standard deviation of the win percentages at the end-of-season league standing. ISD is the idealised standard deviation, defined as $0.5/\sqrt{m}$ where m is the number of matches played by each team in the league (see, for example, Fort & Quirk, 1995). The idealised standard deviation is the theoretical standard deviation of the outcome in a theoretical league, where all teams are expected to have equal ex ante sporting quality, i.e. each team has an expected
win percent equal to 50% (the number of wins and losses are expected to be equal). Moreover, the idealised standard deviation is based on a binomially distributed outcome and is contingent on the number of matches played (see for example Downward & Dawson, 2002; Owen, 2010, 2014).

As for ASD, there has been a discussion about how to incorporate draws (see the discussions in Cain & Haddock, 2006; Fort, 2007; Owen, 2012). This paper follows the traditional way and treats draws as half wins. In addition, the denominator is calculated as in the original expression, defined in Equation (1).17 Hence, the typical European point score system is also ignored in this way to apply the measure.18 Of course, instead of win percentage, point percentage could have been an alternative. The traditional way is chosen to facilitate comparison with other studies.

Performance persistence can be measured, for example, by the Spearman’s rank correlation coefficient (SRCC) (Andreff & Raballand, 2011; Daly & Moore, 1981; Kringstad & Gerrard, 2007). According to Kringstad and Gerrard (2007), the formula for the SRCC can be written as in Equation (2):

$$SRCC = \frac{1 - 6 \sum D_i^2}{N(N^2 - 1)}$$  \hspace{1cm} (2)

where $D_i$ is the difference in league standing from season $t - 1$ to season $t$. $N$ is the number of teams in the league in season $t$.

SRCC was first used in a closed North American league (i.e. Daly & Moore, 1981). Since the Scandinavian football leagues for women follow the traditional European way, with relegation and promotion across the hierarchical structured divisions, adaptions are required. Therefore, the approach taken in Kringstad and Gerrard (2007) is adopted by ranking the teams ex post as the original end-of-season standing in the top division, while the ex ante standing is adjusted into a ranking including a replacement between the relegated top division teams and the promoted teams. The ex ante ranking among the promoted teams is based on the final standing in the second level division(s) in season $t - 1$.

This paper will only look at the distribution of championship winners in the dimension of prize concentration. One of the most applied measures here is the Herfindal index, as in Kringstad and Gerrard (2007), with the following formula (Equation (3)):

$$H - index = \sum_i \left( \frac{W_i}{M} \right)^2$$  \hspace{1cm} (3)

where $W_i$ is the number of championships won by team $i$ in a period of $M$ seasons.

Earlier research has applied goal scoring in the context of competitive balance (Andreff & Raballand, 2011; Groot, 2008). In this respect, Andreff and Raballand (2011, p. 132) claimed: “… if all teams’ sporting strengths were exactly even and equally matched, the most probable league results would be all games emerging as draws”. Further, they found that an increased share of 0–0 and 1–0 scores is associated with improved competitive balance. On this basis, this paper also incorporates goal scoring in a competitive balance context, because it can provide further information about possible differences between men’s and women’s football beyond the traditional competitive balance measures. This is based on the notion that differences in sporting quality can be reflected in the goal differences among the teams. In terms of constructing an appropriate measure indicating the level of goal difference on the basis of the end-of-season table that can be comparable across leagues, this paper suggests the formula in Equation (4):

$$CB_{GD} = \frac{\sum_{i=1}^{N} |X_i - Y_i|}{m \times N}$$  \hspace{1cm} (4)
where $X_i$ and $Y_i$ are the number of goals for and against team $i$. By using $m \times N$ in the denominator, this expression will measure the average (end of season) absolute goal difference per game for all teams and should hence be comparable across leagues. A league in which some teams have big goal differences will hence have a higher $\text{CBGD}$ compared to a league where the goal differences are smaller.

The measure for draw percentage is as follows (Equation (5)):

$$\text{Draw}\% = \frac{\sum_{i=1}^{N} \text{Draws}_i}{m \times N}$$

where the nominator is the aggregated number of draws for team 1 to team $N$, $m$ is the number of games for each team in the league, and $N$ is the number of teams in the league.

When comparing differences between samples of data, methods are chosen on the basis of whether the data are parametric or not. The Shapiro-Wilk W test in Stata is applied to test normality. The two-sample Wilcoxon rank-sum (Mann-Whitney) test is chosen when one or both samples are non-parametric. In the case of parametric samples, the $t$-tests are either done with an assumption of equal or unequal variance. In addition, a simple regression with a gender dummy variable is applied to measure differences in competitive balance in general for each of the measures except the H-Index.

Because the analysis in this paper is done on the basis of different leagues over time, one may argue that there are time invariant elements within each league (entity). Hence, we should have the basis to apply panel data techniques to measure possible systematic differences in competitive balance across leagues over time. In other words, systematic differences between entities can be captured by a fixed effects panel data model. In this study, each league is an entity. Following Wooldridge (2009, p. 481), a fixed effects model with one explanatory factor can be written as in Equation (6):

$$y_{it} = \beta_1 x_{it} + \alpha_i + \mu_{it}$$

The dependent variable $y_{it}$ is the competitive balance measure for league $i$ at season $t$, while $x_{it}$ is the independent variable ($x_{i,t} = x$). This means that the league specific differences are captured in $\alpha_i$. Further, this paper follows the statement of Wooldridge (2009, p. 486), who claimed that “… the fixed effects estimates can be obtained by the dummy variable regression”. Hence, applying a dummy variable regression model gives the fixed effects competitive balance differences between the default league and each of the other leagues. Since the two-sample tests used in this paper indicate that the Danish women’s league is the worst balanced, this league is set as the default league. Therefore, $\alpha_i$ represents the dummy variable for league $i$, where $i$ refers to each of the other leagues in the sample. Robust standard errors are computed because of heteroscedasticity when applying SRCC and $\text{CBGD}$. In general, the diagnostics from the dummy variable models show no autocorrelation problems.

4. Empirical results
Below, the empirical results are presented on the basis of the end-of-season tables for the top division for men and women in Scandinavian football. In addition, the end-of-season tables from the second level divisions are used a part of the calculation of the SRCC. The main sources were retrieved from www.rsssf.com.

4.1. Championship winner concentration (H-Index)
Table 1 shows no clear patterns when comparing the distribution of championship winners among the Scandinavian football leagues, except that the men’s league in Sweden has the best H-Index in all measured periods.
4.2. Differences in competitive balance between women’s and men’s football leagues?
The competitive balance measures in Tables 2 and 3 show that women’s leagues, in general, have significantly weaker balance than the men’s leagues in Scandinavian football. Moreover, comparing the average values of the different measures, the gap between genders in the competitive balance values is large. For example, the RSD is nearly three times as weak (132 percent above the idealised standard deviation) for women compared to men (47% above). The draw percentage for the men’s leagues, on average, is comparable to Dobson and Goddard’s (2011) findings for England, France, Germany, Scotland and Spain.

4.3. Competitive balance in Scandinavian football—League by league
All measures in Tables 2 and 3 show a strong significant difference in competitive balance between women’s and men’s leagues in Scandinavian football. The next step is to apply the same measures on the league level. The presentation is done separately for each of the four competitive balance measures used in Tables 2 and 3. Tables 4–7 are designed as a matrix, in which the upper part compares the average competitive balance index between the leagues, and the cells in the middle of the table refer to significance level and the type of two-sample test applied. The bottom part of the tables splits the sample period into two equal parts for each league. Whether the differences between these sub-periods are significant or not is shown in the bottom row, together with which test is used. Further, Table 8 presents the results from the fixed effects regressions with league dummies, where Denmark’s women’s league is the default league.

The results from measuring RSD in Table 4 show that the Scandinavian women’s football leagues have large win dispersion, particularly the Danish league, with an average RSD of 2.59. However, this Danish RSD is significantly improved (reduced) between the first and the second half of the sample period.

| Table 1. Championship concentration in Scandinavian football measured by the H-Index |
|---------------------------------|---|---|---|---|---|
| Average HI | Denmark | Sweden | Norway |
| Gender | Men | Women | Men | Women | Men | Women |
| Total | 0.30 | 0.36 | 0.12 | 0.24 | 0.46 | 0.19 |
| 96–05 | 0.36 | 0.26 | 0.18 | 0.36 | 0.82 | 0.34 |
| 06–15 | 0.42 | 0.58 | 0.18 | 0.36 | 0.28 | 0.30 |

| Table 2. Average values of competitive balance between women’s and men’s football leagues |
|---------------------------------|---|---|---|---|---|
| RSD | SRCC | CBGD | Draw% |
| Average Men | 1.47 | 0.53 | 0.49 | 0.25 |
| Average Women | 2.32 | 0.78 | 1.56 | 0.14 |
| Difference | 0.85*** | 0.25*** | 1.07*** | 0.11*** |

***Significant at the 1% level, based on Mann-Whitney test.

| Table 3. Competitive balance in Scandinavian football with a gender (men’s leagues) dummy variable |
|---------------------------------|---|---|---|---|---|
| RSD | SRCC | CBGD | Draw% |
| Coeff. | St.E. | Coeff. | St.E. | Coeff. | St.E. | Coeff. | St.E. |
| Season | −0.0034 | 0.0047 | 0.0015 | 0.0028 | −0.0110** | 0.0051 | 0.0004 | 0.0007 |
| Gender | −0.8469*** | 0.0538 | −0.2472*** | 0.0320 | −1.068*** | 0.0589 | 0.107*** | 0.0077 |
| Intercept | 9.131 | 9.424 | −2.298 | 5.645 | 23.72** | 10.33 | −0.6561 | 1.350 |
| Adj. $R^2$ | 0.6759 | 0.3317 | 0.7364 | 0.6167 |

*Significant at the 10% level.  
**Significant at the 5% level.  
***Significant at the 1% level.
### Table 4. RSD for each Scandinavian football league over the period 1996–2015

|       | Women |       |       | Men |       |       |
|-------|-------|-------|-------|-----|-------|-------|
|       | Denmark | Sweden | Norway | Denmark | Sweden | Norway |
| RSD   | 2.59 | 2.17 | 2.20 | 1.61 | 1.40 | 1.41 |
| Women | Den | 2.59 | T-eq*** | Non*** | T-eq*** |       |
|       | Swe | 2.17 | T-eq*** | Non | T-eq*** |       |
|       | Nor | 2.20 | Non*** | Non | Non*** | T-eq*** |
| 1996–05 | 2.80 | 2.22 | 2.19 | 1.58 | 1.29 | 1.41 |
| 2006–15 | 2.41 | 2.12 | 2.22 | 1.63 | 1.51 | 1.41 |
| Difference | T-eq*** | T-eq *** | T-eq *** | T-eq ** | T-eq | T-eq *** |

Notes: T-eq = t-test assuming equal variance; T-un = t-test assuming unequal variance; Non = non-parametric.

*Significant at the 10% level.
**Significant at the 5% level.
***Significant at the 1% level.

### Table 5. SRCC for each Scandinavian football league over the period 1996–2015

|       | Women |       |       | Men |       |       |
|-------|-------|-------|-------|-----|-------|-------|
|       | Denmark | Sweden | Norway | Denmark | Sweden | Norway |
| SRCC  | 0.83 | 0.76 | 0.75 | 0.57 | 0.55 | 0.48 |
| Women | Den | 0.83 | Non* | T-eq** | T-un*** |       |
|       | Swe | 0.76 | Non* | T-eq | T-eq*** |       |
|       | Nor | 0.75 | T-eq*** | T-eq | Non*** | T-eq*** |
| 1996–2005 | 0.85 | 0.77 | 0.72 | 0.57 | 0.52 | 0.48 |
| 2006–2015 | 0.81 | 0.75 | 0.78 | 0.57 | 0.58 | 0.47 |
| Difference | T-eq | T-eq *** | T-eq *** | T-eq ** | T-eq | T-eq *** |

Notes: T-eq = t-test assuming equal variance; T-un = t-test assuming unequal variance; Non = non-parametric.

*Significant at the 10% level.
**Significant at the 5% level.
***Significant at the 1% level.

### Table 6. CB_{eq} for each Scandinavian football league over the period 1996–2015

|       | Women |       |       | Men |       |       |
|-------|-------|-------|-------|-----|-------|-------|
|       | Denmark | Sweden | Norway | Denmark | Sweden | Norway |
| CB_{eq}  | 1.98 | 1.25 | 1.46 | 0.49 | 0.47 | 0.52 |
| Women | Den | 1.98 | Non*** | Non*** | Non*** |       |
|       | Swe | 1.25 | Non*** | T-un** | Non*** |       |
|       | Nor | 1.46 | Non*** | T-un** | Non*** | T-un** |
| 1996–2005 | 1.96 | 1.40 | 1.57 | 0.53 | 0.47 | 0.60 |
| 2006–2015 | 2.00 | 1.11 | 1.36 | 0.46 | 0.47 | 0.44 |
| Difference | T-eq | T-eq*** | T-eq*** | T-eq | T-un | T-eq*** |

Notes: T-eq = t-test assuming equal variance; T-un = t-test assuming unequal variance; Non = non-parametric.

*Significant at the 10% level.
**Significant at the 5% level.
***Significant at the 1% level.
period. Compared to the other women’s leagues, it has significantly worse RSD. Norway and Sweden have relatively similar RSD. These results also seem robust over time. Additionally, for each country, the women’s league has significantly higher RSD than the men’s league. Further, Table 4 indicates that among the men’s leagues, Denmark has the largest win dispersion. Among the men’s leagues, Denmark has a significantly larger win dispersion than the two other leagues at the 5% significance level. There is no significant difference between Norway and Sweden.22

The SRCC measure for performance persistence shows much of the same pattern as for the RSD measure, except that the differences between the Danish and the other two women’s leagues are less significant (Tables 5 and 8). Compared to the men’s leagues, this measure also shows strong significant differences between the men’s and women’s leagues within each country, while no leagues have a significantly different competitive balance in the first half compared to the second half of the sample. Between the men’s leagues, there are no significant differences.23

The results from the average absolute (end of season) goal differences measure also show large differences in competitive balance between the women’s and men’s football leagues within each of the Scandinavian countries (Tables 6 and 8). Among the women’s leagues, Denmark is again the least balanced. Another feature from the results of the women’s leagues is that Norway and Sweden are significantly divided, with the Swedish league as the best balanced. This league also had significant improvement between the first and second half of the sample period. The panel data model in Table 8 suggests a general improvement over the sample period in absolute goal difference (smaller absolute average difference over time). There is no significant difference between the men’s leagues.24

The share of drawn matches is significantly lower in all the women’s leagues in Scandinavia, compared to their respective men’s leagues (Table 7). Within each league, there seems to be little change from period to period. The women’s leagues follow the results from the average absolute goal differences. Sweden is the best balanced, ahead of Norway and Denmark. The differences are significant or strongly significant. There is no significant difference between Sweden and Denmark when it comes to the men’s leagues.25 Norway has a lower draw percentage than the other two, and is weakly significant against Sweden and strongly significant against Denmark.26

Groot (2008) used goals scored per match as an explanatory factor for differences in competitive balance, indicated, for example, by the high number of goals scored in basketball matches. In European football, Andreff and Raballand (2011) calculated the average number of goals in a football

---

Table 7. Draw percent for each Scandinavian football league over the period 1996–2015

| Draw% | Women | Men |
|-------|-------|-----|
|       | Denmark | Sweden | Norway | Denmark | Sweden | Norway |
| 0.10  | Non*** | Non*** | Non*** | 0.26    | 0.26    | 0.23   |
| 0.18  | T-eq** | T-eq** | T-eq** |         |         |        |
| 0.15  | Non*** | T-eq** |        |         |         |        |
|       |        |        |        | 0.26    | 0.26    | 0.23   |

Notes: T-eq = t-test assuming equal variance; T-un = t-test assuming unequal variance; Non = non-parametric.

*Significant at the 10% level.
**Significant at the 5% level.
***Significant at the 1% level.
match to be 2.55 among the big five football leagues over the period 2003–2007. Below, Table 9 shows the average for each of the Scandinavian football leagues over the sample period (1996–2015). Although the last half of the sample has a lower average than the first, the average is slightly higher for the men’s leagues, while it is much higher in the women’s league matches. Hence, there seems to be a difference in the average number of goals scored between men’s and women’s football in Scandinavia.

### Table 8: Dummy variable regressions with various competitive balance measures as dependent variable

| Dep var. | RSD  | SRCC | CBgd | Draw% |
|----------|------|------|------|-------|
| Season   | -0.0041 | 0.0013 | -0.0121*** | 0.0005 |
|          | (0.0043) | (0.0025) | (0.0040) | (0.0006) |
| Den_M    | -0.9885*** | -0.2596*** | -1.4924*** | 0.1577*** |
|          | (0.1017) | (0.0526) | (0.0963) | (0.0117) |
| Swe_M    | -1.1990*** | -0.2825*** | -1.5178*** | 0.1564*** |
|          | (0.0936) | (0.0459) | (0.0967) | (0.0131) |
| Swe_W    | -0.42705*** | -0.0688* | -0.7331*** | 0.0748*** |
|          | (0.0838) | (0.0364) | (0.1021) | (0.0126) |
| Nor_M    | -1.1875*** | -0.3537*** | -1.4696*** | 0.1316*** |
|          | (0.0955) | (0.0576) | (0.0968) | (0.0129) |
| Nor_W    | -0.3929*** | -0.0800** | -0.5226*** | 0.0487*** |
|          | (0.0892) | (0.0388) | (0.1200) | (0.0123) |
| Intercept| 10.767 | -1.7614 | 26.2249*** | -0.9025 |
|          | (8.7275) | (5.0156) | (7.9217) | (1.1495) |
| R²       | 0.7687 | 0.3744 | 0.8606 | 0.7418 |

*Significant at the 10% level.
**Significant at the 5% level.
***Significant at the 1% level.

### Table 9: Average number of goals per game in Scandinavian football over the period 1996–2015

| Goals/Game | Denmark | Sweden | Norway |
|------------|---------|--------|--------|
| Gender     | Men | Women | Men | Women | Men | Women |
| 1996–2015  | 2.87 | 4.42 | 2.69 | 3.50 | 3.15 | 3.95 |
| 1996–2005  | 3.04 | 4.54 | 2.74 | 3.80 | 3.32 | 4.25 |
| 2006–2015  | 2.71 | 4.31 | 2.65 | 3.19 | 2.98 | 3.65 |

Although the last half of the sample has a lower average than the first, the average is slightly higher for the men’s leagues, while it is much higher in the women’s league matches. Hence, there seems to be a difference in the average number of goals scored between men’s and women’s football in Scandinavia.

### 5. Discussions

An empirical analysis by Treber et al. (2013) finds college basketball to have higher championship concentration among women than men. These findings are not supported for Scandinavian football in this paper (Table 1). On the other hand, the measures for the two other dimensions (Tables 2–9) show that competitive balance is much weaker in Scandinavian women’s football leagues compared to the men’s football leagues, both in general and within each country. More specific, among the women’s leagues, Denmark has the weakest competitive balance, while it is more difficult to distinguish between Sweden and Norway when using the traditional measures for win dispersion and performance persistence. However, the draw percentage and the average absolute goal difference indicate greater dispersion in sporting outcomes among the teams in Norway. Taking into account
these results, competitive balance in Scandinavian women’s football leagues seem to be much weaker than for the men’s leagues.

Treber et al. (2013) apply the “evolution-based explanation” (p. 258) by Gould (1986, 1996) as an argumentation for their findings of the championship concentration differences among the genders. They claim that (pp. 258–259):

Gould argued that player ability in any sport is initially widely dispersed. A small number of naturally gifted athletes are near the limit of human capability, while the average players is much farther to the left in the distribution. As a sport’s popularity grows and practice and training intensify, more and more players move closer to the limit of human capability, and the distribution of player ability becomes more compressed.

On this basis, Treber et al. (2013) use difference in the “dispersion of player ability” (p. 260) as a possible explanation for the difference in competitive balance between the genders. Relevant for Scandinavian and European football are the argumentation for these differences, that (p. 260): “Opportunity to engage in competitive basketball have been greater for men than for women”, and that the financial return for success is much higher for men (NBA) compared to women (WNBA). In Europe, men’s football has both a longer history and is at a much higher financial level than what is the situation for women’s football.27

Compared to the results in Kringstad and Gerrard (2007), both RSD and SRCC in the Scandinavian football leagues for men are at the same level as for the big five in Europe. Also compared with the North American major leagues, only one league (NBA) has results that indicate weaker win dispersion than the Scandinavian football leagues for women (RSD from the middle of the 1980s). Moreover, none of the big five football leagues, nor the North American major leagues, had as weak SRCC as the Scandinavian football leagues for women (Tables 2 and 5 above). On this basis, an important question is whether Scandinavian football leagues for women have a competitive balance problem. Related to the uncertainty of outcome hypothesis, this means that further research should focus on whether weak competitive balance leads to low uncertainty of outcomes, which again leads to lower attendance in these leagues, ceteris paribus. The suggestion is hence related to attendance study applying fixed effects panel data model on match level, such as in Meier et al. (2016) for the top division for women in Germany. Moreover, research should also take into account prize structures in these leagues, as described in the competitive intensity concept regarding leagues in, for example, Kringstad and Gerrard (2004, 2007) and Andreff and Scelles (2015).

As shown in for example Kringstad and Gerrard (2007), the level of competitive balance differs between sports. The NFL in North America has an RSD on the level with the big five football leagues in Europe and the Scandinavian football leagues for men (Tables 2 and 4 above), while the NBA has a much weaker RSD. The empirical results in this paper show significant differences in competitive balance between the genders in Scandinavian football. What these examples may reveal is that the framework of a sport is relevant for its level of competitive balance. Therefore, looking at the hypothesis from Groot (2008) that higher number of goals or points on general basis implies weaker competitive balance, ceteris paribus, can be relevant for the context of this paper. Table 9 shows that the average number of goals in women’s matches is higher than for men’s matches in Scandinavian football. Analysis of reasons for this difference should be of interest for future research.

One possibility for differences in both competitive balance and goal scored in Scandinavian football may be related to the general physical differences between the genders. Will the average differences in, for example, height, strength and speed on a football field of approximately the same size as for men affect the relationship between the distribution of playing quality (talent) and the distribution of sporting outcome (competitive balance) in women’s football leagues? In other words, is the relative dispersion in sporting quality a stronger driver for competitive balance in women’s football than in men’s? If so, one will expect different match outcomes when there is an equal difference
in team quality in a men’s and a women’s match. Therefore, this may also be relevant in sports science, as such a relationship can indicate whether the tactics in male football are more effective for sporting outcomes than in women’s football.

The discussion above suggests that there may be differences between the distribution of playing quality (talent) in the top divisions in men’s and women’s football, and/or on the strength of the causality between the distribution of playing talent (quality) and competitive balance (distribution of sporting outcome). If one or both of these issues lead to significant weaker competitive balance in the women’s football leagues than for the men’s leagues, the governing bodies in football (FIFA, UEFA and national football associations) may have tools to compensate for this physical difference between genders in a competitive balance context. On an international basis, they may address the physical structure of the game, such as the size of the field and the goals. They could also address the length of the games. The latter would be an equivalent to, for example, cross-country skiing, where the argument for women to have shorter distances is not because of physical capacity, but because of the weaker competitive balance among the skiers. Related to the first issue mentioned above, the distribution of playing talent can be affected by market regulations (contrary to Rottenberg’s invariance proposition) on the league level. Here, for example, the distribution of team revenues is of particular interest. This distribution varies across leagues, such as between the NFL, where the Gini coefficient on the revenues, on average, was 0.068 from 2000–2007 (Fort, 2011), while it was much higher (between 0.27 and 0.28) in the top division in football for men in Norway over the seasons 2011 to 2013 (Kringstad & Olsen, 2016). The NFL has many market regulations (see, for example, Leeds & von Allmen, 2016), while the Norwegian top division for men is less regulated. One regulation of particular interest in the context of this paper is the distribution of league revenues (e.g. from media). Also, the women’s leagues are receiving revenues on league level, to be distributed among the teams in the league. The results from this paper suggest that if competitive balance is relevant for the football associations, women’s leagues should have a more equal distribution of these revenues than the men’s leagues.

6. Conclusions
To account for the complex nature of competitive balance, this paper has measured three dimensions of the concept. The conclusions from Scandinavian football are that except from championship concentration, the other dimensions have significantly weaker competitive balance among the women’s leagues. In other words, the women’s football leagues have less variation in team performance across seasons and larger within season differences in sporting outcome between the teams, compared to the men’s leagues. Further, among the women’s leagues, Denmark has weakest competitive balance. While the traditional measures of competitive balance do not differ significantly between Norway and Sweden, draw percent, and a new measure on average goal difference, indicate that the Swedish is best balanced. Summarized, the main finding in this paper is that competitive balance in women’s football in Scandinavia seems to be weak, and far weaker than for the men’s leagues.

Limitations of this study is that it does not measure the uncertainty of outcome hypothesis nor drivers for the level of competitive balance. On the other hand, the results and the limitations should be a base for future research. First, related to the uncertainty of outcome hypothesis, analysis of the effects on attendance is relevant. Here, demand study on match level is preferred, where demand can be both spectators at stadium and the number of TV viewers. Second and third, future research should be related to analyse possible drivers for the strong significant differences in competitive balance between women’s and men’s football leagues. This can be on basis of whether the distribution of playing quality (talent) among the teams is different between the genders, and if the causality between the distribution of sporting quality (talent) and competitive balance differs between men’s and women’s football. If competitive balance in women’s football matters, the results in this paper suggest more equal distribution of league level revenues than for the men’s football leagues.
18. As noted in Owen (2012) the RSD will hence be under-

16. RSD has been criticised when it comes to applying it for

15. Overviews of competitive balance measures can be

14. Berri et al. (2005) referred to the Gould hypothesis ap-

12. See Jennett (1984) for a description of the concepts of

11. See also Downward and Dawson (2000).

10. See, for example, Downward and Dawson (2000).

9. See also Borland and Macdonald (2003). They dis-

8. Retrieved from http://www.fifa.com/fifa-tournaments/

7. Retrieved from www.european-football-statistics.co.uk.

6. Retrieved from www.kvinnefotball.no.

5. Retrieved from resources.fifa.com/mm/document/foot-

4. Retrieved from www.forbes.com/sites/niallmccar-

3. Retrieved from www.forbes.com/sites/niallmccar-

2. Retrieved from http://resources.fifa.com/mm/document/foot-

1. Retrieved from http://www.fifa.com/fifa-tournaments/

Notes

1. Retrieved from www.fifa.com/womens-football/mission/

2. Retrieved from resources.fifa.com/mm/document/foot-

3. Retrieved from www.european-football-statistics.co.uk.

4. Retrieved from www.kvinnefotball.no.

5. Retrieved from www.european-association/football/index.html.

6. Retrieved from www.fifa.com/fifa-tournaments/archive/womensworldcup/index.html.

7. See also Borland and Macdonald (2003). They dis-

8. See also Downward and Dawson (2000).

9. See Jennett (1984) for a description of the concepts of

10. See, for example, Downward and Dawson (2000).

11. See also Downward and Dawson (2000).

12. See Jennett (1984) for a description of the concepts of

13. Retrieved from www.fifa.com/about-fifa/who-we-are/the-laws/

14. Berri et al. (2005) referred to the Gould hypothesis ap-

15. Overviews of competitive balance measures can be

16. RSD has been criticised when it comes to applying it for

17. Owen (2012, footnote 12) emphasised that Fort (2007)

18. As noted in Owen (2012) the RSD will hence be under-

19. All tests are done in Stata14 (www.stata.com).

20. Converting variables into the natural logarithm did not alter

21. Other internet sources: www.dbu.dk, www.svenskfot-

22. All tests are non-parametric.

23. All tests are non-parametric.

24. All tests are non-parametric.

25. Based on t-test with equal variance.

26. Based on non-parametric test.

27. The RSD for the Scandinavian women’s football

28. Based on non-parametric test.

29. Based on

30. t-test with equal variance.

31. Based on

32. All tests are non-parametric.

References

Alsiö, M., Frantz, A., Lindahl, J., & Persson, G. (Eds.). (2004). 100

Dr. Svenska fotbollsförbundet jubileumsbok 1904–2004. D.

Statistiken. Vällingby, Swe: Stroemberg Media Group.

Andreff, W., & Raballand, G. (2011). Is European football’s

future to become a boring game? In W. Andreff (Ed.),

Contemporary issues in sports economics. Participation

and professional team sports (pp. 131–167). Cheltenham:

Edward Elgar. https://doi.org/10.4337/9780857930385

Andreff, W., & Scelies, N. (2015). Walter C. Neale 60 years after:

Beyond competitive balance, the league standing effect
tested with French football data. Journal of Sports

Economics, 16(8), 819–834. doi:10.1177/1527002514566621

Berri, D., Brook, S., Fick, B., Fenn, A., & Vicente-Mayoral, R.

(2005). The short supply of toll people: Competitive

imbalance and the national basketball association.

Journal of Economic Issues, 39(4), 1029–1041. doi:

https://doi.org/10.1080/00213624.2005.11505868

Berri, D., & Krautmann, A. (2013). Understanding the WNBA on

and off court. In M. Leeds & E. S. Leeds (Eds.), Handbook

on the economics of women sports (pp. 132–155). Cheltenham:

Edward Elgar Publishing. https://doi.org/10.4337/9781849809399

Berland, J., & Macdonald, R. (2003). Demand for sport.

Oxford Review of Economic Policy, 19(4), 478–502.

https://doi.org/10.1093/oxrep/19.4.478

Cain, L., & Haddock, D. (2006). Measuring parity: Tying into the

idealized standard deviation. Journal of Sports Economics,

7(3), 330–338. doi:10.1177/152700250272937

Cairns, J., Jennett, N., & Sloane, P. (1986). The economics of

professional team sports: A survey of theory and evidence.

Journal of Economic Studies, 13, 3–80.

https://doi.org/10.1108/eb002618

Daly, G., & Moore, W. (1981). Externalities, property rights

and the allocation of resources in major league baseball.

Economic Inquiry, 19, 77–95. doi:10.1111/ecin.1981.19.issue-1

Dansk Boldspil-Union. (1998). Årsberetning 1997. Brøndby, DK:

Dansk Boldspil-Union.

Dansk Boldspil-Union. (1999). Årsberetning 1998. Brøndby, DK:

Dansk Boldspil-Union.

Dansk Boldspil-Union. (2000). Årsberetning 1999. Brøndby, DK:

Dansk Boldspil-Union.
Dobson, S., & Goddard, J. (2004). Revenue divergence and competitive balance in a divisional sports league. Scottish Journal of Political Economy, 51(3), 359–376. https://doi.org/10.1111/sjpe.2004.51.issue-3

Dobson, S., & Goddard, J. (2011). The economics of football (2nd ed.). Cambridge: Cambridge University Press. https://doi.org/10.1080/02331186.2015.1061032

Leeds, M., & von Allmen, P. (2016). The economics of sports (5th ed.). New York, NY: Routledge.

Meier, H. E., Konjer, M., & Leinwather, M. (2016). The demand for women's league soccer in Germany. European Sport Management Quarterly, 16(1), 1–19. doi:10.1080/16184742.2015.1109693

Noll, R. (1988). Professional basketball. (Studies in Industrial Economics Paper No 144). Stanford, California: Stanford University.

Owens, P. D. (2010). Limitations of the relative standard deviation of win percentages for measuring competitive balance in sports leagues. Economics Letters, 109, 38–41. doi:10.1016/j.econlet.2010.07.012

Owens, P. D. (2012). Measuring parity in sports leagues with draws: Further comments. Journal of Sports Economics, 13(1), 85–95. doi:10.1111/j.15270025.2011.00937.x

Owens, P. D. (2014). Measurement of competitive balance and uncertainty of outcome. In J. Goddard & P. Sloane (Eds.), Handbook on the economics of professional football (pp. 41–59). Cheltenham, UK and Northampton, MA: Edward Elgar.

Owens, P. D., & King, N. (2015). Competitive balance measures in the sports league: The effects of variation in season length. Economic Inquiry, 53(1), 731–744. doi:10.1111/ei.12102

Quirk, J., & El-Hodiri, M. (1974). The economic theory of a professional sports league. In R. Noll (Ed.), Government and the sports business (pp. 33–80). Washington, DC: Brookings Institution.

Quirk, J., & Fort, R. (1992). Pay dirt. The business of professional team sports. Princeton, NJ: Princeton University Press.

Robertson, S. (1956). The baseball players' labor market. Journal of Political Economy, 64(3), 242–258. https://doi.org/10.1086/257790

Schmidt, M., & Berri, D. (2003). On the evolution of competitive balance: The impact of an increasing global search. Economic Inquiry, 41(4), 692–704. https://doi.org/10.1111/ecin.12102

Scully, G. (1989). The business of major league baseball. Chicago, IL: University of Chicago Press.

Szymbanski, S. (2001). Income inequality, competitive balance and the attractiveness of team sports: Some evidence and a natural experiment from English soccer. The Economic Journal, 111, 69–84. https://doi.org/10.1111/ecoj.2001.111.issue-469

Treher, J., Levy, R., & Matheson, V. A. (2013). Gender differences in competitive balance in intercollegiate basketball. In M. Leeds & S. E. Leeds (Eds.), Handbook on the economics of women sports (pp. 251–268). Cheltenham: Edward Elgar Publishing. https://doi.org/10.4337/9781848039399

Woodridge, J. (2009). Introductory econometrics. A modern approach (4th ed.). Mason, OH: South-Western Cengage Learning.

Zimbaltis, A. (1992a). Baseball and billions. New York, NY: Basic Books.

Zimbaltis, A. (1992b). Salaries and performance: Beyond the Scully model. In P. Sommers (Ed.), Diamonds are forever: The business of baseball. Washington, DC: Brookings Institution Press.
