The More the Merrier?

A Study Measuring Relative Efficiency of Two Prediction Markets

Authors: Carl Anners & Stefan Saarm
Supervisor: Måns Thulin
2015-01-12
Department of Statistics
Uppsala University
Abstract

Our aim of this paper was to create a method for comparing the overall relative efficiency of a prediction market for the English football league Premier League and a prediction market for the Swedish football league Allsvenskan. The purpose of this was to see how the overall turnover of a prediction market affects the efficiency of it. We conclude that while the implied probability of the two markets on average corresponds well to the win frequency, the Premier League prediction market has statistically significant lower variation than Allsvenskan.

The method we created can also be used to test the relative prediction accuracy of any two prediction markets/bookmakers given enough observations.
# Contents

1 Introduction and Background \.......................................................... 3
  1.1 Prediction Markets \.......................................................... 3
  1.2 Association Football, Premier League and Allsvenskan \........... 3
  1.3 Purpose and Assumptions \............................................. 3

2 Market Inefficiencies on Betting Markets and Terminology \............ 4
  2.1 Favorite Long-Shot Bias \............................................. 4
  2.2 Preference Bias \....................................................... 5
  2.3 “Common knowledge” Pertaining to Home-Draw-Away \.............. 5

3 Data \............................................................................. 5
  3.1 Calculating the Average Belief \.................................... 6
  3.2 Normalization \........................................................... 7

4 Method \........................................................................ 7
  4.1 Test for Which Prediction Market is Best Overall \.................... 7
    4.1.1 Testing for Overall Bias \..................................... 7
    4.1.2 Testing Variation from the Theoretical Optimal Line \....... 8
  4.2 Tests for Possible Biases/Inefficiencies Outlined in Section 2. \.. 9
    4.2.1 Favorite-Long Shot Bias \.................................... 9
    4.2.2 Preference Bias \................................................ 10
    4.2.3 Home-Draw-Away \............................................. 10

5 Results \........................................................................ 10
  5.1.1 Results for Weighted Least Squares \............................ 10
  5.1.2 Results for Test of Variation from Optimal Line \.............. 11
  5.2 Results for Market Inefficiencies Investigation \.................... 15
    5.2.1 Preference Bias \................................................ 15
    5.2.2 Favorite-Long Shot Bias \.................................... 17
    5.2.3 Home, Draw, Away \........................................... 18

6 Discussion \................................................................. 18

7 References \................................................................. 21

Appendix \................................................................. 22
1 Introduction and Background

This thesis includes a method for measuring relative efficiency between two betting markets, a comparison using this method for Premier League and Allsvenskan and lastly a brief investigation of where this difference may arise. We start with formulating a purpose, defining terminology and explaining important concepts. We then proceed to describe our data and the method outlined. Lastly we discuss the results achieved, answer possible criticism and give suggestions for further research.

1.1 Prediction Markets

A prediction market is a speculative marketplace where speculators can buy and sell contracts that give a certain payout at the time of an uncertain future event. What one can bet on is everything from sports and politics to how many people will go watch a certain movie.

Supply and demand completely determine these markets and as the buyers and the sellers have to agree on a price it may be seen as their average belief of the probability of the outcome of the event. [Page and Clemen, 2004] Even though there are some play money markets in existence [Servan-Schriber et. al., 2004], on most prediction markets real money changes hands. The prospect of making money (or honor, in the case of play-money exchanges) should, at least theoretically, act as an incentive to gather information and make informed decisions.

Empirical research show that prediction markets often reflect the true probabilities quite accurately. [Tziralis and George, 2012] Prediction markets can therefore be an alternative to estimate the probability of future events and results, in contrast to more classical statistical methods such as polling. In the last few years, the use of prediction markets has grown in popularity and many firms now use them for this purpose, for example Google. [Tziralis and Tatsiopoulos, 2012]

1.2 Association Football, Premier League and Allsvenskan

Throughout this thesis, whenever we refer to football we are referring to what is called association football or soccer. [Szymanski, 2014] It’s important to point this out so that there will not be any confusion, since for example American football and Australian football are different altogether.

1.3 Purpose and Assumptions

In this thesis we are going to compare the Premier League with Allsvenskan. The Premier League is the top league of the English football league system and Allsvenskan is the top league of the Swedish football league system. We compare these two because the Premier League is a much larger prediction market than Allsvenskan (see section 3, Data) and we are interested in analyzing how this affects the

---

1 There are numerous claims of this to be found in almost any article pertaining to prediction markets. The cited article is a literature review of the subject.
prediction efficiency. Measuring the relative prediction efficiency between Premier League and Allsvenskan is the main purpose of our thesis. A secondary purpose is to see where the differences lie. The information flow in the Premier League is obviously higher than in Allsvenskan. An example of higher information flow is that the average player in the Premier League is better known to the public than in Allsvenskan. This would in turn make the Premier League easier to predict. However, in practice, a large prediction market will always have a higher information flow, therefore we simply see a higher information flow as a characteristic of a larger prediction market compared to a small.

2 Market Inefficiencies on Betting Markets and Terminology

It’s worth noting that the definition of efficiency we use differs slightly from the one used when talking about efficient market hypothesis. One definition by Eugene F. Fama, the often attributed founder of the efficient market hypothesis, goes as follows:

“… actual prices of individual securities already reflect the effects of information based both on events that have already occurred and on events which, as of now, the market expects to take place in the future.” [Fama, 1965]

So market efficiency says that the contract is priced taking all the public information available into account and one can therefore not outperform the market by factoring in this information in one’s model. This, however, does not state that the probability of the outcome is exactly reflected in the price, only that from the information available this is the best guess possible.

When we refer to efficiency what we mean is that the implied probability is equal to the actual probability of the outcome. There are actually two components of efficiency: first of all, a necessary condition is that the odds can be seen as an implied probability of the actual win percentage. Secondly, the variation of these implied probabilities from the actual win proportion should be as small as possible. Interestingly, in the literature pertaining to prediction markets the definition by Fama and the one we use are often equated. One example, out of many, can be found on page 11 in the article by Goddard and Asimakopoulos (2003) where they test for market efficiency with a linear probability model, a technique we will come back to in Section 4.1.1 but as a test for unbiasedness of the prediction.

From reading the literature we have singled-out three main areas where inefficiencies can occur. Our goal is not to make an in depth analysis of these inefficiencies, but if they exist in our data they might affect the efficiency of the prediction and are therefore worth examining.

2.1 Favorite Long-Shot Bias

Favorite long-shot bias is a phenomenon where bettors tend to overvalue “long-shots” and undervalue favorites. The definition of a long and a short odds is always relative but being a bit hand waving, in
this thesis a short odds is almost always lower than 2 (implied probability of 50 %). The explanation of this phenomenon is often given by risk-loving behavior as well as inaccurate estimation by the bettors. Andrikogiannopoulou and Papakonstantinou (2011) find that 6% of bettors systematically bet on biased long-shot and that 2% of bettors exploit this bias on bookmaker markets.

2.2 Preference Bias
Preference bias implies that fans are more willing to bet on their favorite team. There is empirical evidence for the existence of this in the Spanish league and that bookmakers take advantage of it. [Forrest & Simmons (2008)] If this bias exists in our dataset, we should see that the odds for the popular teams are overvalued. The five most popular teams for the Premier League in 2010 were Manchester United, Chelsea, Arsenal, Liverpool and Tottenham according to the research company Sport + Markt. For Allsvenskan there is to our knowledge no statistics showing which teams are the most popular. However, since the stadiums are more or less never filled to their maximum capacities, the attendance of a team should be a good estimation of the popularity of a team. Using this estimation, the five most popular teams in Allsvenskan for 2010 were Malmö FF, AIK, Helsingborgs IF, IFK Göteborg, and IF Elfsborg.

2.3 “Common knowledge” Pertaining to Home-Draw-Away
A well-known fact about football is that the home team has an advantage. This could possibly be due to having stronger support by your fans at home or being tired after travelling to another city. Whatever the reason may be, this common knowledge is supported by our dataset where the home team has an average win of about 46 % in both Premier league and Allsvenskan. Similarly, the average away win is about 28-29 % in both leagues.
Since the bettors are aware of this, one would expect that their beliefs about the probability of the outcome would follow different distributions, depending on the team being at home or away. Likewise, the actual outcomes would follow different probability distributions depending on if the team is home or away.

3 Data
The prediction market that we use for investigating our question is Betfair. It is a marketplace where odds for primarily sports are traded, but they also offer markets for politics and certain competitions, like Swedish Idol. Betfair make money by taking out a commission of roughly 5% of the winning bettor and it is therefore not a bookmaker but an exchange in the same vein as a stock exchange. By the arguments in Section 2.1, the average odds on Betfair can then be seen as a collective prognosis for the probability of an event.
Betfair provides historical odds on their webpage\(^2\). Six years’ worth of data, from 2008-10-19 to 2014-10-26, was collected for a total of 2301 Premier league matches and 1464 games from the Swedish Allsvenskan.

For every match the data included in our analysis is:

- Home and away team
- Scheduled off date and time
- Winner
- Number of bets and volume betted on different odds.
- Date and time of when the first bet on a specific odds was matched.
- Date and time of when the last bet on a specific odds was matched.

In our dataset we have about 2800 bets placed per match on average for the Premier League compared to about 400 bets placed per match on average for Allsvenskan. The average turnover per match for the Premier League is 600 000 GBP compared to an average turnover of 26 500 GBP per match for Allsvenskan.

Unfortunately, the data provided by Betfair has some problems. As seen in the list above, the data includes a time-stamp when the first bet was traded and when the last bet was traded for a certain odds. Betfair also include data on whether a bet was placed pre-event or in-play, but since 2008 an unmatched bet on Betfair can be matched in-play if the bettor ticks a box that the bet can go in-play and still be classified as pre-market. Since we want to investigate how well the belief of the actors on a prediction market corresponds to actual outcomes without information on what actually happens in the game, this poses a bit of a dilemma for us. Ignoring all bets that were both matched pre-market and in-play would lead to a loss of information, but including them all may also lead to false conclusions. However, the bettors have to make a conscious decision to allow a bet to be matched after the game starts and doing this is in general a very bad idea. Given this, apart from all bets that were classified as pre-market but had a first match after about one minute into the game, which were then filtered out, all these bets were included in the analysis.

3.1 Calculating the Average Belief

In Section 2.2 we say that we want to analyze Betfair’s collective prognosis. This collective prognosis has to be calculated since the data do not include an average for every outcome. We do this by weighing the odds for every outcome with how much money was bet on each odds in the following way:

\[
\text{weighted mean odds} = \frac{\sum_{i=1}^{n}(volume_i \times odds_i)}{\sum_{i=1}^{n}(volume_i)}
\]

(1)

\(^2\) http://data.betfair.com
Where \( n \) is the total number of odds for the selection.

We motivate using the amount of money bet on an odds as the weight with that the amount of money should be a good indication of what the market believes. An alternative would be to use the number of bets as the weight, but this would lead to, for example, two one dollar bets being worth more than one 1000 dollars bet. This does not seem very reasonable as betting more money on something should be an indication of a stronger belief. Nonetheless, using the amount of bets compared to the amount of money as weight when calculating the mean odds does not differ. So using one or the other will not give us drastically different results.

The implied probability is just the inverse of the weighted mean odds,

\[
\text{implied probability} = \frac{1}{\text{weighted mean odds}} \quad (2)
\]

3.2 Normalization

As a consequence of our averaging, the Betfair probability prognoses for one match (1, X, 2) often does not sum to one. This seems counter-intuitive since the probability that a match results in either home win, draw or away win should be one. It would be possible to normalize the probabilities so they sum up to one to correct for this, but we have decided against it. We motivate this with that we are evaluating individual beliefs and not the probability for the whole match.

4 Method

To test how well the implied probabilities reflect the actual outcome for the different markets a few different tests and statistical techniques are employed. Firstly we do an overall test of the overall relative efficiency (Section 4.1), then we investigate the biases outlined in Section 2.

4.1 Test for Which Prediction Market is Best Overall

The main purpose of this thesis is testing the difference in efficiency between Premier League and Allsvenskan and to do this we will primarily use two tests. The first test, described in Section 4.1.1, is simply to see if there is any clear bias in the bettors’ beliefs in Allsvenskan and the Premier League prediction markets. That is, are the belief of the bettors’ aligned on an optimal line, a theoretical line with intercept zero and slope one, which is a necessary, but not sufficient, condition that odds can be seen as implied probabilities. The second test, described in Section 4.1.2, is employed to see whether there is any significant difference in variation from this optimal line.

4.1.1 Testing for Overall Bias

Testing whether there is an overall bias or not is done by regressing the outcomes on the implied probabilities for both Allsvenskan and Premier League using weighted least squares [Wooldridge, 2012] in the following way:
\[
\frac{\text{Result}_i}{\sqrt{h_i}} = \frac{\alpha}{\sqrt{h_i}} + \left( \beta \frac{\text{implied probability}}{\sqrt{h_i}} \right) + \frac{\varepsilon}{\sqrt{h_i}} 
\]  
(3)

Where \( h_i \) is the weight, \( \text{Result}_i \) is one if the outcome is a win and zero if not and \( \text{implied probability} \) is the weighted mean odds (See Section 3.1) transformed into a probability. The weights, \( h_i \), are obtained using the fitted values from regular OLS:

\[
h_i = \hat{f}_i(1 - \hat{f}_i) 
\]  
(4)

With no systematic errors, the \( \alpha \) should not be statistically different from zero and \( \beta \) should not be statistically different from one. Our null and alternative hypothesis for this test is therefore given by:

\[
H_0: (\alpha, \beta) = (0, 1) 
\]

\[
H_1: (\alpha, \beta) \neq (0, 1) 
\]

4.1.2 Testing Variation from the Theoretical Optimal Line

To see how implied probabilities relate to actual win proportions we divide the implied probabilities into clusters by minimizing the within-group variance using one-dimensional Ck-means clustering [Wang and Song, 2011] and comparing them to the mean of the actual win frequency corresponding to the clusters [Andrikogiannopoulou and Papakonstantinou, 2011] [Servan-Schreiber et. al., 2004]. This leads to the question: how many clusters are appropriate? The Ck-means algorithm\(^3\) developed by Wang and Song can decide what number of clusters is optimal by using Bayesian information criterion. So, we simply let the algorithm decide between 1-100 clusters for each league separately, and then compare them using the method below.

After dividing the odds into clusters, we compare the mean squared error of the Allsvenskan clusters and the Premier League clusters using a modified F-test. The mean squared error is calculated from the optimal line (intercept of zero and slope of one), i.e. if the MSE is zero, the outcome is perfectly predicted. Since the groups are not of the same size, we weigh the squared error with the group size when calculating the two MSEs, or rather Weighted Mean Square Error.

\[
WMSE = \frac{\sum_{i=1}^{k} n_i * (\text{proportion win}_i - \text{implied probability}_i)^2}{N} 
\]  
(5)

Where \( k \) is the number of groups, \( n_i \) is the number of data points in a group and \( N \) is the total number of data points.

\(^3\) Their R implementation, Ckmeans.1d.dp, can be downloaded from: http://cran.r-project.org/web/packages/Ckmeans.1d.dp/index.html
Since the MSEs are weighted, they are not chi-square distributed and we cannot use a regular F-test directly. The distribution of each clusters weighted squared error under the null hypothesis will follow a gamma distribution with parameters that depend on the number of observations\textsuperscript{4}. Because of the complexity of the distribution of the weighted MSE under the null hypothesis (a sum of many different gamma distributions with different parameters) it is estimated using the Monte Carlo method (see Simulated distributions in Appendix). Analogous to the F-test, the ratio of the MSEs' will follow some distribution, so we can compare the observed WMSE-ratio with the simulated joint distribution and thus compute the probability of observing such a ratio or more extreme, given that the null hypothesis (the MSE's of the different markets is the same) is true. The weighted MSE-ratio is given by:

\[
F^* = \frac{\text{WMSE}_\text{Allsvenskan}}{\text{WMSE}_\text{Premier League}}
\]  

(6)

Given significance of this test we would see that the Premier League’s clusters have a lower variation from the optimal line than Allsvenskan’s. The null and alternative hypothesis for this test is therefore:

\[H_0: \text{Allsvenskan and the Premier Leagues has the same variation from the optimal line}\]

\[H_1: \text{Premier League has a smaller variation than Allsvenskan from the optimal line}\]

4.2 Tests for Possible Biases/Inefficiencies Outlined in Section 2.

Once we have determined whether there is an overall difference between the Premier League and Allsvenskan we will proceed to test the areas where biases/inefficiencies described in Section 2.4 might arise. These are Favorite Long-shot bias, Preference bias and “Common Knowledge” pertaining to home-draw-away. Testing these biases/inefficiencies is not the main purpose of our thesis, but nonetheless, it is interesting to see if we can determine where the Premier League is better.

4.2.1 Favorite-Long Shot Bias

To test for Favorite-Long shot bias we regress the difference between win frequency and implied probability for each cluster on the implied probability of each cluster, using a standard OLS in the following way:

\[
\text{Difference}_i = \alpha + \beta * \text{implied.probabiliy}_i
\]

(7)

Where i are the clusters \{1, 2, 3,…\}

\textsuperscript{4} This follows directly from linear combinations of Moment Generating Functions:

\[M_{nx}(t) = E[e^{t(x)}] = M_x(tn)\]

And the fact that the unweighted squared error is \(\chi^2\) distributed with one degree of freedom.
Favorite-Long shot bias implies that high probabilities are undervalued and low probabilities are overvalued, this means that the coefficient should be positive given that the bias exists in our data. Our null and alternative hypotheses are therefore given by the following:

\[ H_0: \beta \leq 0 \]
\[ H_1: \beta > 0 \]

4.2.2 Preference Bias
To find indication of preference bias we will simply look at the five most popular teams in the Premier League and in Allsvenskan to see if there is an overestimation of their implied probability compared to the actual win frequency. If there is an indication of this we can perform tests to see whether the overestimation is statistically different from zero.

4.2.3 Home-Draw-Away
As seen in Section 5, we do not find any indication that the variance is dependent on how high or low the implied probability is. Given this, to test if there is any inefficiency pertaining to if the selection is home, draw or away is tested by a series of t-tests using Bonferroni-correction [Rice et al, 2008] to adjust for the multiple testing problem. The Bonferroni-correction is given by:

\[ \alpha_{(per\ comparison)} = \frac{\bar{\alpha}}{k} \quad (8) \]

Where k are the number of t-tests done.

It should however be noted that the Bonferroni-correction increases the chance of making a Type-II error.

5 Results
5.1.1 Results for Weighted Least Squares
As can be seen in Table 1 and Table 2, using any of the standard cutoff values, for neither Allsvenskan nor the Premier league the null hypothesis detailed in Section 4.1.1 is rejected. We find no evidence of any overall systematic error.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std.Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.01119</td>
<td>0.01454</td>
<td>0.442</td>
</tr>
<tr>
<td>vwap.implied.prob</td>
<td>1.03102</td>
<td>0.04290</td>
<td>0.470</td>
</tr>
</tbody>
</table>
Table 2: WLS Output Premier League. The p-values are for intercept = 0 and slope = 1.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.004411</td>
<td>0.008782</td>
<td>0.616</td>
</tr>
<tr>
<td>vwap.implied.prob</td>
<td>1.014658</td>
<td>0.025590</td>
<td>0.567</td>
</tr>
</tbody>
</table>

5.1.2 Results for Test of Variation from Optimal Line
As can be seen in Table 3 and Table 4, Ck-means clustering results in five clusters for both the Premier League and Allsvenskan. Figure 1 and 2 show the limits of these clusters plotted to make it obvious which odds belongs to what cluster. Figure 3 show the mean of the implied probability of each cluster plotted against the win frequency associated with this cluster. The plotted line is given by intercept 0 and slope 1 which shows where the mean of the implied probability should have been given that it had a perfect prediction.

In Figure 4 the difference between win frequency and implied probability (vertical axis) is plotted against implied probability (horizontal axis). A visual inspection indicate higher deviation for Allsvenskan. This is confirmed using the F-type test by using the data in Table 1 and Table 2 as we proceed to calculate the WMSE and $F^*$-ratio for both leagues using the method outlined in Section 4.1.2. The WMSE is $3.61 \cdot 10^{-4}$ for Allsvenskan and $4.299 \cdot 10^{-5}$ for the Premier League. The ratio of these is 3.166 and comparing this to the 1.055055 rejection region given by the simulated joint distribution, we clearly reject the null hypothesis stated in Section 4.1.2. We find, at the approximately 0.001-significance level, evidence that the Premier League has a lower variation from the optimal line than Allsvenskan.

Table 3. Size and range of implied probabilities in the clusters calculated by the Wang and Song implementation of the Ck-means algorithm and associated win proportion for Allsvenskan.

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>Count</th>
<th>Mean implied probability</th>
<th>Minimum implied probability</th>
<th>Maximum implied probability</th>
<th>Win frequency</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>737</td>
<td>0.1578693</td>
<td>0.04426737</td>
<td>0.2141328</td>
<td>0.1682497</td>
<td>-0.01038</td>
</tr>
<tr>
<td>2</td>
<td>2014</td>
<td>0.2720647</td>
<td>0.21505376</td>
<td>0.3289474</td>
<td>0.2601787</td>
<td>0.011886</td>
</tr>
<tr>
<td>3</td>
<td>730</td>
<td>0.3875946</td>
<td>0.33003300</td>
<td>0.4484305</td>
<td>0.4000000</td>
<td>-0.01241</td>
</tr>
<tr>
<td>4</td>
<td>567</td>
<td>0.5120664</td>
<td>0.45045045</td>
<td>0.5882353</td>
<td>0.5255732</td>
<td>-0.01351</td>
</tr>
<tr>
<td>5</td>
<td>341</td>
<td>0.6670097</td>
<td>0.59171598</td>
<td>0.8547009</td>
<td>0.6598240</td>
<td>0.007186</td>
</tr>
</tbody>
</table>
Table 4. Size and range of implied probabilities in the clusters calculated by the Wang and Song implementation of the Ck-means algorithm and associated win proportion for Premier League

<table>
<thead>
<tr>
<th>Cluster number</th>
<th>Count</th>
<th>Mean implied probability</th>
<th>Minimum implied probability</th>
<th>Maximum implied probability</th>
<th>Win frequency</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1477</td>
<td>0.1342826</td>
<td>0.02543882</td>
<td>0.2028398</td>
<td>0.1279621</td>
<td>0.00632</td>
</tr>
<tr>
<td>2</td>
<td>3081</td>
<td>0.2717915</td>
<td>0.20325203</td>
<td>0.3472222</td>
<td>0.2765336</td>
<td>-0.00474</td>
</tr>
<tr>
<td>3</td>
<td>1113</td>
<td>0.4239104</td>
<td>0.34843206</td>
<td>0.4975124</td>
<td>0.4132974</td>
<td>0.01061</td>
</tr>
<tr>
<td>4</td>
<td>716</td>
<td>0.5735607</td>
<td>0.50000000</td>
<td>0.6578947</td>
<td>0.5726257</td>
<td>0.00094</td>
</tr>
<tr>
<td>5</td>
<td>516</td>
<td>0.7467393</td>
<td>0.66225166</td>
<td>0.9009009</td>
<td>0.7558140</td>
<td>-0.00908</td>
</tr>
</tbody>
</table>

Figure 1: Plot showing the different clusters implied probability limits for Allsvenskan
Figure 2: Plot showing the different clusters implied probability limits for Premier League.
Figure 3: Cluster mean implied probability plotted against the actual win frequency for Premier League and Allsvenskan. Red dots are Allsvenskan and blue dots Premier league.
5.2 Results for Market Inefficiencies Investigation

5.2.1 Preference Bias

In Figure 5 we show how overrated and underrated the teams in the premier league are in our dataset. The most extreme values, Middlesbrough, Crystal Palace, Q.P.R, Portsmouth and Blackpool can be explained by them being relegated from the Premier league so we simply do not have enough observations. It is also clear from the table that the most popular teams in the premier league, Manchester United, Chelsea, Arsenal, Liverpool and Tottenham do not show a trend of being overvalued. Liverpool, Chelsea and Arsenal are overvalued but Manchester United and Tottenham are undervalued. Therefore we do not conduct a t-test to test for preference bias since we can rule it out directly by analyzing the graph.

Figure 4: The difference between win frequency and implied probability (vertical axis) is plotted against implied probability (horizontal axis). Red dots are for Allsvenskan and blue dots are for Premier league.
Figure 5: Difference in average implied probability and win percent for Premier League teams.

In Figure 6 we see the most overrated and underrated teams in Allsvenskan. Similarly as in Premier League, the most overrated teams are also the teams that have been relegated in our time period, so there are simply not enough observations to analyze them. It is also clear from the table that the most popular teams, Malmö FF, AIK, Helsingborgs IF, IFK Göteborg, and IF Elfsborg show no signs of being generally overvalued. Malmö FF, AIK and Helsingborgs IF are undervalued and IFK Göteborg and IF Elfsborg are overvalued. Therefore we rule out a preference bias in Allsvenskan.
5.2.2 Favorite-Long Shot Bias

The results for the Favorite-Long shot bias investigation is given in Table 5 and Table 6. The results show no indication of Favorite-Long shot bias in neither the Premier League nor Allsvenskan. We therefore do not reject the null hypothesis stated in section 4.2 for the tests.

Table 5. Favorite-Long shot examination for Premier league

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std.Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>0.008039</td>
<td>0.007908</td>
<td>0.384</td>
</tr>
<tr>
<td>implied prob</td>
<td>-0.016808</td>
<td>0.016430</td>
<td>0.382</td>
</tr>
</tbody>
</table>

Table 6. Favorite-Long shot examination for Allsvenskan

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std.Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>-0.007312</td>
<td>0.015021</td>
<td>0.660</td>
</tr>
<tr>
<td>implied prob</td>
<td>0.009685</td>
<td>0.034348</td>
<td>0.796</td>
</tr>
</tbody>
</table>
5.2.3 Home, Draw, Away

Not surprisingly, the implied probability does not exactly correspond to the actual win frequency *(Table 7)*. However, by the results from the t-test shown in *Table 7* and *Table 8* we do not find any indication that there is any statistically significant difference in any of the cases.

*Table 7. Implied probability and win frequency when the selection is home, draw and away.*

<table>
<thead>
<tr>
<th>Selection</th>
<th>Premier League</th>
<th>Allsvenskan</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Implied probability</td>
<td>Win frequency</td>
</tr>
<tr>
<td>Home</td>
<td>0.4582787</td>
<td>0.4632768</td>
</tr>
<tr>
<td>Draw</td>
<td>0.2508199</td>
<td>0.2603216</td>
</tr>
<tr>
<td>Away</td>
<td>0.2919985</td>
<td>0.2764016</td>
</tr>
</tbody>
</table>

*Table 8. Results of t-test of difference between win frequency and average implied probability for home, draw and away for Premier league.*

<table>
<thead>
<tr>
<th>Selection</th>
<th>Premier League</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Home</td>
<td>Premier League</td>
<td>Draw</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Premier League</td>
<td>Away</td>
</tr>
<tr>
<td>df</td>
<td>3002.95</td>
<td>2359.787</td>
<td>2980.143</td>
</tr>
<tr>
<td>t</td>
<td>0.447</td>
<td>1.0318</td>
<td>-1.5588</td>
</tr>
<tr>
<td>p</td>
<td>0.6549</td>
<td>0.3023</td>
<td>0.1191</td>
</tr>
</tbody>
</table>

*Table 9. Results of t-test of difference between win frequency and average implied probability for home, draw and away for Allsvenskan.*

<table>
<thead>
<tr>
<th>Selection</th>
<th>Allsvenskan</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Home</td>
<td>Allsvenskan draw</td>
<td>Allsvenskan away</td>
</tr>
<tr>
<td>df</td>
<td>1731.754</td>
<td>1480.339</td>
<td>1707.268</td>
</tr>
<tr>
<td>t</td>
<td>0.7151</td>
<td>-1.2603</td>
<td>0.2541</td>
</tr>
<tr>
<td>p</td>
<td>0.4746</td>
<td>0.2078</td>
<td>0.7995</td>
</tr>
</tbody>
</table>

6 Discussion

In this thesis we have examined whether there is any difference in relative prediction efficiency between Betfairs’ prediction market for the Premier League and Allsvenskan. Overall, the implied probability on the Betfair prediction markets for both Allsvenskan and Premier league corresponds well to the actual win frequency. However, the Premier league market is significantly better, where better is defined to be having lower variance. Given our assumption that both markets are similarly difficult to predict, what differentiates them is the size of the turnover. Unfortunately, the public data published by Betfair does not differentiate between individuals, but a reasonable guess is that larger turnover is highly correlated with many speculators, so the higher efficiency may be attributed to the market having a large number of participants.
As noted in Section 2, earlier research has found evidence of biases in certain areas of betting markets. For example, Forrest and Simmons (2008) found a preference bias in the Spanish league. In our investigation, however, we do not find any indications of any of these biases in neither Allsvenskan nor the Premier league.

One would expect that preference bias is something that at least would exist in a betting market for a league like Allsvenskan, whose bettors most likely are Swedish people who feel strongly about one team in Allsvenskan, compared to the Premier League which has many foreign bettors who do not have a local team they feel strongly about. Since this bias has been observed by researchers in the past, it is somewhat surprising that we do not find any evidence of this. It might be the case that in today’s prediction markets more and more savvy professional bettors exploit this inefficiency as it is so well documented.

Just as for preference bias, we do not find any indication of neither favorite-longshot bias nor inefficiencies pertaining to home-draw-away. The lack of these biases may be explained in the same way as the lack of preference bias: such obvious biases might already be exploited by professional bettors. Also, placing a bet on Betfair is more complex than, say, at a traditional bookmaker and could possibly scare away many recreational bettors and left are the more knowledgeable ones who bet to a lesser extent according to feelings.

Possible criticism to our choice of comparing Premier League and Allsvenskan to represent a large and a small prediction market is that Allsvenskan is potentially harder to predict because of the different nature of the two leagues. The difference in efficiency between Premier League and Allsvenskan would then not be due to a smaller amount of bettors giving a lower efficiency, but simply because Allsvenskan is harder to predict. If Allsvenskan had the same amount of bettors as Premier League we would then still get the result that the Premier League prediction market is more accurate. The data support this to a degree where the overall win proportion of the four top teams in the premier league from 2008-10-19 to 2014-10-26 is around 60% (Manchester United, Chelsea, Manchester City and Arsenal), while the top teams in Allsvenskan for this period has win proportion of around 50% (Malmö FF, AIK, Elfsborg and Helsingborg). It could therefore be argued that it is clearer in Premier League which team is better. However, when looking at the individual seasons, the win proportion of the top teams in Allsvenskan is generally not lower than for the Premier League. This is possibly due to foreign (or other domestic) teams buying the best players from Allsvenskan, so if a team performs well, it will lose its best players. We therefore make the intuitive assumption that bettors take into account that the best team last season might not be the best team the season after that if that team has sold its star players to other clubs.
The method we have outlined is obviously not limited to comparing Premier league and Allsvenskan, but can be used to analyze the relative efficiency of any two prediction markets. It however requires a large number of observations.

To advance research on prediction markets, there are data which for example solve our problem with time stamps. These data however has to be bought from Fracsoft Ltd, but it includes data on exactly what time a bet was placed. This could then be used to analyze how the efficiency of the prediction market changes over time and how the relative efficiency changes with more and more bettors.

In conclusion, it is especially interesting to note that even though Allsvenskan has about twenty times lower turnover than the Premier League, it still has a good efficiency. This shows the usefulness of even small prediction markets and it is hard to think of a more powerful way to make a prognosis of a future outcome than through prediction markets.
7 References


SPORT+MARKT Football Top 20 2010.


Appendix

As the distribution of the weighted mean squared errors is unknown, the distribution is simulated using the Monte Carlo method.

Each cluster consists of a number of observations, $n_i$. An unweighted cluster’s squared error, $(\text{proportion win} - \text{implied probability})^2$ is assumed to follow a $\chi^2$ distribution with one degree of freedom. The weighted squared error will then follow a gamma distribution with $\alpha$ and $\beta$ of $n/2$ and $2 \cdot n$, respectively.

For each cluster, $k$, 100 000 samples are drawn from the gamma distribution associated with the number of observations in the cluster. The matrix below illustrates how the samples are drawn, where $nX$ is a realization from the respective gamma distribution:

$$
\begin{pmatrix}
    n_1X_{1,1} & \ldots & n_1X_{1,100000} \\
    \vdots & & \vdots \\
    n_kX_{k,1} & \ldots & n_kX_{k,100000}
\end{pmatrix}
$$

(9)

Each column of the matrix is summed and we are left with a simulated distribution of the weighted mean squared error under the null hypothesis. In Figure 7 and Figure 8 we see visually the distributions for Premier league and Allsvenskan, respectively, using five clusters.

The F-type distribution is simply constructed by dividing each of these simulated distributions. The ratio for the above example is visualized in Figure 9.
Figure 7. Simulated distribution. Sum of weighted chi-squared random variables. Weights from the number of observations per cluster for Premier league.
Figure 8. Simulated distribution. Sum of weighted chi-squared random variables. Weights from the number of observations per cluster for Allsvenskan.
Figure 9. Simulated distribution. Ratio of figure 7 and figure 8.