Bookmakers in Agreement: A Comment on Deschamps and Gergaud

Elliot Christophers

Abstract

Deschamps and Gergaud (2007) suggest that the disagreement between bookmakers, when pricing a soccer match outcome, has a positive effect on the financial yield from betting on that outcome. This signal would be useful for prospective bettors, helping them increase their profits. We scrutinise the authors' methodology and conclusion against a larger dataset, and find their claims unsubstantiated.

1 Introduction

In their 2007 paper "Efficiency in Betting Markets: Evidence from English Foootball", Deschamps and Gergaud (DG) present a broad array of facts and features of betting data from English association football (soccer) leagues, between 2002 and 2006. Bookmaker margins consistently decreased across time. The famous favourite-longshot bias was identified. A draw bias was discovered, where betting on the draw outcome gave higher returns than betting for home or away wins, on average. Most notably, for our purposes, they found a positive relationship between the disagreement displayed by bookmakers in their collective pricing for an outcome, and the mean returns that a bettor would receive when betting on these outcomes. The intuition is clear: if, controlling for the price bet at, there is larger uncertainty between bookmakers, then the most exposed opinion is increasingly likely to be incorrect. Their finding is relevant for prospective sports bettors, as it provides a clear signal to act and trade upon.

In this paper, we investigate this finding across a different dataset. Accessed from football-data.co.uk, we analyse 23 seasons' worth of soccer games from 22 of the biggest leagues across Europe. From a total of 174,475 games spanning the 2001/2002 season through 2023/2024, the dataset contains the result of each game and match-outcome odds (home win, draw, away win) from as many as thirteen individual bookmakers. We find no evidence that the DG uncertainty-returns effect exists. Having quantified the favourite-longshot bias, we demonstrate how it confounds the relationship, where returns to uncertainty are actually returns to the favourite-longshot bias, due to heteroscedasticity in the DG regression. We seek to improve upon their methodology, but cannot nullify the bias' confounding impact. Therefore, we apply an entirely different mechanism, totally controlling the bet price and hence isolating the impact of bookmaker uncertainty on returns. There is no significant effect whatsoever, leading us to conclude that the DG results are, unfortunately, unfounded.

2 Favourite-Longshot Bias

The favourite-longshot bias in sports betting (see, for instance, Law and Peel (2002), Ottaviani and Sørensen (2008), Williams (2001) for discussion), is the observation – consistent in data across sports, leagues and time – that financial returns to bets placed at longer odds (which have low probabilities of realisation and are therefore underdogs) are much lower than the corresponding returns when betting on favourites with short odds. The reason for this persistent phenomenon is likely a combination of multiple effects. Bookmakers may place extra margin on long odds to protect themselves from large liabilities, especially given how punters enjoy the thrill of backing underdogs (Hurley and McDonough, 1996). Moreover, since odds are inverses of probabilities, if bookmakers are uncertain about the true probabilities, underestimation at low probabilities means an inversely greater error in the odds, which can be mitigated with large margins on low probabilities. Alternatively, this phenomenon may be a symptom of bookmakers negating the insider knowledge of certain informed bettors (Shin, 1993).

Regardless of origin, we here reproduce this bias. For each outcome in the soccer win market, we take the longest (most profitable) odds from each of m games, indexed i. For each outcome j, we sort these odds o_{ij} in ascending order such that $o_{ij} \leq o_{(i+1)j}$ and create ten bins G_{gj} of equal size, the first group containing the first decile of odds, and so on. Formally,

$$G_{gj} = \{(i,j) \mid o_{(q_{g-1})j} \leq o_{ij} < o_{(q_g)j}\} \tag{1}$$

where $q_g = \lfloor g\frac{m}{10} \rfloor$ for $g = 1, \ldots, 10$ and $q_0 = 0$. The mean odds for each outcome-group $\bar{o}_{gj}$ is then calculated

$$\bar{o}_{gj} = \frac{1}{|G_{gj}|} \sum_{(i,j)\in G_{gj}} o_{ij} \tag{2}$$

where $|G_{gj}|$ is the number of elements in the g-th bin for outcome j. Then, if we let $\boldsymbol{I}(i,j)$ be the indicator function

$$\boldsymbol{I}(i,j) = \begin{cases} 1 & \text{if the } i\text{-th match ends with outcome } j \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

then we can calculate the mean returns to unit stakes at odds in each bin

$$\bar{r}_{gj} = \frac{1}{|G_{gj}|} \sum_{(i,j)\in G_{gj}} \boldsymbol{I}(i,j) o_{ij} - 1 \tag{4}$$

$r_{ij} = o_{ij}$ if $\boldsymbol{I}(i,j) = 1$ and -1 otherwise.

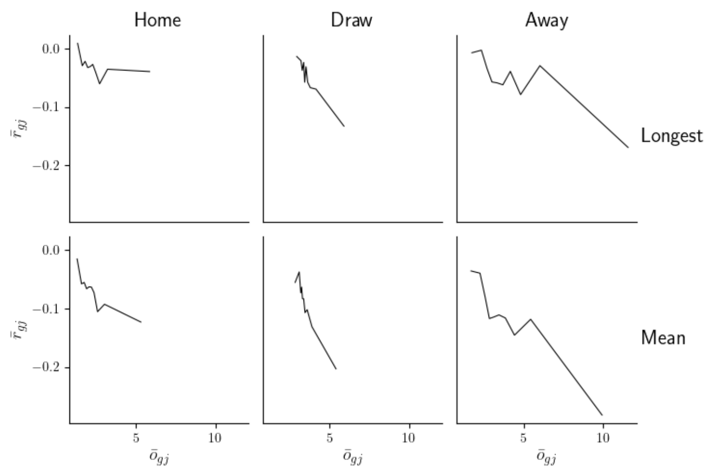

Figure 1: The Favourite-Longshot Bias: returns to 10 odds bins, by outcome, by longest and mean odds. As odds lengthen, returns for the bettor deteriorate. Scale is the same for all six subplots, for comparability.

As mentioned, we do this for the longest odds from each match. Figure 1 plots the $\bar{r}_{gj}$ versus the corresponding $\bar{o}_{gj}$, for each outcome j. It also illustrates the same relationship, but instead of taking the longest odds from each match, considering the arithmetic mean odds from each match (that is, o_{ij} is the mean of all odds offered in the dataset for outcome j in match i). There are thus six subplots, all having the same scale on both axes, to ensure that they are commensurable. Evidently, as the mean odds increase for an outcome, the returns bettors receive decrease. This is true across outcomes as well as both odds categories (mean and longest). That returns are uniformly negative even at longest odds demonstrates the presence of margin on all outcomes and lengths of odds. Clearly, then, the favourite-longshot bias exists even in our large dataset.

3 Returns to Bookmaker Uncertainty: Replication

The more novel finding (beyond similarly reproducing the favourite-longshot bias) from DG is their identification of increasing returns to disagreement between bookmakers. Their methodology has two steps: first, define bookmaker disagreement as the mean absolute deviation of the odds offered on an outcome from their mean, and second, regress this mean absolute deviation on the mean odds, across all matches, from which the residual will quantify any disagreement above expected, for that magnitude of odds. Their dataset consisted of games with exactly six bookmakers. To increase the veracity of the replication, we imitate this and take only the 45,693 games in our dataset with odds from exactly six distinct bookmakers.

Let o_{hij} be the odds offered by the n_{ij} bookmakers indexed h on the i-th game, outcome j. With this notation, the o_{ij} from above for a particular game-outcome pair (I, J) would then be $o_{ij} = \max(\{o_{hij} \mid (i,j) = (I,J)\}) = l_{ij}$ in the longest-odds case, while the mean odds case would have $o_{ij} = \frac{1}{n_{ij}} \sum_{h=1} o_{hij} = \bar{o}_{ij}$. We then have that the mean absolute deviation, as calculated by DG, is given by

$$Mad_{ij} = \frac{1}{n_{ij}} \sum_{h=1} \mid o_{hij} - \bar{o}_{ij} \mid \tag{5}$$

where, in this case, n_{ij} is constrained to equal 6. They run the regression

$$Mad_{ij} = \beta_0 + \beta_1 \bar{o}_{ij} + \beta_2 \bar{o}_{ij}^2 + \varepsilon_{ij} \tag{6}$$

The reasoning is that, the larger is $\bar{o}_{ij}$, the larger, naturally, would be Mad_{ij}. By regressing the latter on the former, we can control for this magnitude inconsistency. We can then classify a match-outcome pair by its residual ε_{ij}, which quantifies the bookmaker disagreement above expected, given $\bar{o}_{ij}$. As above, we construct bins G_{gj}, this time by ordering the residuals for each outcome ε_{ij}, such that $\varepsilon_{ij} \leq \varepsilon_{(i+1)j}$, giving five bins, where the first contains the first quintile of residuals, and so on. Similar to (1),

$$G_{gj} = \{(i,j) \mid \varepsilon_{(q_{g-1})j} \leq \varepsilon_{ij} < \varepsilon_{(q_g)j}\} \tag{7}$$

where $q_g = \lfloor g \frac{m}{5} \rfloor$ for $g = 1, \ldots, 5$ and $q_0 = 0$. We can then similarly calculate mean residuals in each group

$$\bar{\varepsilon}_{gj} = \frac{1}{|G_{gj}|} \sum_{(i,j) \in G_{gj}} \varepsilon_{ij} \tag{8}$$

while the calculation for $\bar{r}_{gj}$ is the same as (4), only replacing, as discussed above, o_{ij} with $\bar{o}_{ij}$.

Figure 2 shows, in the upper panel, the results obtained by DG on their dataset, presented in their Table 8. They categorised the bookmaker disagreement as we have, but chose not to give the mean residual in each group, instead labelling disagreement as “Very Low” through “Very High”, as demonstrated on the horizontal. They find a consistently positive relationship, where increased bookmaker uncertainty, as measured by the residual from regressing the magnitude of uncertainty on the match-outcome’s odds, leads to increased returns when betting. Our results, however, show nothing of the sort. Instead, returns are negatively correlated with the magnitude of the residual. Given our knowledge of the favourite-longshot bias, a possible explanation would be that, in our dataset, the magnitude of the residual increases with the magnitude of the odds: if there is heteroscedasticity of this kind, then betting at “very high” or “very low” uncertainty is actually betting, simply, at long odds, and therefore, due to the favourite-longshot bias, returns will be inferior. Figure 3 confirms that this is, indeed, the case, showing, for

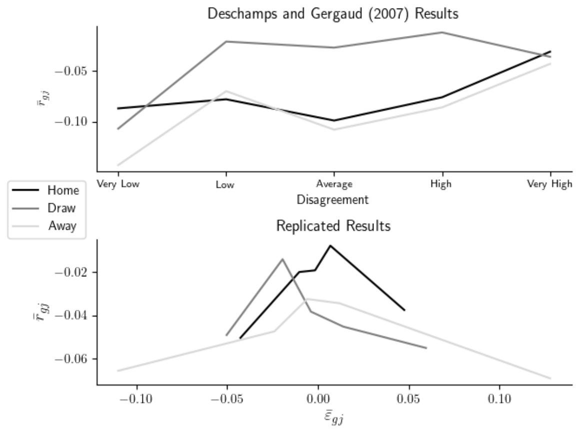

Figure 2: Returns to unit stakes versus bookmaker disagreement, "controlling" for odds. Table 8 in Deschamps and Gergaud (2007) in the upper panel, our replicated results, via (8) and adjusting (4), in the lower panel.

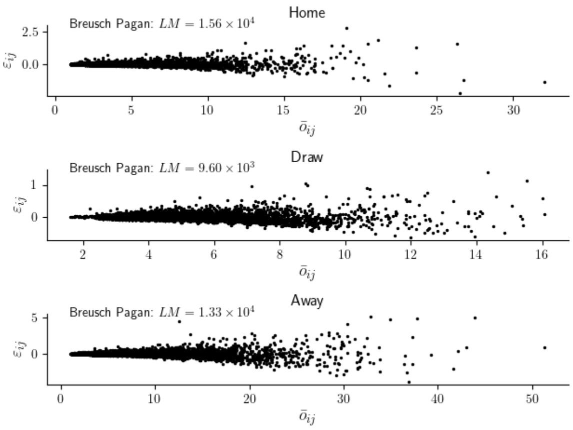

Figure 3: Residuals versus the independent variable in the regression from (6), with test statistics from the Breusch-Pagan test. Both demonstrate the extreme heteroscedasticity in the regression.

each outcome, ε_{ij} versus $\bar{o}_{ij}$. The Breusch-Pagan test statistic LM (Breusch and Pagan, 1979), with alternative hypothesis of heteroscedasticity, is given. We have that $LM \sim \chi^2$, with two degrees of freedom (two independent variables in (6)), giving critical values for significance levels $\alpha = 0.01$ and $\alpha = 0.001$ of $\chi^2_{0.99,2} = 9.21$ and $\chi^2_{0.999,2} = 13.82$, respecitvely. The null of homoscedasticity is emphatically rejected in each case, while heteroscedasticity is corroborated visually, the variance of residuals clearly increasing with the odds.

We thus call into doubt the results obtained by Deschamps and Gergaud. Their regression does not at all seem capable of remaining calibrated to the same degree at all magnitudes of the independent variable. Recall, also, that our analysis above only includes matches from our dataset with 6 bookmakers offering odds, as had DG. In the following section, we evaluate an extended regression as a means for deciphering the true causal relationship between bookmaker uncertainty and betting returns, as well as a different mechanism altogether.

4 Returns to Bookmaker Uncertainty: Our Model

We now include matches with any number of bookmakers, not only the six considered by DG. The first method aimed to amend the incompleteness of the authors' attempt is simply an extended multi-linear regression:

$$Mad_{ij} = \beta_0 + \beta_1 \frac{1}{\bar{o}_{ij}} + \beta_2 \frac{1}{\bar{o}^2_{ij}} + \beta_3 \frac{1}{l_{ij}} + \beta_4 \frac{1}{l^2_{ij}} + \beta_5 n_{ij} + \beta_6 n^2_{ij} + \varepsilon_{ij} \tag{9}$$

where n_{ij}, as discussed above, is the number of bookmakers offering odds on match i, outcome j. This variable is included because, since we are not only analysing games with six bookmakers precisely, there could be some variation in Mad_{ij} relating to the number of bookmakers that this mean covers. We regress on the inverses of the odds, rather than the odds themselves since this provides a much more narrow range for Mad_{ij}, given that these inverses are all bounded by 0 and 1. The longest odds are included because the bets for which returns are calculated are placed at the longest odds, not the mean odds, meaning that the residual should be independent of this variable as well. Due to likely multicollinearity issues from including both l_{ij} and $\bar{o}_{ij}$, we perform principal component analysis (Pearson, 1901; Hotelling, 1933) across the four odds-related variables. Then, when these principal components have been derived, we regress Mad_{ij} on these data, combined with a constant and n_{ij}. Table 1 compares this model to the original DG model. This latter model was run across all numbers of bookmakers, not only six.

Disagreement	Home		Draw		Away	
	DG	(9)	DG	(9)	DG	(9)
Very Low	3.27	2.48	3.98	3.82	5.91	4.37
Low	2.27	2.61	3.48	3.69	3.54	4.54
Average	2.07	2.62	3.47	3.64	3.21	4.45
High	1.99	2.58	3.54	3.65	3.12	4.29
Very High	3.18	2.50	4.13	3.79	5.94	4.07
$s_{\bar{l}_{gj}}$	0.621	0.066	0.311	0.083	1.449	0.177
Breusch Pagan: LM	4.2×10^5	1.7×10^4	5.4×10^5	1.3×10^4	6.0×10^5	1.8×10^4

Table 1: $\bar{l}_{gj}$: the mean odds bet at in each group, and their standard deviation s. If the regression residuals were homoscedastic, standard deviations would be close to 0. Although heteroscedasticity remains present, as evidenced by the Breusch-Pagan test statistics, our model vastly improves upon that of DG. For six degrees of freedom, $\chi^2_{0.99,6} = 16.82$ and $\chi^2_{0.999,6} = 22.46$. Note that the DG results are with any number of bookmakers, not exclusively six.

Although there is heteroscedasticity even here, the improvement in enormous. The deviation between odds in the different groups, for each outcome, is lower, meaning that there is more signal between bookmaker uncertainty and returns, with the dominating noise produced by the favourite-longshot bias less present. This is corroborated by the Breusch Pagan test statistic being more than an order of magnitude lower for the extended model, for each outcome. Note that these improvements are not due to the n_{ij} variables controlling for the number of bookmakers offering odds on a game and outcome: β_5 and β_6 in (9) are essentially zero for all three outcomes, and our model performs just as well in their absence. Ultimately, despite this significant improvement, the confounding effect of the odds remains. We conclude that a regression-based approach is suboptimal for deciphering the effect of bookmaker uncertainty on returns.

Instead, we decide to control for the bet price directly. By virtue of having an extremely large dataset, as well as there only being so many prices that the bookmakers typically set, these different individual prices (1.25, 4.6, 2.33, and so on) will be seen numerous times. More specifically, for the three outcomes, respectively, 248, 164 and 273 distinct prices have been the longest odds for their outcome at least 100 times. For Home, $l_{ij} = 2.1$ was the longest odds in 6,575 matches, for Draw, $l_{ij} = 3.4$ was the longest price for 18,626 games, while for Away, $l_{ij} = 4.0$ was the most common longest price with 5,147 such games in the dataset.

We have 248, 164 and 273 odds with at least 100 games. For each, we can construct five groups G_{gjk} with at least 20 games, split by Mad_{ij}, which is still calculated as in (5). There is no need to regress Mad_{ij} now, since we control directly for the price, meaning that there is no favourite-longshot Bias. This is why we can bin by Mad_{ij}, rather than ε_{ij}. Taking $m_{jk} = |(i,j) \mid l_{ij} = l_{jk}\}|$, we first sort the matches such that $l_{ij} = l_{jk} \ \forall \ i \in \{1, \ldots, m_{jk}\}$. Then, after sorting the matches such that $Mad_{ij} \leq Mad_{(i+1)j} \ \forall \ i \in \{1, \ldots, m_{jk} - 1\}$, we have

$$G_{gjk} = \{(i,j) \mid Mad_{(q_{g-1})j} \leq Mad_{ij} < Mad_{(q_g)j}, i \leq m_{jk}\} \tag{10}$$

where $q_g = \lfloor g \frac{m_{jk}}{5} \rfloor$ for $g = 1, \ldots, 5$ and $q_0 = 0$. Note that we only consider l_{jk} such that the corresponding $m_{jk} \geq 100$. For each of the five groups, in similar vein to above, we calculate the mean return $\bar{r}_{gjk}$ (equation (4), replacing G_{gj} with G_{gjk} and o_{ij} with l_{jk}) and $\bar{Mad}_{gjk}$ (equation (2), replacing o_{ij} with Mad_{ij}, while again replacing G_{gj} with G_{gjk}), giving five data points $(\bar{Mad}_{gjk}, \bar{r}_{gjk})$. If Deschamps and Gergaud's hypothesis (and dubious conclusion) is correct, on average, there will be a positive relation in these five data points, with the mean absolute deviation as the independent variable and returns dependent. Thus, for each individual longest odds l_{jk}, by outcome, we regress

$$\bar{r}_{gjk} = \alpha_{jk} + \beta_{jk}\bar{Mad}_{gjk} + \varepsilon_{gjk} \tag{11}$$

For each of the $m_{jk} = 248, 164$ and 273 different l_{jk}, respectively, by outcome, we gather this β_{jk} coefficient, expecting their distribution to be centered significantly above 0, since this would mean that, on average, returns increase with bookmaker disagreement, controlling for price. Thus we test the two-sided hypotheses $H_0 : \beta_j = 0;\ H_1 : \beta_j \neq 0;\ j \in \{\text{Home, Draw, Away}\}$, with $\bar{\beta}_{jk}$ the sample mean and population estimator. Applying the one-sample t-test (Student, 1908), we obtain a test statistic t, and hence a p-value $p = Pr(|T| \geq |t| \mid H_0)$, where T is a stochastic variable drawn from the t-distribution with ν degrees of freedom: $T \sim t_\nu$. Figure 4 shows the distribution of β_{jk}, for each outcome. Table 2 shows the results from the hypothesis test, for each outcome.

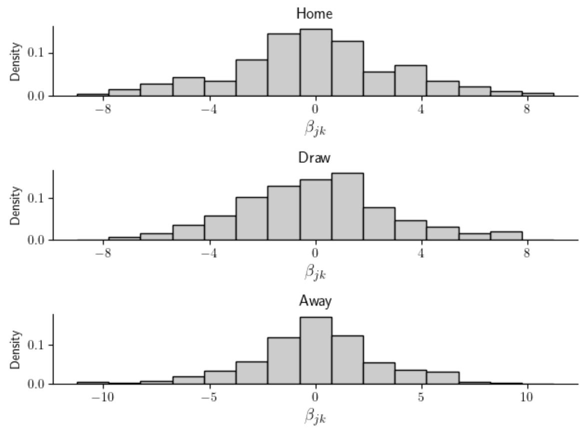

Figure 4: Each l_{jk} gives five bins G_{gjk}. We regress mean returns on bookmaker uncertainty, across the bins. The figure plots the distribution of slope coefficients β_{jk} from these regressions. Centered at zero, no relationship seems to exist.

	Home	Draw	Away
t	0.251	-0.328	0.546
p	0.802	0.743	0.586
$\bar{\beta}_{jk}$	0.065	-0.079	0.116
$s_{\beta_{jk}}$	4.080	3.084	3.493
n	248	164	273

Table 2: Each individual odds l_{jk} give a slope coefficient β_{jk} capturing the effect of bookmaker disagreement $\bar{Mad}_{gjk}$ on betting returns $\bar{r}_{gjk}$. For each outcome j, we perform a two-sided one-sample t-test, obtaining a test statistic t and p-value p, using sample mean and standard deviation $\bar{\beta}_{jk}$ and $s_{\beta_{jk}}$, respectively. We fail to reject the null hypothesis that there does not exist a relationship between the two variables.

When completely isolating the relationship between bookmaker uncertainty and betting returns, contrary to the findings of DG, we find no evidence of a positive effect of the former on the latter. Visually, the distributions of $\bar{\beta}_{jk}$ appear to be centred at zero for each outcome. This is confirmed by the t-test, which gives p-values $p \geq 0.586$, for the three outcomes. We therefore conclude that the relationship identified by DG, although entirely reasonable, and even expected, intuitively, does not exist in reality.

5 Conclusion

We have thus seen that the positive effect of bookmaker uncertainty on betting returns which Deschamps and Gergaud found cannot be replicated across our dataset. Extreme heteroscedasticity was identified in their regression, which was reduced but remained excessive in our extended regression. Although not a problem necessarily, in this case, heteroscedasticity causes the betting returns to be related not with bookmaker uncertainty, as intended, but instead with the bet odds. Due to the favourite-longshot bias, where betting returns are decreasing in odds, this confounds the desired relationship. Indeed, when a different approach was attempted, controlling for the price directly, no effect whatsoever was found, for any outcome. We suspect that DG were simply the subjects of a sample-size trick, their 8,377 matches significantly fewer than the 174,475 present in our dataset, and lower, also, than the 45,693 games with exactly six bookmakers. For instance, the draw bias which the authors discovered, where "draw odds yield a much higher return than home or away odds", also has no support in our data (see Figure 1). Regardless, we conclude, despite the appealing intuition, that disagreement amongst bookmakers does not generate greater betting returns.

6 References

Breusch, Trevor S, and Adrian R Pagan. "A simple test for heteroscedasticity and random coefficient variation". *Econometrica: Journal of the econometric society*, 1979, pp. 1287–94.

Deschamps, Bruno, and Olivier Gergaud. "Efficiency in betting markets: evidence from English football". *The Journal of Prediction Markets*, vol. 1, no. 1, 2007, pp. 61–73.

football-data.co.uk. "Data Files: All Countries", www.football-data.co.uk/downloadm.php. Accessed 15 Jan. 2024.

Hotelling, Harold. "Analysis of a complex of statistical variables into principal components." *Journal of educational psychology*, vol. 24, no. 6, 1933, p. 417.

Hurley, W., and L. McDonough. "The favourite-longshot bias in parimutuel betting: A clarification of the explanation that bettors like to bet longshots". *Economics Letters*, vol. 52, 1996, pp. 275–78. https://doi.org/10.1016/S0165-1765(96)00870-1.

Law, David, and David A Peel. "Insider trading, herding behaviour and market plungers in the British horse–race betting market". *Economica*, vol. 69, no. 274, 2002, pp. 327–38.

Ottaviani, Marco, and Peter Norman Sørensen. "The favorite-longshot bias: An overview of the main explanations". *Handbook of Sports and Lottery markets*, 2008, pp. 83–101.

Pearson, Karl. "LIII. On lines and planes of closest fit to systems of points in space". *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, vol. 2, no. 11, 1901, pp. 559–72.

Shin, Hyun Song. "Measuring the incidence of insider trading in a market for state-contingent claims". *The Economic Journal*, vol. 103, no. 420, 1993, pp. 1141–53.

Student. "The probable error of a mean". *Biometrika*, 1908, pp. 1–25.

Williams, Leighton Vaughan. "Can Bettors Win?" *World Economics*, vol. 2, no. 1, 2001, pp. 31–48.