

Scary Estimators and the Tentacle Property, or the Return of Paul's Forecasts at the 2010 World Cup

Konstantin M. Wacker^a

^a*Chair of Development Economics & Centre for Statistics, University of Göttingen;
konstantin.wacker@stud.uni-goettingen.de*

Abstract

This paper addresses the question whether the octopus Paul is a reliable predictor for the outcome of 2010 FIFA World Cup games. We find statistically convincing evidence for this hypothesis and for the theory that Paul's decision actually influences the results of the games.

Accordingly, we conclude that Spain is going to be World Champion with a probability of 98.4 %.

Keywords: Oracle Property, Hypothesis Testing, FIFA World Cup, octopus Paul

“Lies, damned lies, and statistics”

(Mark Twain)

1. Introduction

During the FIFA World Cup 2010, the octopus Paul was introduced as a controversial predictor for the outcome of soccer games.¹ We thus investigate the oracle property² of the tentacle and find statistically convincing evidence that puts the nature of football into a new perspective.

The article is outlined as follows: section 2 reviews the literature, section 3 introduces our hypothesis based on a well based theory. In section 4 we derive the central test statistic. Section 5 presents the outcome and concludes.

2. Literature Review

The sparse contribution of *Nolan* (2010) is the only contribution in this context to the best of our knowledge. This means that we have been too lazy to do an extensive literature review and that our contribution may be groundbreaking, maybe even leading us to tenure.

However, in order to suck up to the editor, we want to mention the works of *Friend of Editor* (2006), *Editor's Wive* (2008), and *Potential Referee* (2009) in related fields.

Concerning the sceptical arguments of *Miske* (2010) we want to point out that minor traces of flavor at the boxes and small variations in the mussel size can be assumed as being independently and identical distributed deviations resulting in a normal distribution with mean zero according to the central

¹Yes, we know, it was already used during the Euro Cup 2008 (otherwise, the wordplay in the title would not have worked, man!) but considering this would make the derivation of the test statistic more complicated and our results would not nearly be as neat...

²Not to be confused with the oracle property advertised by *Fan and Li* (2001) for their SCAD estimator which is clearly bullshit, see *Leeb and Pötscher* (2008), nor with the oracle inequality frequently referred to in the literature.

limit theorem. These factors thus have no influence asymptotically - Paul definitely is a consistent estimator.

3. Theory and Hypothesis

Our provocative theory is that the oracle Paul is not only the best non-linear estimator for World Cup results but that the evolution of football games itself is off the hands (better: feet) of the players but lies in the tentacles of Paul.

To come up with a test statistic that a potential referee may buy, we specify the converse of our theory as our null hypothesis, i.e. our null hypothesis is that Paul is only a stupid octopus and has no predictive power nor any influence on football games.

4. Derivation of the Test Statistic

Under our null hypothesis, the choice of Paul is a Bernoulli random variable X (just as a toss of a coin) with the following outcomes:

$$X = \begin{cases} 1, & \text{if Paul predicts correctly} \\ 0, & \text{otherwise} \end{cases}$$

where the probability of success is $p = 0.5$.

Now let

$$Y_n := \sum_{i=1}^n X_i \tag{1}$$

be the sum of n realizations (i) of this random variable X . It is seen immediately that under our null hypothesis Y_n follows a binomial distribution.

In order to derive the critical value Z of our one-sided test statistic at the 5 % level of significance we take the 95 % quantile of a binomial distribution. Remember that Paul predicted 6 out of 6 World Cup games correctly (up to July 10, 2010 5 a.m. CET). For these $n = 6$ predictions with the realization $Y_6 = 6$ at the given success probability $p = 0.5$ under our null hypothesis, the according 0.95 quantile of the distribution is given at $Z = 5$.

5. Results and Conclusion

Due to the fact that we observed $Y_6 = 6 > 5 = Z$ correct predictions of Paul, we can clearly reject the null hypothesis that Paul is just a stupid octopus and has no predictive power nor any influence on football games on the critical 5 % level. We can thus conclude the following:

- The octopus Paul is not just a stupid octopus that picks a team randomly - he knows what he does!
- Accordingly, Spain is going to be World Champion and anyone who argues the converse is just an ignorant esoteric hippie who has no understanding for hard-fact science and the laws of universe. The accuracy of this prediction is equivalent to $1 - p^n$ - the p-value of our test statistic which is $1 - p^n$, i.e. $1 - 0.5^6 = 0.984375$.
- In the (unrealistic) case of anyone being interested in the petite final or the performance of Germany: yes, they win.
- The results of World Cup soccer games are not driven by the efforts of players but by the preferences of Paul.

To identify the possible transmission channels of the latter effect lies beyond the scope of this paper. Nevertheless, we want to point out that a reason may lie in the mere fact that a football player has only two legs, Paul has 8. Future research may also focus on the question whether there are structural breaks in the series, i.e. whether the forecasting power of Paul and its impact on outcomes also holds in non-World-Cup games or was only a one-time happening specifically bound to South Africa. We are confident to find four-year funding to investigate these crappy ideas.

- [1] Fan, J. & R. Li (2001): “Variable selection via nonconcave penalized likelihood and its oracle properties”; in: Journal of the American Statistical Association Vol. 96: 1348-1360
- [2] Leeb, H. & B. M. Pötscher (2008): “Sparse estimators and the oracle property, or the return of Hodges’ estimator”; in: Journal of Econometrics, Vol. 142: 201-211
- [3] Miske, Volker (2010): “Obergescheites Kommentar”; in: Welt Online, 10. Juli 2010
- [4] Nolan, Hamilton (2010): Blog Contribution; Gawker.com