Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 26 (2010) 551–563
international journal of forecasting
www.elsevier.com/locate/ijforecast

Are differences in ranks good predictors for Grand Slam tennis matches?

Julio del Corral[a,*], Juan Prieto-Rodríguez[b,1]

[a] *Fundación Observatorio Económico del Deporte—Department of Economics, University of Castilla la Mancha, Ronda de Toledo s/n, 13071, Ciudad Real, Spain*

[b] *Fundación Observatorio Económico del Deporte—Department of Economics, University of Oviedo, Avda. del Cristo s/n, 33006, Oviedo, Spain*

Abstract

This paper tests whether the differences in rankings between individual players are good predictors for Grand Slam tennis outcomes. We estimate separate probit models for men and women using Grand Slam tennis match data from 2005 to 2008. The explanatory variables are divided into three groups: a player's past performance, a player's physical characteristics, and match characteristics. We estimate three alternative probit models. In the first model, all of the explanatory variables are included, whereas in the other two specifications, either the player's physical characteristics or the player's past performances are not considered. The accuracies of the different models are evaluated both in-sample and out-of-sample by computing Brier scores and comparing the predicted probabilities with the actual outcomes from the Grand Slam tennis matches from 2005 to 2008 and from the 2009 Australian Open. In addition, using bootstrapping techniques, we also evaluate the out-of-sample Brier scores for the 2005–2008 data.

Keywords: Sports forecasting; Probit models; Prediction intervals; Tennis; Brier scores; Bootstrapping

1. Introduction

Several papers have studied sports data predictions, given that sports events are quite distinct from completely randomized events such as lotteries. The previous literature on sports forecasting can be divided into several groups, according to the type of forecasts made. For example, several papers have aimed to predict the result of a particular match between two contestants (Boulier & Stekler, 2003; Caudill, 2003; Klaassen & Magnus, 2003). Other papers have aimed to predict the point spread between two contestants (Smith & Schwertman, 1999), and still other papers have aimed to predict the winner of sports events involving several contestants, such as tournaments (Anderson, Edman, & Ekman, 2005; Clarke & Dyte,

* Corresponding author. Tel.: +34 926295300x3521; fax: +34 926 295 211.
E-mail addresses: julio.corral@uclm.es (J. del Corral), juanprieto@uniovi.es (J. Prieto-Rodríguez).
[1] Tel.: +34 985103768; fax: +34 985104871.

0169-2070/$ - see front matter © 2010 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ijforecast.2009.12.006

2000), leagues (Rue & Salvesen, 2000) or races (Bolton & Chapman, 1986).

Two main methods for predicting the outcome of a sports event exist, namely statistical models and expert evaluations. Thus, some scholars have compared the accuracies of these competing methods (Anderson et al., 2005; Boulier & Stekler, 2003; Forrest, Goddard, & Simmons, 2005). With regard to statistical models, a myriad of different models have been used. For instance, if the objective is to construct winning probabilities for contestants in a match, the most prominent approach is to use either logit (Clarke & Dyte, 2000; Klaassen & Magnus, 2003) or probit (Abrevaya, 2002; Boulier & Stekler, 1999) models. Another alternative is to use the maximum score estimator (Caudill, 2003). If a tie is possible (e.g., in soccer), either ordered probit models (Goddard & Asimakopoulos, 2004) or multinomial logit models (Forrest & Simmons, 2000) can be used. However, if the probabilities for a match are based on the number of points, goals or run probabilities, Poisson regression (Dixon & Coles, 1997) and negative binomial (Cain, Law, & Peel, 2000) models should be used to take the discrete nature of the data into account.

With regard to the variables that enter these statistical models, Goddard and Asimakopoulos (2004) and Goddard (2005) used several variables related to past results, as well as information about the number of goals scored and conceded. Forrest and Simmons (2000) used the performance in previous matches as an explanatory factor in their logit models of the English national soccer league. Dyte and Clarke (2000) used *Fédération Internationale de Football Association* (FIFA) ratings to predict the numbers of goals scored by national teams competing in the 1998 FIFA World Cup. Similar to FIFA, other sports-governing bodies also produce rankings based on the past performances of contestants. These rankings have been used as predictors of victory in several different settings. For instance, Boulier and Stekler (1999) found that the ranking difference between contestants is a good predictor in professional tennis and collegiate basketball. Lebovic and Sigelman (2001) demonstrated the accuracy of collegiate football rankings in predicting match outcomes. Smith and Schwertman (1999) showed that the difference in rankings is a good predictor of the victory margin in collegiate basketball. Caudill and Godwin (2002) developed a heterogeneous skewness model that takes into account not only differences in rankings but also their degree.

In tennis in particular, Klaassen and Magnus (2003) proposed a method of forecasting the winner of a match at the beginning of the match, as well as during it. For this purpose, they used a measure based on nonlinear differences in rankings, similar to that used by Caudill and Godwin (2002). Clarke and Dyte (2000) used tennis rankings to estimate the chance of winning as a function of the difference in rating points, and were able to estimate a player's chance of a tournament victory once the draw for the tournament became available.

The aim of the present paper is to extend the previous literature in the following ways. First, we test whether the ranking difference is a good predictor of tennis victories, using an approach which is different from those used in previous studies. Specifically, we classify our variables into three groups, namely a player's past performance, a player's physical characteristics and the match characteristics. We then estimate three alternative probit models for men and women separately using Grand Slam tennis match data from 2005 to 2008. In the first model, all three explanatory variables are included, whereas in the other two specifications, either the player's physical characteristics or the player's past performances are not considered. Subsequently, the forecasting accuracies of the different models are evaluated by computing the Brier scores and comparing the predicted probabilities with the actual outcomes from the 2005 to 2008 Grand Slam matches, as well as from the 2009 Australian Open matches. Moreover, using bootstrapping techniques, we evaluate the out-of-sample Brier scores from 2005 to 2008. Using probit estimates, we also study the effect of ranking differences on predicting Grand Slam tennis outcomes, and in particular we analyze whether this effect varies by gender.

Section 2 presents the empirical model specifications. The data are presented in Section 3, followed by the probit results in Section 4. In Section 5, the predictive accuracies of the models are analyzed. Our conclusion is given in Section 6.

2. Empirical model specifications

In order to obtain the determinants of match outcomes, we estimate probit models in which the