Probabilistic Forecasting – The National Hockey League Totals Market

*Alex Traugutt, Fontbonne University*

**Finance/Economics - Economics (Professional Sport)**

**Abstract 2019-091**

Probabilistic forecasting within the sports setting has become an increasingly popular topic, particularly for individuals who believe that they can forecast outcomes to a degree which will allow them to earn above-average returns. Despite this growing popularity, much of the previous empirical research attention has focused on measures of efficiency and the identification of systematic biases (e.g., favorite-longshot) in the various markets. The primary purpose of this paper is to determine whether historical betting data can be utilized to accurately forecast wager outcomes in the National Hockey League (NHL) totals market to a degree which would yield profitable returns. For reference, totals betting involves a wager on whether the combined number of goals will be over or under the line set by the oddsmaker. In comparison to other professional sporting leagues in North America, the NHL and the totals market specifically has received considerably less research attention making it an ideal market for research consideration.

For this study, a logistic regression model concerned with forecasting winning under wager outcomes was employed. Consistent with forecasting analysis, the dataset, which consisted of betting data from nine seasons (2007/08 – 2015/16), was divided into in-sample and out-of-sample subsets. To analyze the forecastability of this market, a three-step process was utilized which involved (1) estimating a model, (2) forecasting outcomes using model estimates, and (3) calculating profits based on the results from the devised betting strategies. The model was first estimated based on the in-sample data (n = 6,480). The estimated parameters were then utilized to calculate predicted probabilities in the out-of-sample subset of data (n = 3,243) and assess the profitability of three betting strategies which measured (1) all games in the out-of-sample period, (2) games in which the under total was favored based on the closing odds, and (3) games which featured specific closing totals (e.g., 4.5, 5, 5.5, 6). Since outcomes from a logistic model are estimates of probability, for each strategy, wagers were placed on games that featured predicted probabilities greater than or equal to .50, .5238, .55 and .575.

Results indicated that there were opportunities to forecast betting outcomes to a degree which would allow for profitable returns, however such outcomes were not widespread. Applying the estimated model to the first and second strategies resulted in cumulative win percentages of 49% and 77% with net losses totaling $6,662.05 and $97.28, respectively. In the third strategy, however, wagering on the under when the total closed at 5.5 resulted in a positive return of $3,474.29. Thus, it can be surmised that the market may be susceptible to profitable forecasting under certain conditions. Additional results and applications of the model estimates will be presented.

Given that the totals market in the NHL had not been previously researched in this manner, these results provide a foundation upon which to build forecasting models for this market. Moreover, these findings contribute to the growing body of literature, both within and outside of sport, concerned with utilizing historical information and statistical modeling to forecast future outcomes.