Thirst for Glory: Are Players Motivated by Milestones?

Arseniy Stolyarov Gleb Vasiliev

New Economic School, Higher School of Economics, Toros Research

1. Motivation

Today the increased media activity along with the high availability of data makes players focus more on individual stats. Several unofficial strikers' clubs (e.g.Fedotov club in Russia, Blokhin club in Ukraine) collect statistics on the number of goals scored by players. Do players score more frequently when they are close to the inclusion into such ceremonial clubs? Our paper analyzes this issue via statistical and econometric methods.

If the hypothesis of better performance of players reaching a milestone is true, then it may be beneficial:

1. To coaches, that may take into account a higher likelihood of a goal by a player that enters the ceremonial club while choosing the line-up.
2. To both gamblers and booking companies as they may better estimate the probability of a goal by a player, who is close to the entrance into a ceremonial club.

2. The Data

The sample consists of players from top-5 European leagues that scored 100 or more goals in one league during the last 20 years. The data from leagues' websites is used along with the news from sports newspapers.

3. Hypotheses and Statistical Tests

The main hypothesis is that players score goals more frequently if the number of total goals scored by them in one tournament approaches a round number (e.g. 100, 200. etc). This hypothesis is tested using following statistical procedures:

1. Assuming exponential distribution of the time between two subsequent goals scored by the player, the estimate $\hat{\lambda_1}$ of λ_1 coefficient is obtained. Then, we obtain another estimate $\hat{\lambda_2}$ of λ_2 using the sample of 100^{th} goals and test $\lambda_1 = \lambda_2$ against $\lambda_1 \neq \lambda_2$.
2. A regression of the form $t_i = \beta_0 d_i + \sum_{i=1}^{n} \beta_i x_i$ is estimated, where t_i stands for the time between two goals, d_i stands for the dummy variable which is equal to 1 in case of next goal being 100^{th} in the league by the player, x denote control variables including the strength of the opponent, form of the player, average time between the goals of the player, etc. Several specifications are used and test for $\beta_0 = 0$ is performed in each case.
3. The goodness-of-fit test that assumes the same distribution of time before the next goal in all subsamples under the null hypothesis and tests it against the alternative hypothesis of different distributions in various subsamples.