Eli Metzner
11/14/2018
Econ 438 Paper 3 Outline

Introduction

For the third paper, I plan to expand on my second paper, which attempted to estimate the marginal revenue product of NFL quarterbacks and found that quarterbacks tended to be paid almost an order of magnitude under their predicted revenue. The models used in this paper had some flaws, though, so I plan to improve both the model that predicts wins added by a quarterback as well as the prediction of revenue added based on wins added.

1) Wins Model

The current model I use for predicting wins added by a quarterback is as follows:

$$W_i = \alpha_i + \beta_1 Yd_i + \beta_2 Cmp_i + \beta_3 TD_i + \beta_3 Int_i + \beta_5 RZCmp_i + \beta_6 RZTD_i + \beta_7 RZInt_i + \beta_8 {Age_i}^2 + \epsilon_i$$

where α_i represents team-specific fixed effects, Yd_i is net yards gained per pass attempt, Cmp_i is pass completions per pass attempt, TD_i represents touchdowns per attempt, Int_i is interceptions per attempt, Age gives the quarterback's age in years, and RZ signifies the red-zone-split statistics for each of these categories. i indexes each NFL quarterback across 10 years, a total of 331 observations.

Looking to improve on this model, I plan to add several more statistics to improve fit − notably, rushing among quarterbacks has become a bigger part of the game in recent years, so I plan to look for QB rushing statistics as well to add to the model. This data is readily available online. I also plan to add time-fixed effects in addition to the team-fixed effects already incorporated into the model. The original model worked under the assumption that playing and coaching styles differ across teams, so each quarterback has a different opportunity to add wins based on the team they play for, but that these difference do not vary across years. In a 10-year sample size, though, player and coach turnover will happen much more frequently than once every 10 years, so adding year-fixed effects should help control for unobserved factors such as offensive line strength, etc, which probably does vary across time. Therefore, our updated wins model will look like:

$$W_i = \alpha_i + \gamma_i + \beta_1 Yd_i + \beta_2 Cmp_i + \beta_3 TD_i + \beta_3 Int_i + \beta_5 RZCmp_i + \beta_6 RZTD_i + \beta_7 RZInt_i + \beta_8 {Age_i}^2 + \\ \beta_9 RuYd_i + \beta_1 0 RuTD_i + \beta_1 1 Ru1D_i + \epsilon_i$$

where all variables are the same as indicated in the first model, with additional γ_i represents time fixed effects, $\beta_9 RuYd_i$ and $\beta_1 0 RuTD_i$ and $\beta_1 1 Ru1D_i$ represent net rushing yards per rush attempt, TD per rush attempt, and first downs gained per rush attempt.

2) Revenue Model

We used a very basic model to predict the effect of added wins on revenue generated per team:

$$Rev17_{it} = \alpha + \beta_1 Wins_{it-1} + \epsilon_i$$

where i indexes teams across the 10-year sample size, and t is the current period. While this model did produce significant results, its R^2 was around 0.01, suggesting that we can improve the fit substantially.

I have two main ideas for improving the fit of the revenue model: the first, relatively simple, is adding controls for market size and other economic indicators using year-by-year census MSA-level data on population and per capita income, with year fixed-effects to take into account the growing popularity of the league year-over-year. It also makes sense to use the log of revenue instead of the raw data, as logs can smooth out the wide variation in revenues across teams.

$$log(Rev17_{it}) = \alpha_i + \beta_1 Wins_{it-1} + \beta_2 Pop_i + \beta_3 IncomePC_i + \epsilon_i$$

where α_i represents year fixed effects, Wins are wins added in the last season, Pop is area population, and IncomePC is per-capita income in the team's area. This should improve the revenue model and adjust for differences in each team's home market. It might also be worth investigation an autoregression-based model for revenue, in that a team's revenue in the current season is probably highly correlated to their revenue in the past season – an autoregression would give us an idea of what percentage of revenue "carries over" to the next year. For this model, we would have:

$$log(Rev17_{it}) = \alpha_i + \beta_1 Wins_{it-1} + \beta_2 Rev17_{it-1}$$

We can ideally compare both models and choose the one with a better fit – other factors, such as dummies for big free agent signings, etc, could also be included to improve the R^2.