Davis Kim

ECON 570

Final Project Document

5-01-2020

For this project, I'm working with NFL team data. The main factor I want to investigate is the effectiveness of the NFL draft as a league balancing device. The NFL draft is a method installed by the league to create parity within the league. Based on you position in the league standings the higher or lower the pick you receive. I want to investigate the impact of high draft picks awarded to poor performing teams on overall success or wins.

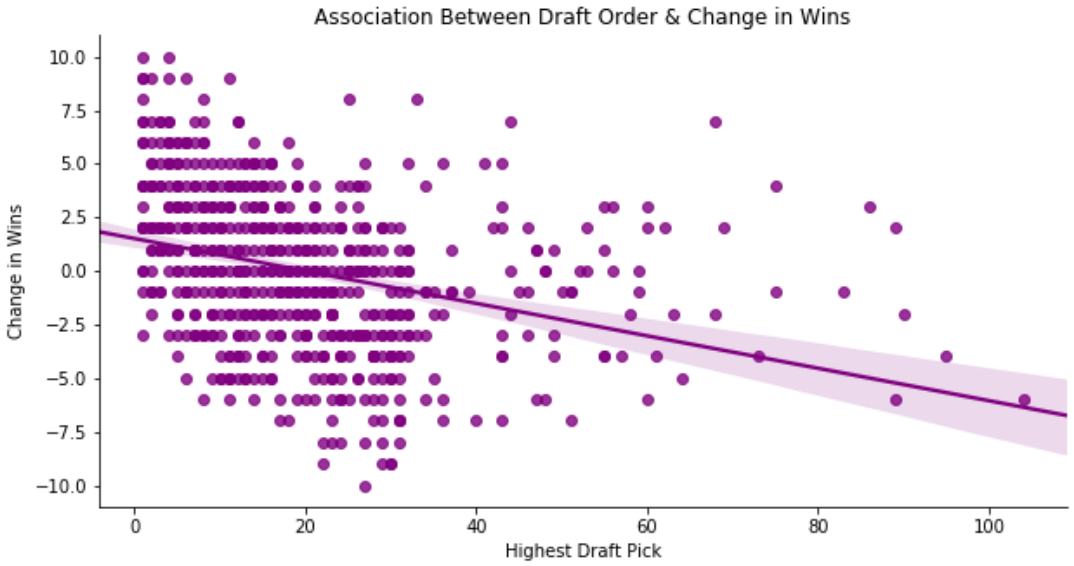
The data I'm working with is NFL data compiled from the NFL or profootballreference. The variables I would like to work with are (highest draft pick), the highest ordered draft pick for the team that year, (wins from last year), (wins this year), and (change in wins). I will implement OLS to see the impact of picks on change in wins. And, I will measure the effectiveness of the draft based on change in team performance or (change in wins).

League parity is an interesting concept to investigate, because I believe you can view and interpret the results in many different ways. First, I plot all of the data for all teams, and then utilize OLS if I can identify a trend with (draft pick positioning) and (change in wins). However, I thought it would also be interesting to also see if there is a trend present in each conference. Separating the data could show that certain groups perform better than others, which could show that the draft may not create league parity as much as initial findings may suggest.

I also investigate the effectiveness in the draft over time. One idea I thought of implementing is running regressions on the data from 1999-2009 and then running regressions on the data from 2010-2019. I think through this I can identify change over time by looking at

the coefficients from the regressions. I also introduce interaction terms with the dummy variables to see if I can measure marginal effects.

Figure 1: Draft Ordering & Change in Wins



This is a regression plot containing on the Y-axis: (change in wins) and on the X-axis: (highest draft pick). The scatterplot consists of all NFL Team data from (1999-2019). Through OLS, there is a distinguishable trend that I have found through preliminary research. I've also implemented a 95% confidence interval in the picture as well. Through the model that OLS generated I think it's shows that likely the draft is a good system to implement parity, but some groups may still perform better than others.

Simple Model: Change in Wins = $ß_0 + ß_1$(Highest Pick)

Change in Wins	Coefficient	Standard Error	T-Statistic	P-Values
Intercept	1.5114	0.216	7.002	0.000
Highest Pick	-0.0757	0.009	-8.884	0.000

This model was helpful to start because it confirmed my initial thoughts by seeing a low p-value for the coefficient. However, I also wanted to look at change over time, so I added a time dummy. I also generated and interaction term to find the marginal effect.

Figure 4: Dummy Model: Change in Wins = $ß_0 + ß_1$(Highest Pick) + $ß_2$(Time Dummy)

Change in Wins	Coefficient	Standard Error	T-Statistic	P-Values
Intercept	1.5366	0.251	6.112	0.000
Highest Pick	-0.0757	0.009	-8.878	0.000
Time Dummy	-0.0525	0.268	-0.196	0.845

Unfortunately, I was not able to gather more information from this model. The (time dummy) was not statistically significant. So, I generated an interaction term (Time Dummy X Highest Pick) to measure the marginal effect of the time dummy and highest pick.

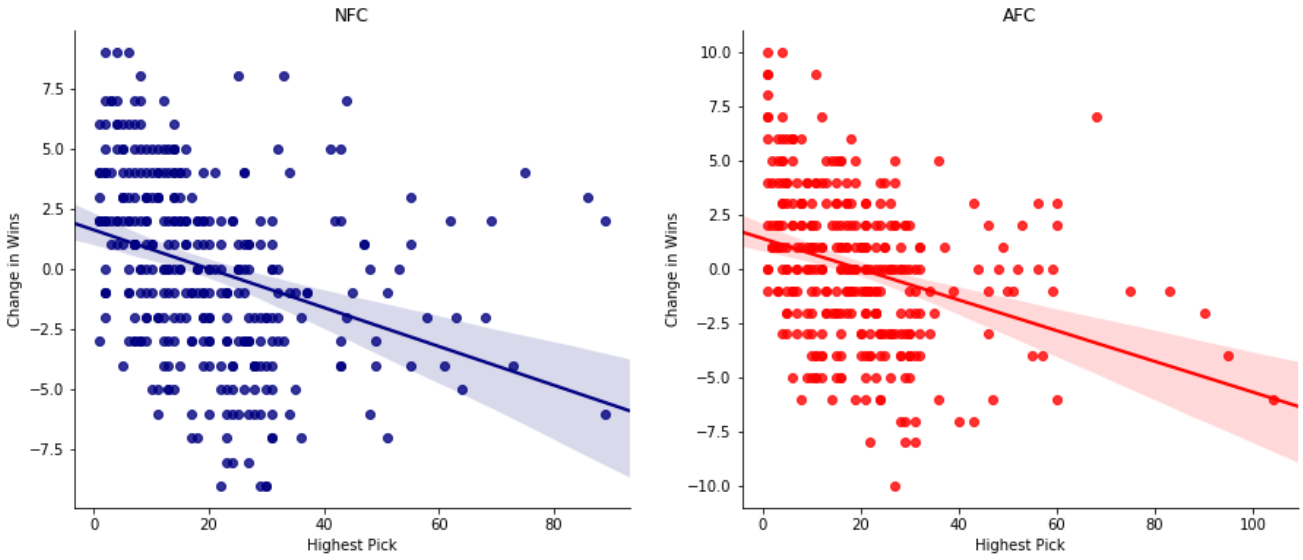
Figure 5: Interaction Model

Change in Wins = $ß_0 + ß_1$(Highest Pick) + $ß_2$(Time Dummy) + $ß_3$(Interaction)

Change in Wins	Coefficient	Standard Error	T-Statistic	P-Values
Intercept	1.8399	0.301	6.111	0.000
Highest Pick	-0.0909	0.012	-7.627	0.000
Time Dummy	-0.6701	0.431	-1.553	0.121
Interaction	0.0310	0.017	1.824	0.069

Unfortunately, the interaction term did not generate statistically significant results. However, a p-value of (0.069) was nearly statistically significant. If it were statistically significant based on the coefficient value (0.0310), late draft selections in the dummy time group would not be as penalized by later draft selections. Furthermore, I decided to investigate the same modeling concept after splitting the data into conference groups.

Figure 6: Conference Plots (With Basic Model)



First, I generated regression plots for the NFC and the AFC to see if there were any discernable visual differences from the basic models. It appears that there are no discernable differences between the two when only including (Change in Wins) as the dependent variable and (Highest Pick) as the independent variable. However, things may look different after including the interaction terms in the models.

Figure 7: NFC Interaction Model

Change in Wins	Coefficient	Standard Error	T-Statistic	P-Values
Intercept	1.6843	0.452	3.725	0.000
Highest Pick	-0.0870	0.018	-4.720	0.000
Time Dummy	-0.1154	0.639	-0.181	0.857
Interaction	0.0121	0.025	0.480	0.632

Figure 8: AFC Interaction Model

Change in Wins	Coefficient	Standard Error	T-Statistic	P-Values
Intercept	1.9906	0.402	4.955	0.000
Highest Pick	-0.0943	0.015	-6.104	0.000
Time Dummy	-1.2536	0.584	-2.148	0.032
Interaction	0.0524	0.023	2.254	0.032

The results were very powerful. The regressions showed that there were no discernable differences over time for the NFC group, but there were discernable differences in the AFC group. In the AFC group now both the (Time Dummy) and the (Interaction) coefficients are statistically significant. The (Time Dummy) coefficient suggests that AFC teams performed worse in the draft during 2010-2019 with a coefficient of (-1.2536). However, the (Interaction) coefficient also suggests that AFC teams were less penalized by later draft picks during 2010-2019. This effect is outweighed by the extreme impact shown in (Time Dummy) coefficient.

Overall, the things that were found present in this dataset were a strong relationship between high draft picks and a change in win performance. The AFC conference performed worse in the draft during 2010-2019 than they did in 1999-2009. Both these results are interesting when considering the initial question: does the NFL draft creates parity in the league? The first result strongly suggests that draft is good tool to pull the teams performances on both extremes closer to each other. However, the second result shows that one group the AFC performed worse during a certain given time period. While this does not say anything about the NFC's performance during this time period, it's very possible that over time NFC has found a way to outperform the AFC. This could show that NFL draft may be tool better utilized by some than others.