**Run Productivity of Different Hit Outcomes**

**The Background**

When evaluating a player’s offensive value for trades, there are many new metrics that can be used to gauge a player’s value. The most common of which are offensive WAR, OPS, and WRC+. All of these are useful when comparing players to one another, however, they do not give much insight into how they may help a specific team. There are a variety of different ways players can achieve high statistics in these categories, whether they are a contact hitter, a slugger, or someone who controls the zone and walks at an elite rate.

What this does not address, is that an offense needs a diverse set of hitting styles. A solo home run has much less impact on a game than a multi-run homerun. A multi-homerun not only has the potential to put the game out of reach score-wise but is also extremely demoralizing. As Earl Weaver said, “The key to winning baseball games is pitching, fundamentals, and three-run home runs”. This means that a team cannot solely consist of sluggers who may slug over .450, however, have an OBP of less than .300. On the other side of the coin, teams need extra base hits to score runs consistently. Research has proven that home runs lead to more wins. With that being said, it is important to find a balance. Singles make extra-base hits more productive, and vis-versa.

The question then becomes, what does the team need more of? Teams will get varying degrees of productivity from the different hits based on how the team is currently built. It is possible to figure out what type of batter is needed, however when comparing batters within the same category or batters who may be in-between categories it would be useful to quantify who would bring more run production to the team. This is where the Cobb-Douglas production function comes in. The function in an economic sense is as follows:

Y is the economic output, measured in the real-value of goods. A is the total factor productivity (TFP), which includes the assumptions on technology and other factors that affect the productivity of labor and capital. The TFP scales labor, L, and capital, K. Beta and alpha are the output elasticities of their respective inputs. If an output elasticity is less than 1, this means that there is a decreasing marginal return on that input. An output elasticity of 1 is a linear return, and greater than 1 is an increasing marginal return on the input.

What makes the Cobb-Douglas function useful is that it is an interaction term of multiple inputs, and more than two inputs can be used. This means that the inputs are complementary, so the marginal productivity is dependent on the value of other inputs. This has been used in business cases for problems such as finding the optimal amount managers and associates at a retail store. Managers bring more value, however, are more costly and their added value is needed less the more managers that are present. For this reason, the most profitable combination of associates and managers varies based off of productivity estimates. It is for these reasons that the Cobb-Douglas function could be modified to be used for expected runs over the course of a season based on the production of outcomes from an at-bat.

**The Baseball Application**

In order to apply this function to baseball, some modifications need to be made. In order to do this, I propose two possible equations:

**Function 1:**

**R:** The predicted amount of runs over the course of a season

**E:** The run-scoring environment

**X:** Total extra-base hits for a team

**S:** Total singles and base on balls for a team

**β and α:** run-production elasticities

**Function 2:**

**R:** The predicted amount of runs over the course of a season

**E:** The run-scoring environment

**S:** Total singles for a team

**D:** Total doubles for a team

**T:** Total triples for a team

**H:** Total home runs for a team

**B:** total base on balls for a team

**β, α, γ, λ, and δ:** run-production elasticities

Both of these equations estimate the number of runs scored over the course of a season based on the different outcomes from an at-bat. They use the same principles of the Cobb-Douglas function that each outcome affects one another, and uses exponents to modify how productive an additional occurrence of an outcome is. The difference is in the number of outcomes being used to create the estimate. Both of these equations have their merits and should be tested. It is unclear yet the size of the data set, and by bucketing similar outcomes it may lead to a more accurate prediction as rarer outcomes such as triples may be weighed incorrectly on their own due to their rarity. The second equation, however, is more precise as it evaluates each outcome individually.

**Methodology and Reasoning**

Through prospective research, I have found similar studies using similar functions to predict wins in baseball, however, I have not seen the use of this style of function for run prediction. I believe that this is grounded in sound reasoning as the economic-baseball parrels are logical. Runs are equivalent to real value in economics as they are a way to measure the offensive outcome of a team. The run scoring environment is akin to the TPF as it is a constant that has an effect on the productivity of outcomes. In a lower run-scoring environment, the productivity of extra-base hits, singles, and base on balls is going to be lower as the likely hood of batting someone in or being batted in is lower. The run-production elasticities are good parallels for output elasticities as they estimate the amount of additional runs created through an additional outcome, and in economics, the output elasticities estimate an percent change in real value based on an additional unit of input. The background presented the argument as to why an interaction function is useful in this situation, as each outcome has a different productivity based on the number of other outcomes.