A pretty controversial issue in the NBA right now is parity between the eastern and western conference. This issue has especially become even bigger due to the rise of the dominance of the Golden State Warriors. There is a theory espoused by pundits, however, that this lack of parity is good for the league. To clarify, there is a controversial belief that a less competitive league can mean better viewership NBA final wise. Pundits often point to the Magic-Bird era as evidence.

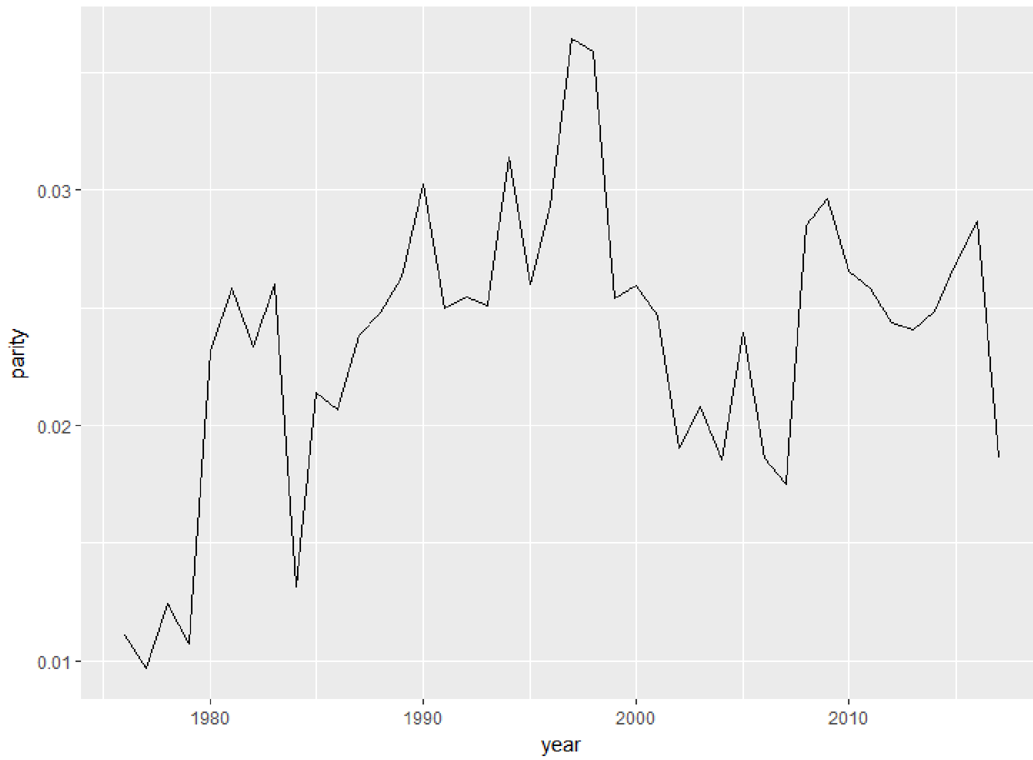
Let us look at historical data to see if a lack of parity leads to better ratings for the NBA. Specifically, I want to test the hypothesis that "the lack of overall parity leads to better final NBA television ratings".

To test this hypothesis, I plan on looking at two sets of time series data. First is a measure of league parity from 1976-2017. This period was chosen since this marks the merger of the NBA-ABA and coincides roughly with the beginning of Nielson NBA Finals Rating collection. There is no official metric to measure overall league parity. However, for this data analysis the metric I have decided to use is the variance of teams win/loss ratio. This is because variance measures the dispersion of the win/loss ratio. In other words, if the variance is low then teams were winning and losing at approximately the same level. To me, teams winning and losing at a similar or dissimilar rate is a pretty importance component in measuring parity. For this reason, I used variance as a measure of parity. Using python I web scraped the win/loss data from www.basketball-reference.com. I then used R to calculate the variance of each year.

The second-time series data is the NBA Finals Nielson Ratings (as a relative percentage to account for population inflation) from 1976 – 2017. This period of data was chosen because it is the same time over which league parity was calculated. It also roughly overlaps with the beginning of NBA Finals Rating collection by Nielson. I webscraped this data from Wikipedia using python.

I then used the statistical/economic time-series tool of "granger causality" to test whether time series data 1(higher yearly win/loss variance and less parity) seems to predict time series data 2(Higher NBA Finals Nielson Ratings). In more layman terms, does past league disparity predict better NBA Finals Ratings.

A plot of Parity/Variance throughout the years:



A plot of Finals Nielson Ratings throughout the years:

The “Granger Causality” Test can only be performed on stationary data. Based off the plots, I am willing to bet the data isn’t stationary. But I tested with an augmented Dickey-Fuller(ADF) test to further confirm non-stationarity.

I performed the adf test in R on the parity data and received a “p-value = .08214”. This means a non-rejection of the null hypothesis at a significance level of 5%. This implies the potential existence of a unit root and non-stationarity.

Similarly, I performed the adf test in R on the ratings data. I received a “p-value = .5465”. This means a non-rejection of the null hypothesis at a significance level of 5%. Again, this implies the potential existence of a unit root and the non-stationary of the data.

```
        Augmented Dickey-Fuller Test

data:  parity
Dickey-Fuller = -3.321, Lag order = 3, p-value = 0.08214
alternative hypothesis: stationary
```

```
        Augmented Dickey-Fuller Test

data:  rating
Dickey-Fuller = -2.0689, Lag order = 3, p-value = 0.5465
alternative hypothesis: stationary
```

Based off an adf test, I should transform the data to something stationary. So I took a first difference. For the parity data I got a “p-value =0.01805”. For the rating data I got a “p-value = 0.01831”. At a significance level of 5%, the null hypothesis of the existence of a unit root can be rejected. I assume that after first-differencing, the time series data is now stationary.

```
        Augmented Dickey-Fuller Test

data:  paritydiff
Dickey-Fuller = -4.037, Lag order = 3, p-value = 0.01805
alternative hypothesis: stationary
```

```
        Augmented Dickey-Fuller Test

data:  ratingdiff
Dickey-Fuller = -4.0304, Lag order = 3, p-value = 0.01831
alternative hypothesis: stationary
```

When performing the granger causality test I also need to select an optimal lag. In this case, I use the AIC, BIC and a series of F-tests to determine the optimal number of lags. Out of 8, the optimal number of lags determined is 1. (I used the code from this stats.stackexchange.com answer to determine the number of lags: https://stats.stackexchange.com/questions/160671/estimate-lag-for-granger-causality-test)

With stationary data and an optimal number of lags, I can now perform the “granger causality” test.

```
Granger causality test

Model 1: ratingdiff ~ Lags(ratingdiff, 1:1) + Lags(paritydiff, 1:1)
Model 2: ratingdiff ~ Lags(ratingdiff, 1:1)
  Res.Df Df      F Pr(>F)
1     37
2     38 -1 0.7527 0.3912
```

I fail to reject the null hypothesis. Ratings “do not cause” Parity.

```
Model 1: paritydiff ~ Lags(paritydiff, 1:1) + Lags(ratingdiff, 1:1)
Model 2: paritydiff ~ Lags(paritydiff, 1:1)
  Res.Df Df      F Pr(>F)
1     37
2     38 -1 1.7747 0.1909
```

I fail to reject he null hypothesis. Parity “do not cause” Ratings.

Granger Causality is only met if Parity “causes” Rating and Rating does not “cause” Parity. In this case, our analysis does not match the “Rating does not seem to cause Parity” requirement. Other words, parity does not seem to “granger cause” ratings. Overall parity is not a predictor of NBA Finals Ratings at least as measured by the standards of Granger Causality.