APPLYING FRACTAL MEASUREMENTS TO ECONOMIC MODELS

VINCENT CASEY

Abstract. This paper explores fractal measurements, rescaled range analysis, and Hurst exponents as modeling techniques for economic systems using scripted methodologies so as to easily be applicable quickly.

1. Introduction

Modeling economic systems is an important part of most statistical, economic, and even general modeling systems. By applying a method to determine or measure how "random" a certain dataset is, you can easily predict how reliable a projection of said data can be. By applying methods used to measure fractals to some system, we can measure those same systems' randomness and quantify it into a reliable and easily understandable scale.

2. Overview

In order to accomplish this, we must first define some terminology and the usage of such methodologies. The material in this section was largely taken from [wah20], other sources will be cited individually. As for the rest of the paper, it will focus on applying R/S analysis and Hurst exponents to several different models and systems and the techniques used, including the process used by a calculator to do so and how the information resolved from the calculations is pertinent in its application and usage.

2.1. **Fractals.** When defining a set of data it is important to define the data relative to some measure, to have a series reference itself is to define a fractal.

Definition 2.1. Fractals are self-referencing time series. Self-referencing is defined as something referring to oneself or ones own rule set.

Some examples of fractals include the Mandelbrot set, the Koch Curve (Figure 1), and the Golden Ratio Spiral. Fractals can be found in many areas of study, from Mathematics and Physics, to Biology and Art.

There are many different types of fractals, some of which are bound to a finite space, others can go on for a near-infinite span. There are several methods for measuring fractals, all of which depend on what type of fractal they are and in which space they belong [ItF20]. The next subsection will focus on how to measure fractals and how to apply said measurements to the applications in sections 4 and 5.

2.2. **Fractal Dimensions and Ratios.** Being able to measure fractals is crucial for their study and classification.

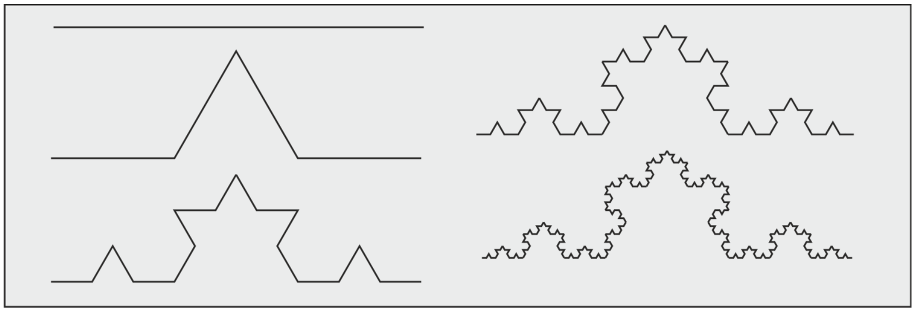

FIGURE 1. Koch Curve example taken from [Koc20]. Koch Curves have a simple rule set for generation but yield great visual complexity.

Definition 2.2. Fractal dimensions are a measure of fractals as a ratio provided through statistical analysis via the measure of complexity relative to the given scale on a time series (or smoothness).

The process to measure a fractal ratio is as follows.

(1) Begin by counting the number of circles of a fixed diameter necessary to cover the entire time series.
(2) Increase the previous diameter and again, measure how many circles are required to measure the time series.
(3) Repeat this process until only one circle is necessary.

Thus the number of circles needed to measure the time series has a direct exponential relationship with the radius of the circles. This can be expressed as:

$$D = \frac{\log(N)}{\log(\frac{1}{d})}$$

where D equals the fractal dimension, N equals number of circles, and d equals the diameter of the circle.

2.3. **R/S Analysis.**

Developed by Hurst, R/S analysis (also called Rescaled Range Analysis [nas21]) is used to determine the distance of some value over a time-series relative to itself.

Definition 2.3. We define S to be the standard deviation of some time series and R as the range of the same time series, thus R/S is the "Rescaled Range" of some time series.

From Nasdaq [nas21], they make the following comment on Rescaled Range Analysis in reference to its origin and usage:

> "The analysis developed by H.E. Hurst to determine long-memory effects and fractional Brownian motion. Rescaled range analysis measures how the distance covered by a particle increases as we look at longer and longer time scales."

2.4. **Hurst Exponents.**

The actual measure of how random a set of data is relative to itself is called the Hurst exponent. Directly correlated to fractal dimensions, Hurst exponents are also a type of measurement.

Definition 2.4. [MK00] Hurst exponents are a direct measurement of the smoothness or relative complexity of fractal dimension spaces. The measure of a fractal dimension and Hurst exponent is expressed as:

$$D = 2 - H$$

where D is the fractal dimension and H is the Hurst exponent. Thus by definition, $0 \leq H \leq 1$. When given a fractal time series, the Hurst exponent can be classified into three major categories:

- $H < 0.5$: Represents ergodic motion, meaning that with each up period the next period generally goes the opposite direction down, and vice versa. The closer that H is to 0, the more powerful this phenomenon is. The closer H is to 0.5, the more random it is.
- $H > 0.5$: Represents trend-reinforcing motion, meaning that with each up period the next period generally goes up as well, and vice versa. The closer H is to 1, the more powerful this phenomenon is. The closer H is to 0.5, the more random it is.
- $H = 0.5$: Represents random motion, meaning that given the previous period, the next period in the series has no dependency on what the previous one was.

The Hurst exponent itself is estimated by taking $\log(R/S)_n$ versus $\log(n)$ and finding the slope of best fit, where $R = \sqrt{T}$ is the distance covered by a time series T, and $(R/S)_n = cn^H$. n refers to the R/S value for $x_1, x_2, \ldots, x_n$, c refers to some constant of best fit, and H refers to the Hurst exponent.

2.5. **Time Series.** In order to measure values relative to one another in a linear fashion, we'd need to be able to define the values on a range relative to time.

Definition 2.5. A time series is a set of values measured over a period of time such that no two values have the same time position. A time position is simply a unique numerical value that represents some position on some period of time.

It is important to clarify that no two values of a single time series can hold the same time position because if there were multiple values in the same location it could cause some paradoxical calculations and wouldn't be set in real observations. For example, given a point on a car traveling from one location to another, you cannot define the vehicle to be both 4 miles and 5 miles away from its destination at any given time position.

2.6. **Economics.** When giving examples of how we can apply some function to a set of data, we must first understand what that set of data is and use a common example such that the use of the function can be understood by a wider audience. Some commonly known examples of time series include stock prices and ticker values; for this reason we will be using them in our examples below. From Merriam Webster, the definition of economics is [Web21a]:

> "a science concerned with the process or system by which goods and services are produced, sold, and bought."

To be more precise our definition for economics will focus specifically on the modeling aspect of economics, the definition of use in this document will be:

Definition 2.6. The modeling of financial systems, specifically the flow of goods and services in actively monitored markets.

And our definition of stock, taken from Merriam Webster is [Web21b]:

> "a share of the value of a company which can be bought, sold, or traded as an investment"

Note that we define a stock ticker as a specific case of some stock signifying the value of some companies stock.

Some additional definitions that may be of use for ease of reading include those for absolute value and relative value:

Definition 2.7. Absolute value is defined as the value of some commodity relative to some absolute measure. In the case of most stocks prices, this value is measured in AUSD, or Adjusted United States Dollar which simply means the value of the United States Dollar adjusted to compensate towards inflation.

Definition 2.8. Relative value is defined as the value of some commodity relative to itself represented as a percentage change from the previous absolute value to the current absolute value (the value being measured).

3. Explanation

At its very essence we are essentially applying measurement methods from fractal geometry to financial systems. This should be somewhat direct as most financial systems can be described as some form of fractal (albeit a random fractal). A random fractal, as opposed to some of the more traditional forms of fractals may not be self-referencing, instead they are a combination of rules chosen at random given different scales.

In other words, as opposed to traditional fractals where every aspect of them can be defined at each of their points, these random fractals while they do have a well-defined set of rules to follow do not have to follow a specific pattern at every point and are instead random when choosing certain aspects of themselves. So while traditional fractals are always self-referencing, random fractals while not being self-referencing, still apply the same set of rules to every aspect of themselves.

Due to the behaviors of random fractals, we can still apply the measuring techniques used when measuring non-random fractals to them. This allows us to represent any financial model as a random fractal, which in turn allows us to measure the model through fractal dimension and ratio measurements used on traditional or non-random fractal models.

4. Application

Now that we have a general idea of how to apply fractal measurements to financial systems, we must find a way to do so directly. In addition to this, given the variety of different financial systems and how their data is collected and measured relative to other systems, we must find unique fractal measurement solutions to different forms of data collected.

There are two methods used here for measuring the value of each time series, the first is absolute, and the other is relative. Absolute values are the completely unedited data inputted into our formula. Relative values are the absolute values taken relatively to the absolute time series values immediately prior.

Note that all calculations performed in the examples are done using a program I developed in the Python scripting language [Cas21]. Complete references to how the program was developed are cited in the program itself. Below in the calculation subsection 4.1 is a brief overview of how it works.

4.1. **Calculation.** It is explained here how to find the Hurst exponent of some one-dimensional time series. This time series can be absolutely or relatively defined, so long as it has one dimension (meaning that it only has one value per time "location") a Hurst exponent can be found using these methods.

The first method explained is through manual calculation (Manual Calculation 4.1.1), it explains how to find the Hurst exponent manually. The second method explained (Calculator Explanation 4.1.2) is that used by my calculation script [Cas21], it explains how the script calculates the Hurst exponent given some CSV data.

4.1.1. *Manual Calculation.*

(1) Estimate the maximum number of R/S values viable by taking the square root of the number of values in the time series
(2) Find the standard deviation and range of the time series for each value in the series from the beginning of the series to said value.
(3) Calculate $\log(R/S)$ for each value.
    (a) Define S to be the standard deviation of the time series from some point in the series including all the points prior to it inclusively.
    (b) Define R to be the range of the same series.
    (c) Plug S and R into $\log(R/S)$ and calculate
(4) Calculate $\log(N)$ for each value.
    (a) Define N to be the index of a value in the time series relative to the first value.
    (b) Plug N into $\log(N)$ and calculate.
(5) Plot $\frac{\log(R/S)}{\log(N)}$, the line of best fit is the Hurst exponent.

4.1.2. *Calculator Explanation.* Firstly, to perform a calculation of the Hurst exponent we need some data, thus we get the data from a CSV file (a type of file used to store large amounts of data, formatted similar to Microsoft Excel or Google Sheets). We used CSV files with data in the format used by Yahoo Finances downloadable informational files on stock prices in the program. Then, we calculate the relative data by finding the percentage change from each value to the value prior (for the first value, we use itself, i.e. no change or a relative value of 1). If, for example, the prices for some stock over a times series were \$2.00, \$3.00, and \$2.00 the relative values would be 1.00, 1.50, and 0.66 (100%, 150%, and 66% respectively).

After this we calculate the Hurst exponent of either the values themselves or their relative change in value (as mentioned prior) depending on what values we decided to calculate. To do this we find the maximum number of R/S values technically viable given a time series, then calculate the range and standard deviation for the time series given each new value and append each of these to a list of values along with their reference value n (an integer referencing each set of values). After this we calculate $\log(R/S)$ where R/S is found by defining S to be the standard deviation of the time series prior to the data point (inclusive) and R as the range of the same time series, thus R/S is found by dividing R by S (Subsection 2.3). After taking $\log(R/S)$ for each of the sets of values mentioned prior we finally find the line of best fit for $\frac{\log(R/S)}{\log(N)}$ of each unique value of N, thus calculating the Hurst exponent of our given time series.

After this, the results are then graphed $\frac{\log(R/S)}{\log(N)}$ and finally displayed to the user along with the aforementioned Hurst exponent. With slight modification to the code it is also possible to graph the raw data, graph the relative values of the data, or to simply just display the Hurst exponent alone. Again, for further explanation on how these calculations are performed, please visit the program itself [Cas21] where the code is documented per each step of the process.

Please note that the methodology used in this program to calculate the Hurst exponent can be applied to almost any data in the correct format and that it is relatively easy to do so. In addition to this, it is almost trivial to apply the program to other downloadable data taken from Yahoo Finance on almost any stock, listed cryptocurrency, or index fund.

5. Examples

Below are examples of several financial stocks and their values represented in $\frac{\log(R/S)}{\log(N)}$ recorded over different spans of time from the past 1 to 5 years or the stock's value throughout its entire lifespan. All values are measured in their relative dollar value, meaning that the values recorded are adjusted to the inflation rate of the United States Dollar (USD); this is done so as to remove some bias.

Some interesting things of note include how relative to SBUX (Starbucks stock ticker) (Figure 2), GME (Gamestops stock ticker) (Figure 3) has an exceptionally low Hurst exponent, this would make sense as over this past year (2020-2021) there has been much turmoil in its financial value and professional opinion on its value relative to pretty much any other stock currently on the market.

Note that all financial data in this paper was gathered from Yahoo Finance [Yah21] and that all graphs and Hurst values from this point forward (unless mentioned otherwise) were generated using my Hurst exponent calculator [Cas21] mentioned in subsection 4.1.

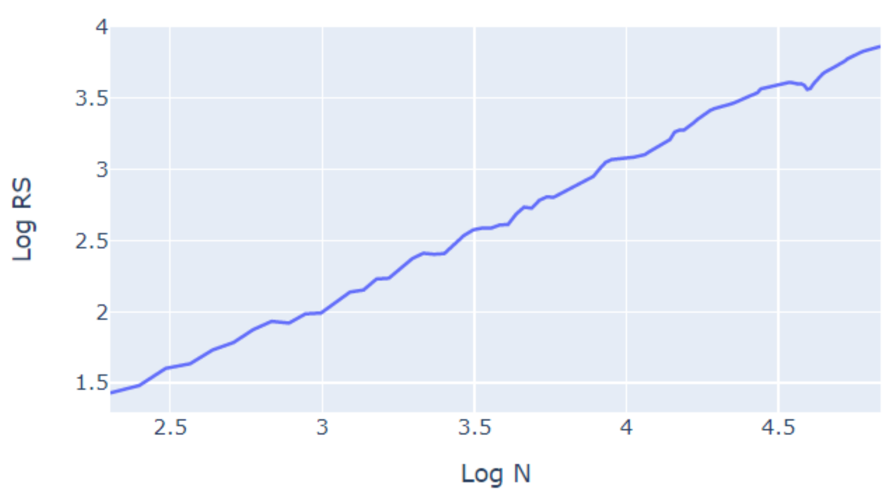

Figure 2. $\frac{\log(R/S)}{\log(N)}$ from SBUX (Starbucks stock ticker) adjusted closing value over the past year (2020-2021) with a resulting Hurst exponent of 0.9964 (rounded to 4 digits)

FIGURE 3. $\frac{\log(R/S)}{\log(N)}$ from GME (Gamestop stock ticker) adjusted closing value over the past year (2020-2021) with a resulting Hurst exponent of 0.9041 (rounded to 4 digits)

Another example, MSFT (Microsofts stock ticker) (Figure 4) has a significantly higher $\log(R/S)$ value compared to either of the previous examples. This is sensible for the reason that the data gathered for the information displayed is taken from a 5 year span as opposed to a single year span, which causes there to be a bias in increased value due to there being a higher chance of it being valued at a higher price over longer periods.

Some interesting things of note with these figures is that they all have seemingly high Hurst exponents, this makes sense. However as they're all based on measurements of absolute adjusted values as opposed to relative adjusted value (percentage change of the adjusted value from one time series point to the next). When measuring the relative adjusted value rather than the absolute value of some stock, we will find that the resulting Hurst exponent can be drastically different.

Given MSFT's (Microsoft stock ticker) relative monthly value however, the Hurst exponent is much more in line with what is expected for the company's valuation, with it being roughly equal to 0.5347 (rounded to 4 digits). Meaning that while its valuation seems to be relatively random from a month to month basis, it is still far from being completely such. Below are two charts, MSFT's monthly relative value (Figure 5) and MSFT's monthly relative value $\frac{\log(R/S)}{\log(N)}$ chart (Figure 6) visualizing our findings.

Taken from the same period of time (from Microsoft's initial public listing in 1986 to 2020, a span of around 34 years) the Hurst exponent of its yearly relative values (Figure 8) is significantly high at 0.5626 compared to its monthly relative values over the same span of time with a Hurst exponent of 0.5347 (both results rounded to 4 digits) (Figure 5). This observation is delved into much more depth in Observations (subsection 5.1), where it attempts to explain why this phenomena occurs.

5.1. **Observations.** An interesting observation from the MSFT example is that depending on how often you decide to measure the given data collected, the Hurst exponent will differ. This means that daily, monthly, or even yearly relative values over a long enough span of time seem more random (or to have more Brownian motion) compared to values with larger

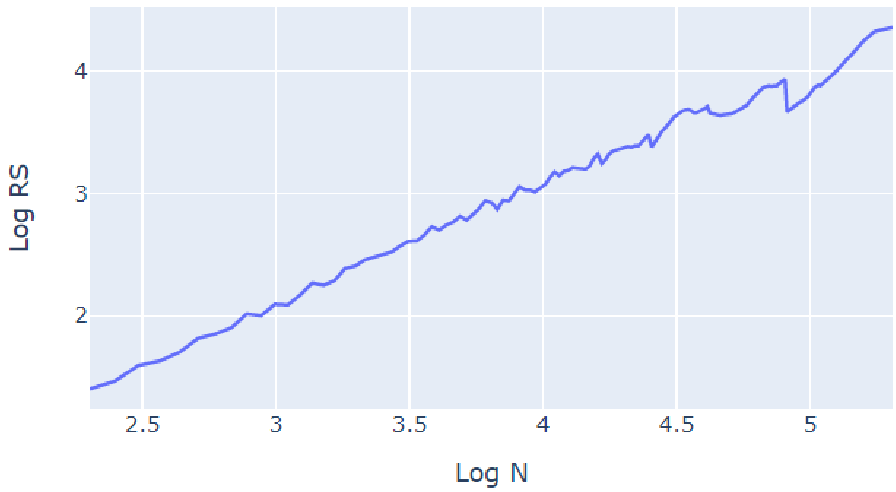

FIGURE 4. $\frac{\log(R/S)}{\log(N)}$ from MSFT (referring to the company Microsoft's stock ticker) adjusted monthly closing value since its creation with a resulting Hurst exponent of 0.9302 (rounded to 4 digits)

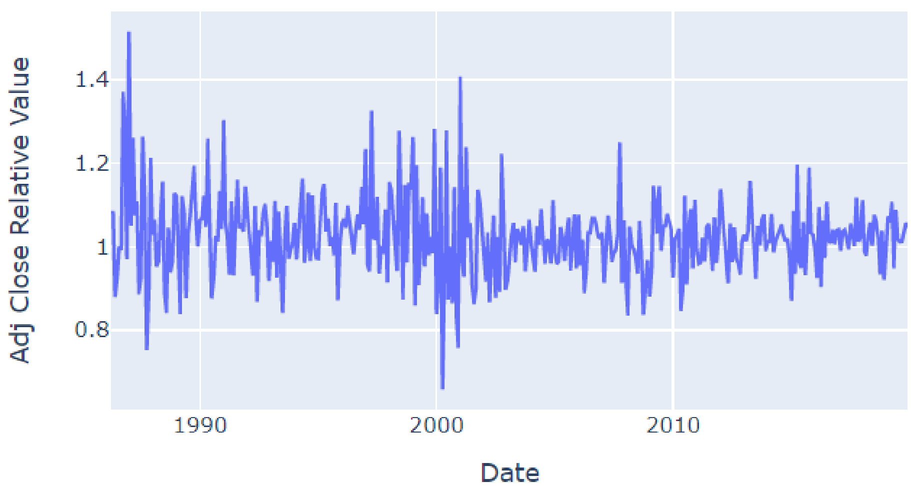

FIGURE 5. MSFT's all time (data measured from 1986-2020) adjusted closing monthly (meaning the value of said stock at the end of each month adjusted to inflation) relative values

spans of distance between them. For example, given monthly versus yearly values, it is more likely that the daily relative values will be closer to 0.5 than the monthly relative values over the same span of time. This phenomena is directly observed in our MSFT examples where we find that given yearly observations, the Hurst exponent is significantly higher at 0.5626 (Figure 8) relative to its monthly observations over the same span of time with a Hurst exponent of 0.5347 (Figure 5).

Some possible applications for these findings include such things as finding the optimal frequency to adjust investment strategies found by comparing Hurst exponents with data measured at different intervals and when to decide how risky the current market is (which

FIGURE 6. MSFT's all time adjusted closing monthly relative values plotted to $\frac{\log(R/S)}{\log(N)}$ resulting in a Hurst exponent of 0.5347 (rounded to 4 digits)

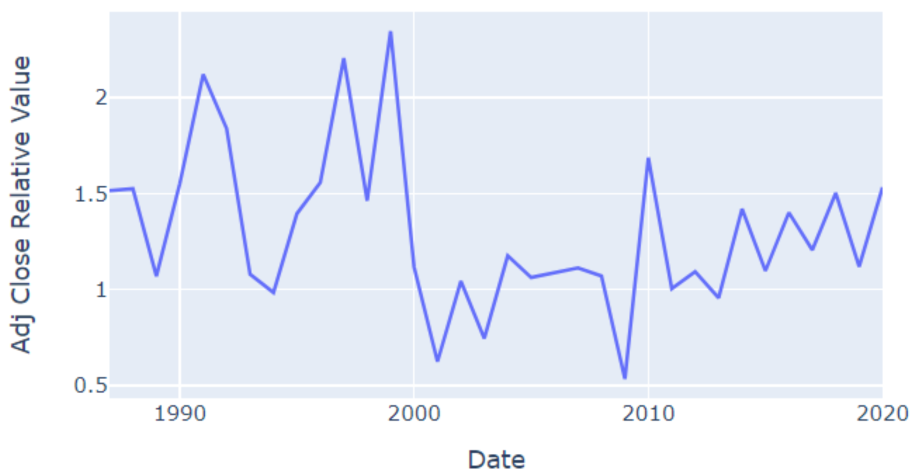

FIGURE 7. MSFT's all time adjusted closing yearly relative values (meaning the relative values of the stocks price at the end of some year relative to the year priors portrayed as a percentage change)

might be found by finding how much (or how little) the Hurst exponent changes over time). This is due to the fact that when the Hurst exponent is closer to 0.5, there is a higher chance that either the investment of purchasing or the selling of some stock tends to be more difficult to predict.

An additional somewhat interesting observation is how much higher SBUX's Hurst exponent was compared to GME's. This aligns with expert opinions such as that found on Aswath Damodaran's (an educator of corporate finance at the Stern School of Business at New York University) blog [Dam53] where he states that:

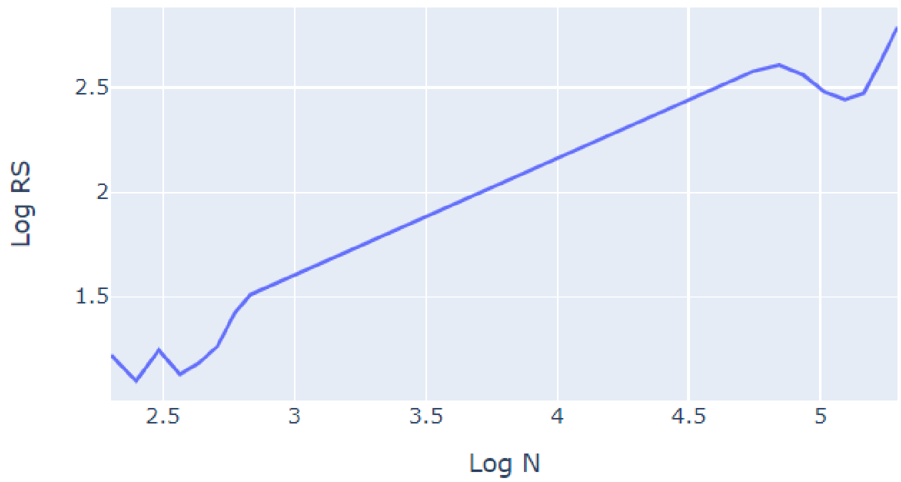

Figure 8. MSFT's all time adjusted closing yearly relative values plotted to $\frac{\log(R/S)}{\log(N)}$ resulting in a Hurst exponent of 0.5626 (rounded to 4 digits)

> "You may get lucky and be able to exit before everyone else tries to, but the risk that you will be caught in a stampede is high, as everyone tries to rush the exit doors at the same time."

in reference to the valuation of GME relative to other stocks and the near impossibility of predicting whether you could sell a stock for some profit. Comparing GME to a relatively safe investment of SBUX (Starbucks stock ticker) it is easy to see the correlation between relative perceived safety and the Hurst exponent of some investment.

Another important observation is that the relative value and actual value of any stock portrayed as a time series can have massively different Hurst exponents. There are a few reasons for this bias. First, there is likely some bias in the selection of stocks themselves as any company that has a decent amount of data must be multiple years (or even decades) old and thus must be competing against the rest of the market relatively well as if they were not the company would not exist. Another reason for this is that when measuring value over time, a company that increases in value will have an absolute change in value from year to year, whereas the same company could still have a lower percentage return from one year to the next (as compared to the year prior) and still gain in actual value. Thus the actual and relative values can result in drastically different results.

6. Conclusion

While there might be some use of this methodology in determining how "risky" an investment is, it should be noted that while it may seem statistically wise to do so all data is taken strictly from previous price points and thus has absolutely no influence from the real world and should not be the only source used for testing reliability. Any method used for modeling value over time can be an incredible tool, if that tool can accurately define the future of some value however is completely up to opinion as said value cannot truly be known until it is reached (and thus needs no prediction).

References

[Cas21] Vincent Casey. hurst-exponent-calculator, 2021. https://github.com/vmc-7645/hurst-exponent-calculator-from-yahoo-csv.

[Dam53] Aswath Damodaran. The gamestop frenzy: Lessons and consequences, 29-Jan-2021 12:53. http://people.stern.nyu.edu/adamodar/pdfiles/blog/GameStop.pdf.

[ItF20] ItF. Introduction to fractals, 2020. http://paulbourke.net/fractals/fracintro/.

[Koc20] KochCurveIMG. Koch curve illustration, 2020. https://sites.google.com/a/rug.nl/aperiodic-tiling/2x2-supertiles-1/von-koch.

[MK00] Ferry Butar Malhar Kale. Fractal analysis of time series and distribution properties of hurst exponent, 2000. Journal of Mathematical Sciences & Mathematics Education, Vol. 5 No. 1.

[nas21] nasdaq. Rescaled range analysis, 2021. https://www.nasdaq.com/glossary/r/rescaled-range-analysis.

[wah20] wahl.org. Calculating fractal dimension, 2020. https://www.wahl.org/fe/HTML_version/link/FE4W/c4.htm.

[Web21a] Meriam Webster. Economics definition, 2021. https://www.merriam-webster.com/dictionary/economicss.

[Web21b] Meriam Webster. Stock definition, 2021. https://www.merriam-webster.com/dictionary/stock.

[Yah21] Yahoo. Yahoo finance, 2021. https://finance.yahoo.com/.

Advisor: Phd. Indranil Sengupta

vincent.casey@ndsu.edu

Department of Mathematics, North Dakota State University, PO Box 6050, Fargo, ND 58108-6050, USA