Nucleation of Asset Bubbles

Jacob Rider, *Arun Suresh, Chima Ani, and Zach Langford*

December 2019

Abstract

This project explored the relationship between homogeneous bubble nucleation in a physical sense and the formation of asset bubbles in a financial domain. The structure of the model was created using a combination of qualitative and dimensional analysis with the JMAK equation as a starting point. Data was then gathered from the federal reserve and bloomberg terminal API. Subsequently, the model was implemented using R, and results were interpreted to indicate a correlation between variables and the formation of asset bubbles. This naturally gives the future direction to quantify this correlation and test the predictive accuracy of our model.

1 Introduction

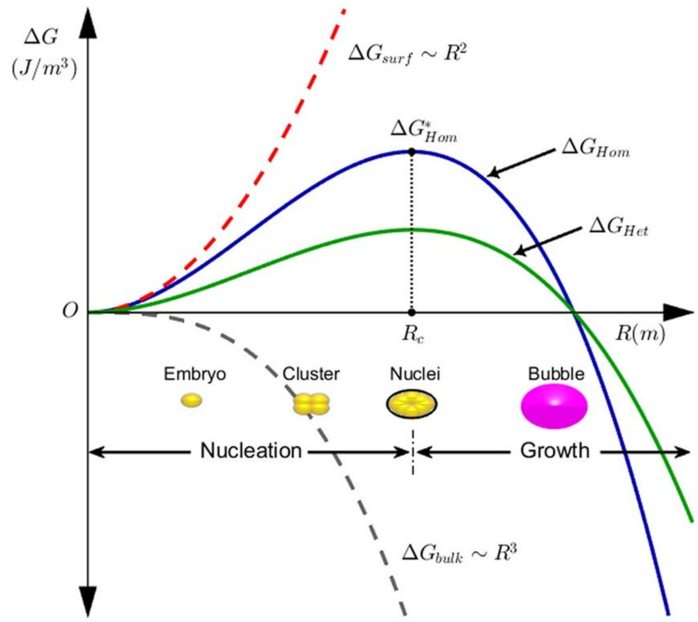
Financial asset bubbles have been around for centuries, and for much of that time, people have been puzzled by the problem of predicting the behavior and existence of these peculiar events. There is a massive amount of research on the prediction of asset bubbles, but none of it attempts to model asset bubbles using the structure of physical bubble nucleation. During the preliminary research phase of this project, we explored many different asset bubble models. While these models gave us a direction to start, they all fundamentally differed in structure from our project, so we were unable to cite any specific contributions from research papers. This paper will propose a model that connects physical equations from the homogeneous nucleation model with key economic variables that may correlate with financial bubbles. Additionally, this paper will attempt to answer the following question:

To what extent can homogeneous nucleation equations be used to model financial bubbles?

2 Background

2.1 Physics Background

The purpose of this paper is to determine the extent to which homogeneous nucleation equations can be used to model financial asset bubbles. In order to accomplish this, we must first introduce the homogeneous nucleation model that will provide a structure for our economic model [1]. The following picture depicts a bubble forming after its radius grows past the critical radius r^*. This is a critical component of both the homogeneous nucleation model and our asset nucleation model.



$$\frac{\text{Volume of bubbles in system}}{\text{Volume of system}} = \frac{V_\beta}{V} = Y = 1 - \exp(K \cdot t^4) \tag{1}$$

$$K = \frac{\pi \dot{N} \dot{G}^3}{3} \tag{2}$$

$$\textbf{Nucleation Rate: } \dot{N} = a \exp\left(\frac{\Delta G^*}{f_o T}\right) \tag{3}$$

$$\textbf{Bubble Growth Rate: } \dot{G} = \frac{dr}{dt} \tag{4}$$

$$\Delta G^* = \frac{16\pi\gamma^3}{3(\Delta G_v)^2} \tag{5}$$

$$\textbf{Critical Radius: } r^* = \frac{2\gamma}{\Delta G_V} \tag{6}$$

$$\textbf{Change in Gibb's Free Energy: } \Delta G_V = \frac{L_V \Delta T}{T_m} \tag{7}$$

The physical properties of this model are undoubtedly important to grasp, but my role focused on the behavior of each variable and the economic parallels. As a result, this paper will primarily focus on the economic explanation of this model.

2.2 Economics Background

Base assumptions:
1) There are two possible phases for an asset: the normal phase and the bubble phase. We hypothesized that a stock enters bubble phase when $r > r^*$
2) The market cap of a stock in our economic model is assumed to be equivalent to volume in the physical model. For lack of a better term, we define a "prepice" to be equivalent to radius in a similar way.
3) The assets in an industry can be indexed as follows:

$j = 1, 2, ..., n$: companies in an industry
$i = 1, 2, ..., m$: industries in a market
V_j : market cap of a company; V_i : market cap of an industry

$$\sum_{j=1}^{n} V_j = V_i$$

Qualitative Analysis
The goal of qualitative analysis is to identify the role of each variable in the nucleation equation and to determine an economic parallel variable that has a corresponding role in the asset bubble model. Before introducing the main equations, a few definitions will be helpful to understand the key variables in our model.

Definition. *Asset Bubble: A period of market speculation characterized by a large increase in price followed by a sharp correction known as a "crash".*

Definition. *Market Cap: The total value of a company calculated by price per share * shares outstanding.*

Definition. *Book Value: The net value of all tangible assets on a company's balance sheet.*

Definition. *Cost of Revenue: The total expenses of a company minus the cost of building inventory.*

Definition. *Free Cash Flow: The cash remaining in a company after expenses.*

Definition. *Volatility: The standard deviation of a stock's price over a specific time period.*

Definition. *Prepice: A variable proportional to the cubic root of market cap.*

The equations are systematically introduced as follows:

$$(1) \Rightarrow Y = 1 - \exp[-k^2 \cdot t^4] \tag{8}$$

$$(2) \Rightarrow k = a \cdot \dot{N}\dot{G}^3 \tag{9}$$

$$(3) \Rightarrow \dot{N} = f \cdot c_0 \cdot \exp(\frac{\Delta G^*}{T}) \tag{10}$$

$$(4) \Rightarrow \dot{G} = \frac{dr}{dt} \tag{11}$$

$$(5) \Rightarrow \Delta G^* = b \cdot \frac{\gamma_j^3}{(\Delta G_{V_j})^2} \tag{12}$$

$$(6) \Rightarrow r_j^* = \frac{\zeta \cdot \gamma_j}{\Delta G_{V_j}} \tag{13}$$

$$(7) \Rightarrow \Delta G_{V_j} = |\nu_{actual} + \nu_{implied}| \cdot |\frac{F_j}{V_j}| \tag{14}$$

$$\gamma_j = \frac{B_j}{V_j} \sqrt[3]{C_j} \tag{15}$$

$$T = \frac{V_i}{\psi + 1} \tag{16}$$

$$c_0 = \frac{n \cdot r_i^{*3}}{V_i^2} \tag{17}$$

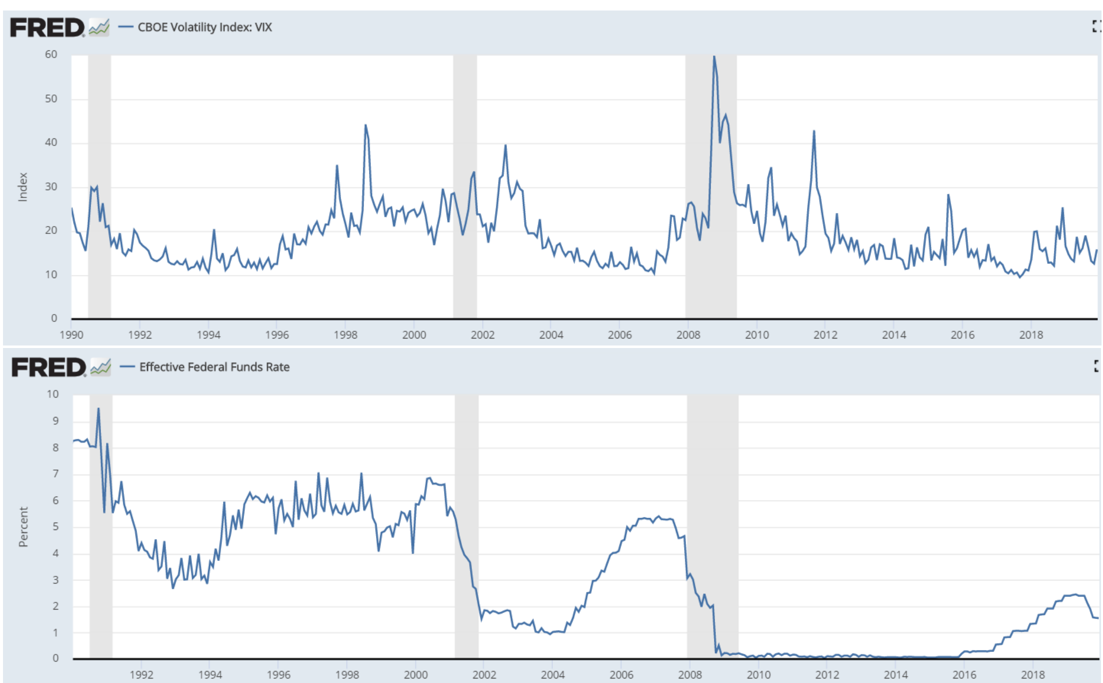
These variables are defined in the table 1 in the appendix and the justification for our choice of economic variables is explained in table 2. In equation 16, the denominator is interest rate plus 1 to avoid divide by zero errors in the case of extremely low interest rates. It is unlikely for interest rates to fall below -1% so this is a reasonable assumption to make.

3 Method (to the madness)

3.1 Data

Datasets are separated into two categories: Macro and Micro. Macro denotes variables that are constant across the population such as interest rate, and micro denotes variables that are company specific such as book value.

Macro: The federal reserve website is one of the most reputable and exhaustive sources of macroeconomic data. Two important datasets from this source include the federal funds rate (interest rate) and the CBOE implied volatility index (shown below) [2][3].



Micro: Company specific data was obtained using the bloomberg terminal API. Data was taken for the following variables from 12/31/1989 to 11/21/2019 for 2500 stocks in the New York Stock Exchange (NYSE): Market Cap, Shares Outstanding, Book Value, Cost of Revenue, Free Cash Flow, and 30 Day Volatility. Two complications arose from the microeconomic data: Data Cleaning and Survivorship Bias.

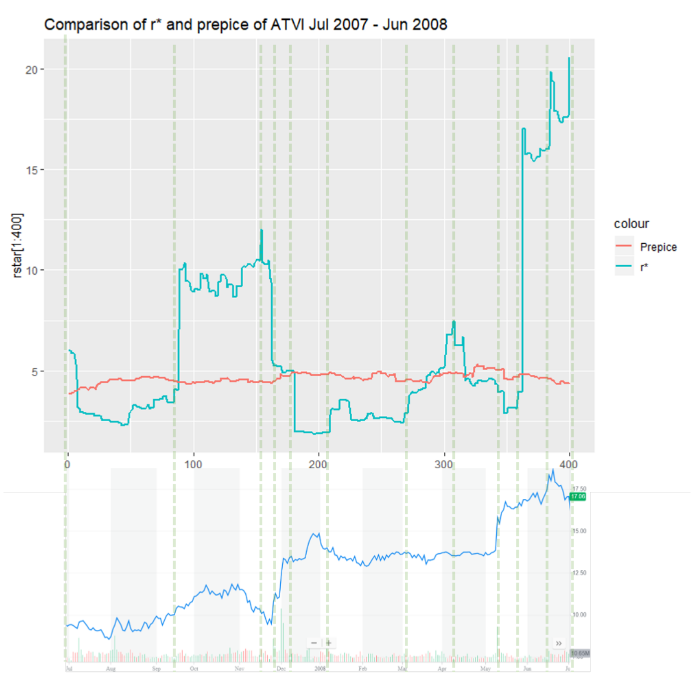
Data Cleaning: Bloomberg variables are updated with varying frequencies so many rows were either repeated values or N/A. This forced us to use a broad sample of the dataset to avoid redundancy and missing values. Since missing values were less common closer to present-day, we focused our analysis on the 2007 crisis.

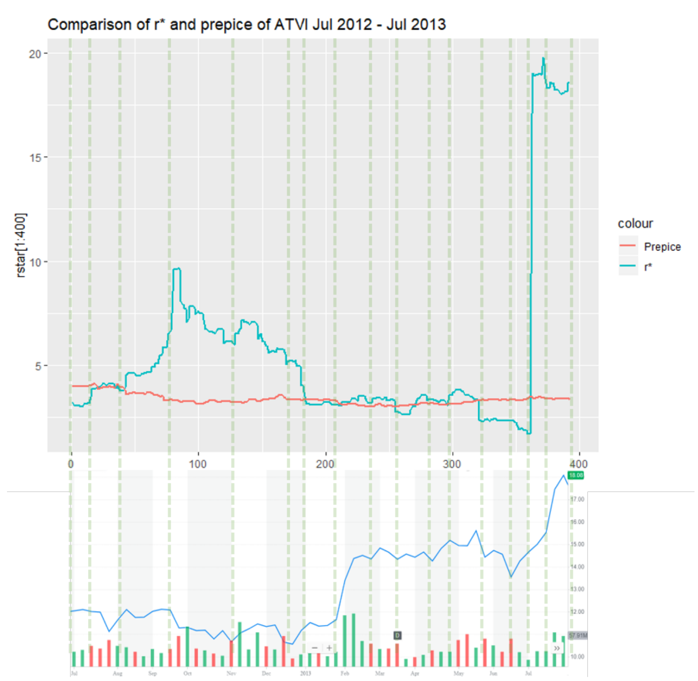
Survivorship Bias: Another limitation of our dataset was that it only included companies that exist in the present. This causes our model to incur an implicit, but unavoidable survivorship bias. This bias certainly skews the results of our model and eliminating it would be crucial to the future success of this model.

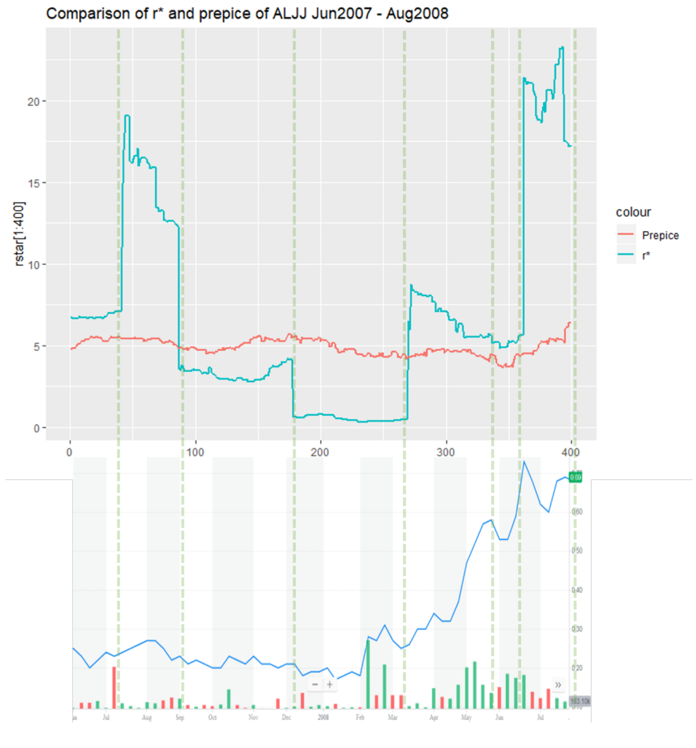
3.2 Model

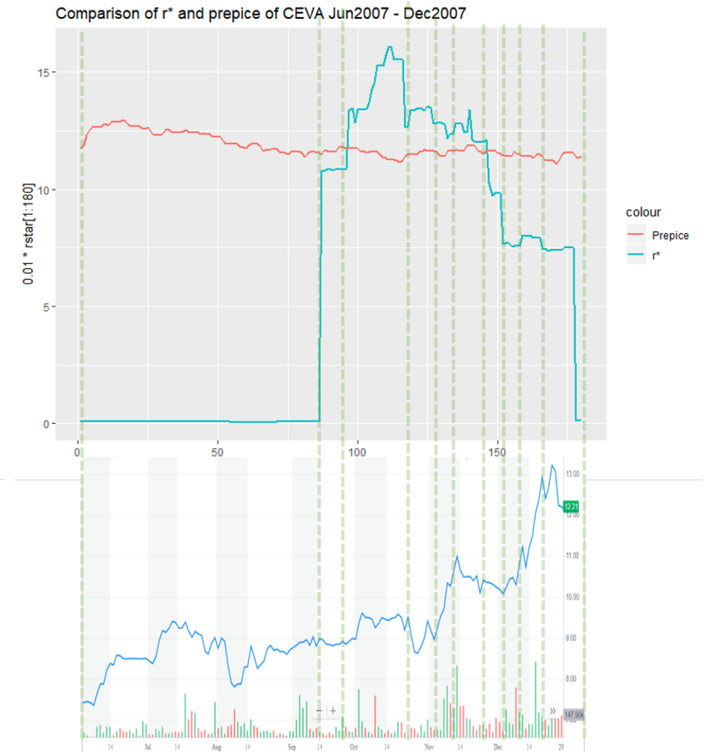
Our model was programmed in R to maximize the efficiency of our limited computational capacity. Stocks were analyzed within set time intervals because it was not feasible to run the model over the whole time period (1989 to Present). The period following July 2007 was the most commonly analyzed period, but we also looked at data from 2012-2013 and 2018-2019.
Some additional alterations were made to are model in the process of implementation. First, $\dot{G}$ was set to be a constant over a given time period as follows: $\dot{G} = \dfrac{dr}{dt} = \dfrac{r_f - r_0}{\Delta t}$. Additionally, c_0 and n were averaged over the time period to keep the model consistent without altering the dimensions. Averaging these variables was the most accessible option to simplify our model. This decision allowed us to focus primarily on the change in r^* over specific time intervals. Given these limitations, we were still able to extract some conclusions from the r^* graphs for specific stocks and the Y graphs for industry averages. These graphs are interpreted in the following section.

4 Results





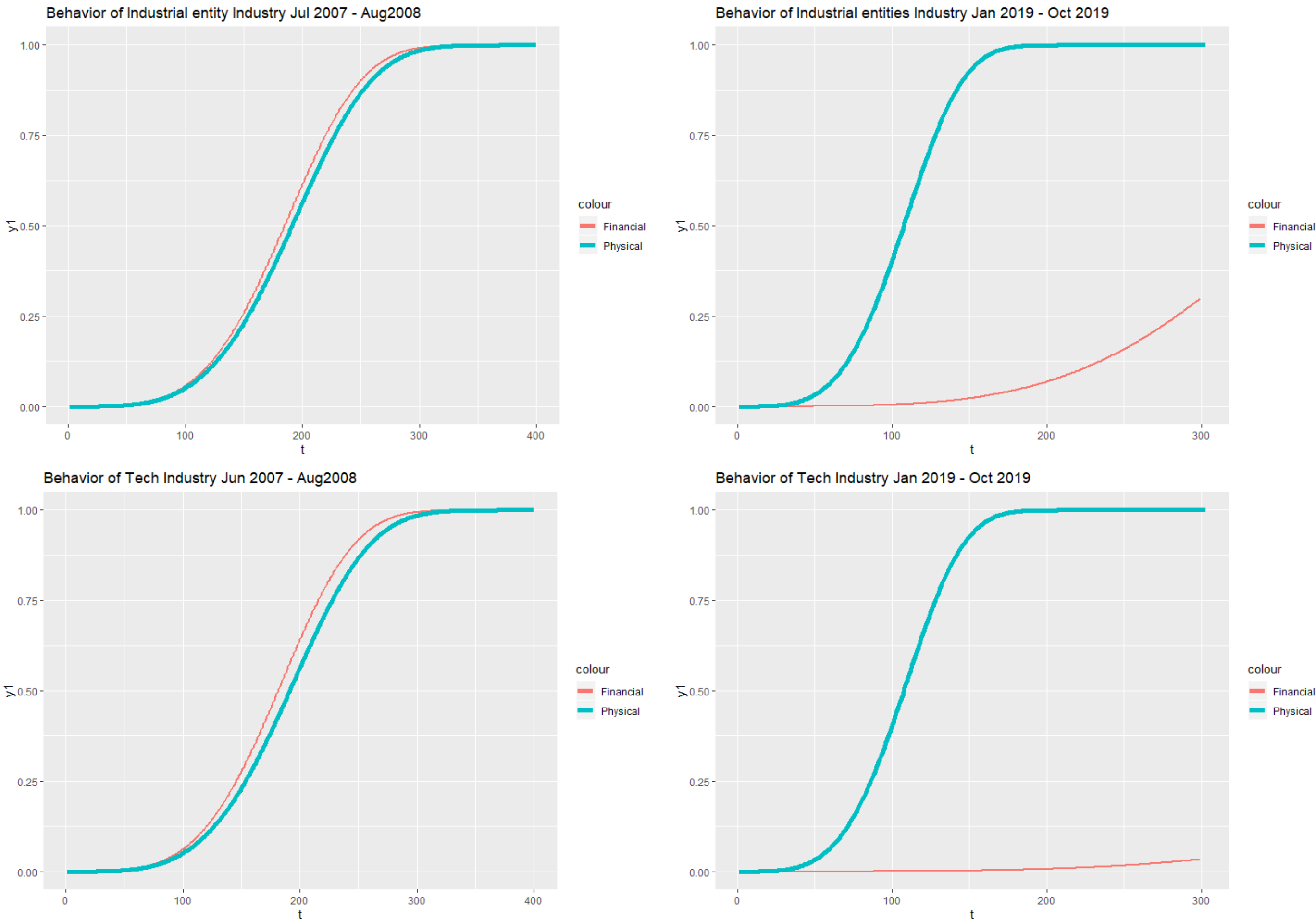




4.1 Stock Analysis

Description: Each of these graphs plots the critical prepice (top, light blue) vs. the price (bottom, dark blue) over a specific time period. The vertical green lines were inserted to compare specific points in time between each pair of graphs. The top graphs depict Activision (ATVI) from 07/01/2007 to 07/31/2008 (top left)[4] and from 07/01/2012 to 07/31/2013 (top right)[4]. The bottom graphs show the behavior of ALJJ [5] and CEVA [6] during the 2007 crisis. These stocks were selected as a random sample of our data.

Hypothesis Evaluation: Initially, we expected the critical prepice to signal a stock entering the bubble phase when $r > r^*$. Our model failed to support this hypothesis. Instead, we observed that sharp changes in r^* often (but not always) indicated local critical points (maxima or minima) on the price over time graph. r^* appears to be most sensitive to changes in free cash flow (which can be positive or negative). Our model takes the absolute value of this term to avoid an r^* term with negative dollars. As a consequence to this decision, the direction of changes in r^* is not significant, but the magnitude of the change is. This is far from a rigorous analysis of the predictive accuracy of our model, but it is the most significant conclusion that we could find given a time constraint and lack of computational resources. This conclusion does provide the insight that the variables (that make up r^*) play a significant role in predicting the critical points of price. One specific example of a spike in r^* providing insight into the behavior of a stock is the last spike in the top right graph. This spike coincides with Blizzard's acquisition of Activision which was followed by a sharp increase in stock price. This is simply one example of a phenonema that shows up throughout these graphs.



4.2 Industry Analysis

The graphs above show the aggregate Y value for two industries (Technology and Industrial Entities) over two time periods (2007 - 2008 and 2019 - Present). The red curve represents the Y value for the industry and the blue curve represents a theoretical physical model that is scaled to represent the nucleation process over the same time period. During the 2007 crisis, the economic model closely matches the theoretical physical model with errors of 8.01% (Technology) and 11.2% (Industrial). This indicates that the economic model closely matches the physical model during a defined bubble (the 2007 crisis). In the present-day case, the curves clearly do not indicate that the industries are in a bubble phase, but there is a trend that may be showing the genesis of this process.

5 Discussion and Further Directions

5.1 Discussion

Collectively, our group was introduced to topics including, but not limited to: homogeneous nucleation, qualitative and dimensional analysis, bloomberg API, and programming in R. Although the end result of our model did not closely match our hypothesis, we were still able to acquire valuable experience on the modelling process and extrapolate some meaningful insights from our results.

5.2 Takeaways

The hypothesis that a company enters the bubble phase when $r > r^*$ did not hold, but we observed that abrupt changes in r^* often correlated with critical points in the price over time graph. This conclusion indicates that the variables that make up r^* (Book Value, Market Cap, Cost of Revenue, Free Cash Flow, and Volatility) are correlated with changes in bubble phase which are indicated by critical points on the price over time graph.

To what extent can homogeneous nucleation equations be used to model financial bubbles?

There is some evidence (specifically the aggregate industry graphs in 2007) to suggest that financial bubbles can be effectively modelled using homogeneous nucleation equations. That being said, our model certainly has limitations that prevent us from answering this question with certainty. For example, we were unable to define an equation for r^* that confirmed our hypothesis, but we cannot conclude that such an equation does not exist. Therefore, the homogeneous nucleation structure has potential to model financial bubbles, but our model is not robust enough to conclude any predictive capabilities.

5.3 Further Directions

Unfortunately, we were unable to quantify the significance of each variable to determine the most important components of our model. With more time and computational capacity, we would ideally test the significance of each variable as well as the predictive accuracy of our model in different scenarios. Additionally, we could attempt to restructure an equation for r^* that behaves in the way our hypothesis originally predicted. Our model includes certain parameters that are constant over a specified time interval, but change values when the time interval changes. This indicates the existence of exogeneous variables that our model fails to incorporate. In other words, there is evidence to suggest our model both includes and neglects significant variables for the prediction of asset bubbles. Lastly, there may be a way to gather data on ”dead” companies that would eliminate the survivorship bias in our model.

6 References

1. Porter, David A., Kenneth E. Easterling, and Mohamed Sherif. Phase Transformations in Metals and Alloys, (Revised Reprint). CRC press, 2009.

2. Chicago Board Options Exchange, CBOE Volatility Index: VIX [VIXCLS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/VIXCLS, December 10, 2019.

3. Board of Governors of the Federal Reserve System (US), Effective Federal Funds Rate [DFF], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DFF, December 10, 2019.

4. Activision Blizzard, Inc. (ATVI). December 10, 2019. Yahoo! Finance. Retrieved from https://finance.yahoo.com/chart/ATVI?p=ATVI

5. ALJ Regional Holdings, Inc. (ALJJ). December 10, 2019. Yahoo! Finance. Retrieved from https://finance.yahoo.com/chart/ALJJ?p=ALJJ

6. CEVA, Inc. (CEVA). December 10, 2019. Yahoo! Finance. Retrieved from https://finance.yahoo.com/chart/CEVA?p=CEVA

7. *Data and Code files are available upon request.*

7 Appendix

Table 1: Variables

Symbol	Physics Meaning	Economics Parallel
$\mathbf{Y} = \frac{V_\beta}{V}$	Ratio of the volume of bubbles in a system to the volume of the entire system.	Ratio of the market cap of bubbling assets in an industry to the market cap of the whole industry.
$\dot{N}$	Rate of nucleation (i.e. rate of new bubbles forming in the system).	Rate of companies changing from normal phase to bubble phase.
$\dot{G}$	Rate of radius growth of bubbles in the system.	Rate of ”prepice” growth of stocks in the bubble phase.
$V = \frac{4\pi r^3}{3}$	Volume as a function of the radius.	Market cap as a function of the ”prepice.”
γ	Interfacial free energy; multiplied by area to give surface tension.	Market forces that prevent bubbles from forming.
ΔG_V	Free energy per unit volume	Market forces that cause bubbles to form.
r	Radius of bubble	The ”prepice” is a variable proportional to the cubic root of market cap.
r^*	Critical radius; the radius at which a bubble is formed.	Critical prepice; the prepice at which an asset transitions from to the bubble phase.
ΔG^*	Critical free energy; the free energy at which a bubble is formed.	The ratio of balancing markets forces that indicates a transition to the bubble phase.
c_0	Number of atoms in a system per unit volume.	The ratio of companies entering bubble phase and the market cap of an industry.
n	Number of atoms in a system.	Number of companies in bubble phase.
T	Temperature	Variable that measures the macroeconomic environment with the ratio between the market cap of an industry and interest rate.
ψ	N/A	Interest Rate
L_V	Latent heat of fusion per unit volume.	N/A
ΔT	Undercooling	N/A
T_m	Melting temperature.	N/A
V_B	N/A	Book Value
ν	N/A	Volatility (implied or actual)
F_j	N/A	Free Cash Flow
C_j	N/A	Cost of revenue
a, b, f, ζ	Constant parameters	Constant parameters

Table 2: Variables

Variable	**Positive or Negative**	**Correlation Justification**	**Qualitative Justification**
V_i	Positive	$V_i \uparrow \Rightarrow T \uparrow \Rightarrow \dot{N} \uparrow$	An increase in the market cap of an industry should increase the rate of new bubbles forming because there is more total capital to supply each bubble.
V_j	Positive	$V_j \uparrow \Rightarrow \gamma_j \downarrow \Rightarrow \Delta G^* \downarrow \Rightarrow \dot{N} \uparrow$	An increase in the market cap of a company should increase the rate of nucleation because a bubble forms when market cap reaches a critical point.
F_j	Positive	$F_j \uparrow \Rightarrow \Delta G_V \uparrow \Rightarrow \Delta G^* \downarrow \Rightarrow \dot{N} \uparrow$	An increase in free cash flow should increase the rate of nucleation because this is a commonly used metric that is known to inflate the value of a stock.
ν	Positive	$\nu \uparrow \Rightarrow \Delta G_V \uparrow \Rightarrow \Delta G^* \downarrow \Rightarrow \dot{N} \uparrow$	An increase in volatility (both implied and actual) should increase the rate of nucleation because historical data shows that volatility is increased during asset bubbles.
C_j	Negative	$C_j \uparrow \Rightarrow \gamma_j \uparrow \Rightarrow \Delta G^* \uparrow \Rightarrow \dot{N} \downarrow$	An increase in the cost of revenue decreases the rate of nucleation because this metric is hypothesized to deflate the value of a stock.
B_j	Negative	$B_j \uparrow \Rightarrow \gamma_j \uparrow \Rightarrow \Delta G^* \uparrow \Rightarrow \dot{N} \downarrow$	An increase in book value decreases the rate of nucleation because companies with a higher value of tangible assets are less likely to develop into bubbles.