

Entropy, the Information Processing Cycle, and the Forecasting of Bull and Bear Market Peaks and Troughs

Edgar Parker, New York Life Insurance Company, Bayonne, USA

ABSTRACT

Many econophysics applications have modeled financial systems as if they were pure physical systems devoid of human limitations and errors. On the other hand, traditional financial theory has ignored limits that physics would impose on human interactions, communications, and computational abilities. The entropic yield curve blends the physical and human financial worlds in a new, powerful, and surprisingly simple way. This article uses this information theoretic perspective to provide a new explanation of the dynamics and timing of financial cycles. Additionally, the entropic yield curve offers a new method of forecasting market peaks and troughs.

KEYWORDS

Bear Market, Bond, Bull Market, Burnashev, Econophysics, Entropy, Equity, Information Theory, Shannon Entropy

INTRODUCTION

The modeling and forecasting of the current and future states of financial systems is of extreme importance to a wide variety of financial actors. The appropriateness of portfolio weightings and other assumptions vary greatly depending on the financial cycle. More accurate determination of the timing of the emergence of bull and bear markets could vastly improve the performance of financial products while simultaneously reducing overall risk.

A bear market is indicative of a severe slowdown in the economy. A popular definition is a decline of 20 percent or more in a representative equity market. On the other hand, a bull market is generally characterized by a growing and expanding economy. Equities tend to rise during these periods. Economies spend on average about 80% of their time in bull markets and 20% in the bearish time periods (Albuquerque, Eichenbaum, Papanikolaou, Rebelo, & Sergio, 2015).

Importantly Albuquerque et al. (2015) found that during bull markets equity returns average 14.2% but tend to fall to -15.5% in bear markets. Reducing risk exposure before the onset of a bear market and relaxing those constraints before the beginning of a bull market would be extremely beneficial to the portfolio returns of market participants.

Several methods have been used to model and forecast large movements in the financial markets. Kole and Dijk (2017) broadly group these methods into parametric and nonparametric categories.

DOI: 10.4018/IJPMAT.2019010105

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

They find that non-parametric models perform well in in-sample identification of market states while parametric models such as Markov switching perform better in the forecasting of the future states of markets. Kritzman, Li, Page, and Rigobon (2011) derived a formula for an economy's absorption ratio which describes the degree of coupling in financial markets. The greater the coupling the greater the current and future potential riskiness of financial markets. Jorion (1995) studied the predictive power of the implied volatility skew for market volatility and while it was found to be biased the volatility skew outperformed times series models. Risso (2008) utilized entropic concepts and found that as financial market informational efficiency decreases the probability of a market crash rises.

The aim of this paper is to introduce a new tool to help practitioners determine when equities are near the start of a bull or bear market or somewhere in-between. Using a new technique Parker (2017) estimated the information processing efficiency of the economy. This new measure was shown to be useful in the prediction of financial downturns. That analysis is extended in the current paper to explain and predict the current and future states of the full cycle of financial markets in a new and intuitive way. This new variable can be used by actors in real-time to get a daily snapshot of the current and probable future states of financial markets and more accurately time the emergence of bull and bear markets.

MATERIALS AND METHODS

Parker (2017) developed an alternative derivation of the yield curve. This derivation is based on Shannon type entropy or information loss as described by Ben-Naim (2017), and combined with Burnashev's formula for the error exponent of communication systems (Burnashev, 1976). An estimate of the information processing efficiency of the economy (R/C) could found using actual yield rates.

Using this alternative derivation, Parker (2017) demonstrated that differing levels of R/C could generate the different regimes of the entropically derived yield curve. These regimes have an equivalent representation in the popular Nelson-Siegel specification of the yield curve (Nelson & Siegel, 1987). The current paper extends the previous research by more closely examining the time evolution R/C during bull and bear markets. As demonstrated empirically R/C rises, reaches a maximum, and then falls in a cyclical pattern. The evolution of this information process provides a new and intuitive explanation of the boom and bust financial cycles as seen from an information theoretic perspective.

The entropic yield curve was derived in Parker (2017) by blending concepts from information theory with traditional financial ideas. Information whether it is communicated through space as traditionally imagined or through time can be lost, corrupted, or otherwise misprocessed. An economy's efficiency in processing and communicating information is related to the economy's interest rates through the following entropic yield curve equation:

$$r_{parker} = B_0 + \frac{\ln(\sqrt{t})}{t} \left(1 - e^{-c_1 \left(1 - \frac{R}{C} \right)} \right) - \frac{\ln(\sigma)}{t} \left(e^{-c_1 \left(1 - \frac{R}{C} \right)} \right) \quad (1)$$

$\frac{R}{C}$ is the implied relative information processing rate, where from Burnashev (1976):

$$p = e^{-c_1 \left(1 - \frac{R}{C} \right)} \quad (2)$$

In the entropic yield curve t represents the term to expiration, B_0 is the asymptotic long rate such as the 30-year bond yield, R is the current economic information set, and C is the economy's

information processing capacity. The remaining terms C_1 and σ are adjustment constants as described in (Burnashev 1976) and (Parker 2017) respectively. Burnashev’s error exponent p determines the relative influence of the true distribution compared to the error distribution in the entropic yield curve as explained in Parker (2017).

The Implied Information Processing Rate (IIPR) or (R/C) can be estimated by matching the entropic yield curve to the observed yields in the markets and then solving for (R/C). Despite their seemingly disparate theoretical origins, the entropic yield curve and the popular Nelson-Siegel model share deep similarities (Nelson & Siegel, 1987; Parker, 2017). The various well-known shapes of the yield curve can be generated simply by varying the information processing ratio R/C as seen in Figure 1.

When $R/C \approx 1$ the normal upward sloping yield curve is generated as seen in the upper left of Figure 1. This structure is indicative of a healthy growing economy. Additionally, this value of R/C indicates that the information processing capabilities of the economy are approximately equal to the amount of information that must be processed.

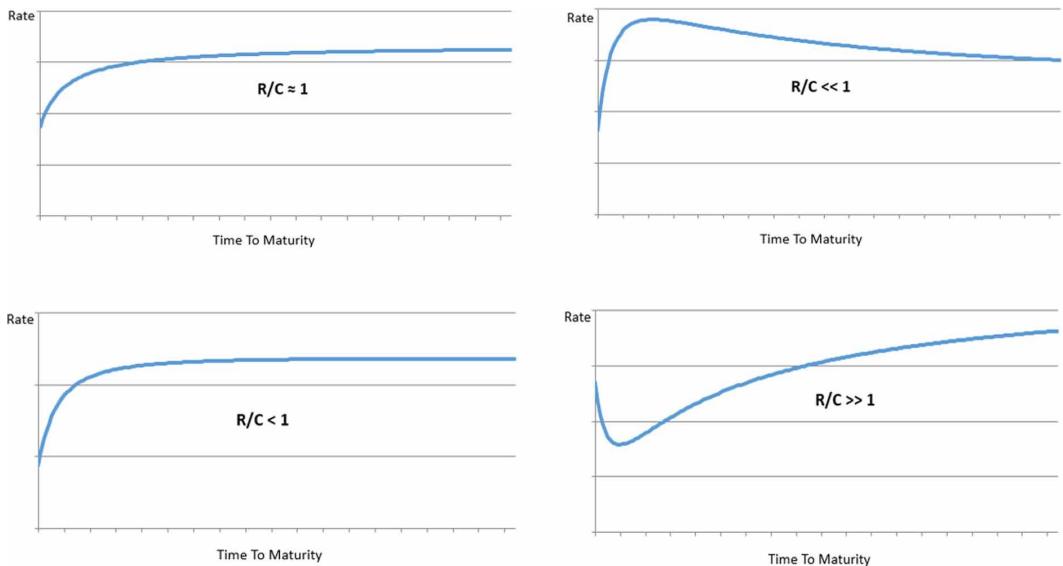
On the bottom left side of Figure 1 a yield generated by a value of R/C slightly less than 1 is produced. In this curve, the yields flatten out slightly compared to the previous case. This structure may be signaling a transition is approaching in the growth rate of the economy. The capacity of the economy to process information is slightly higher than the available information to be processed.

At the top right of Figure 1 is a curve where R/C is much less than 1. The yield curve is inverted with this value of R/C . Short term yields are higher than the longer-term ones. This structure has been highly correlated with economic downturns. The associated R/C indicates that the economy is information starved as processing capacity vastly outstrips the available information.

Finally, at the bottom right the value of R/C is much greater than 1. The yield curve has taken on a “check mark” type configuration. This structure of the yield curve is less well known than the normal, flat, and inverted curves. While the other shapes are discussed extensively in the literature, the “check mark” style curve is rarely mentioned.

However, the “check mark” style yield has a natural explanation from the perspective of the entropic yield curve. When R/C greatly exceeds 1, much more information is arriving than the economy can process. As explained later in the paper this curve structure although rare in actual occurrence is also an important indicator of the state of the economy.

Figure 1. Regimes of the entropic yield curve



When R/C is much greater or much less than 1 the economy is out of balance. Information is either unavailable or greatly exceeds the economy's ability to process that information. This leads to information loss or the growth of entropy as described in (Parker, 2017). This growth of entropy causes the economy to function less efficiently and to be more susceptible to an economic contraction.

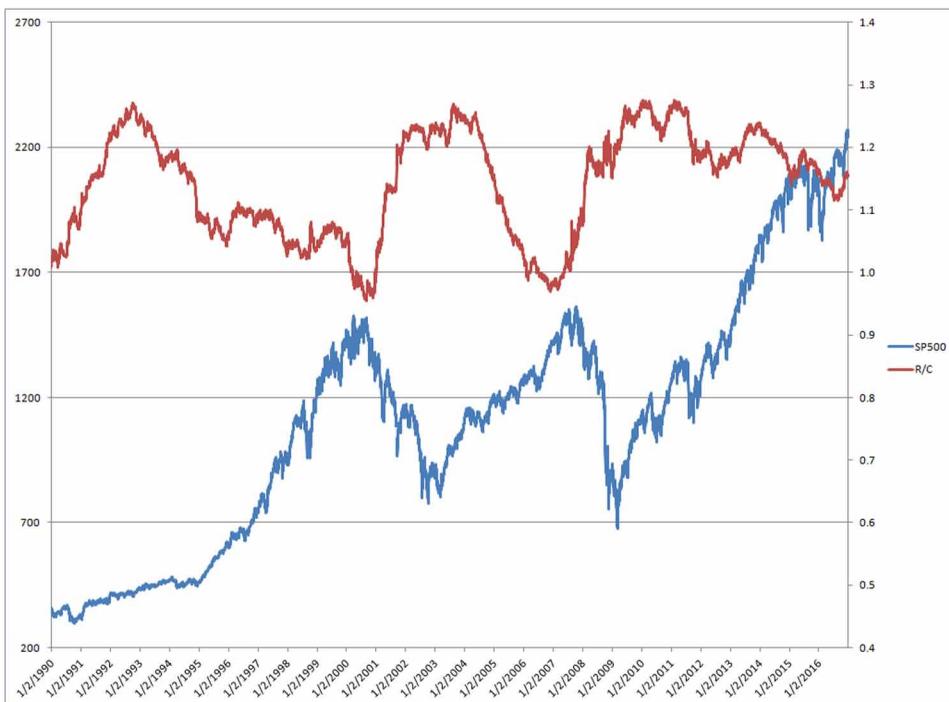
INPUT DATA

Any country with a long range of consistent and reliable data could be used to illustrate the evolution of R/C over a long-time period. The United States was chosen as such a representative country for this long-range study. For this study, freely available daily values of the SP500 from www.yahoo.com finance and yield curve rates from the US Treasury Department website were collected from 1990 through the end of 2016. A future avenue of study would be to see if similar economic patterns exist in other indices and in other countries.

RESULTS

Implementing a simple excel solver based macro as described above and in Parker (2017) the implied information processing rate R/C was estimated from the entropic yield curve. The daily estimated values of R/C over the years 1996-2016 are presented in Figure 2. The utilized excel file with the bond data and macro can be downloaded from the following link listed at the end of the references section. Note the parameterization is modified from that used in Parker (2017). The values of the SP500 are also included in the diagram. Immediately apparent is the counter cycle behavior of R/C compared to the SP500. Information theory provides a new perspective to analyze the relationship between the equity and bond markets as explained in the next section.

Figure 2. R/C versus SP500 from 1990 through 2016



PEAKS AND TROUGHS

In the last two bull markets the peaks are caught between the collapse and recovery phases of the R/C cycle as seen in in Figure 2. During the collapse businesses are squeezing out the last bits of information of the R/C cycle. After the collapse, a new technology is exploited and R/C rapidly rises again. However, it takes companies time to reorganize processing capabilities and capacity to exploit the new information stream. During this reorganization period information is lost, entropy grows and stock prices fall. The bottom of the market occurs near the maximum value of R/C in both of the last two bear markets. At this point the growth of capacity has finally matched that of information growth. Information processing is much more efficient than at the beginning of the R/C recovery, entropy growth slows, and stock prices rise. This continues until the new technology is fully exploited and the cycle begins again.

PEAKS AND TROUGHS AND INFORMATION CYCLE STAGES

There are several obvious components to the R/C cycle. These include phases of recovery, increasing at a decreasing rate, maximization, decline and collapse. Below are quick bullet points detailing each phase:

- Information Cycle Phases:
 - **Increase (at a decreasing rate):** New information is discovered at a declining rate. Additionally the capacity to process that information increases. The growth of R/C slows. As the growth of R/C slows and nears its peak, the decline in the SP500 also slows as the new technology and information begins to be exploited;
 - **Maximization:** Diminishing returns on information discovery and continuing increases in capacity leads R/C to eventually reach a maximum value. As R/C reaches its maximum the SP500 approaches its minimum value as the new technology and information begins to be effectively utilized;
 - **Decline (At an accelerating rate):** The rate of information discovery will continue to be outpaced by the capacity to process that information. As the new technology and information stream is consumed the value of the SP500 increases;
 - **Minimization:** The new technology is exhausted and can no longer reach greater efficiencies in terms of R/C. There is no obvious way to channel investments into the exploited technology to boost R/C. SP500 reaches a maximum as the current technology and information has been fully exploited.

The information processing cycle starts again when a new technology is discovered and implemented.

TIMING OF PEAKS AND TROUGHS

The SP500 peaks occur in or near the zone where the slope of R/C flattens out and R/C rises above 1.00. Note the floor of R/C is approximately 0.96 as seen in Figure 3.

R/C AND SP500 PEAKS

See Figure 3.

R/C AND SP500 TROUGHS

The SP500 troughs occur in or near the zone where the slope of R/C flattens out and R/C rises to its peak. The past three peaks of R/C occur at 1.27 as seen in Figure 4.

Figure 3. SP500 peaks and R/C

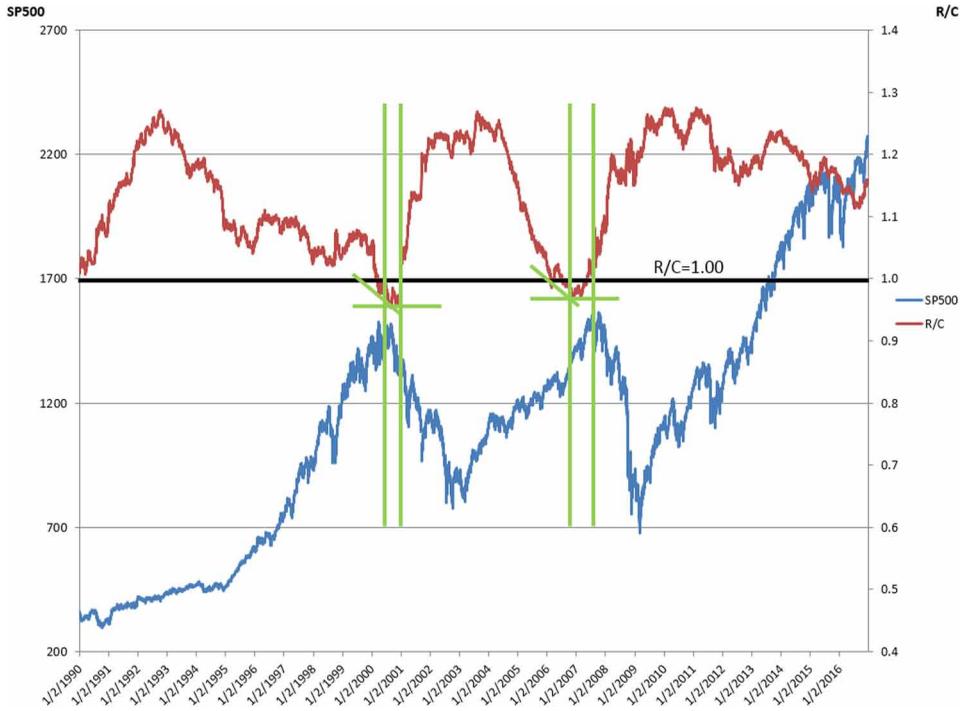
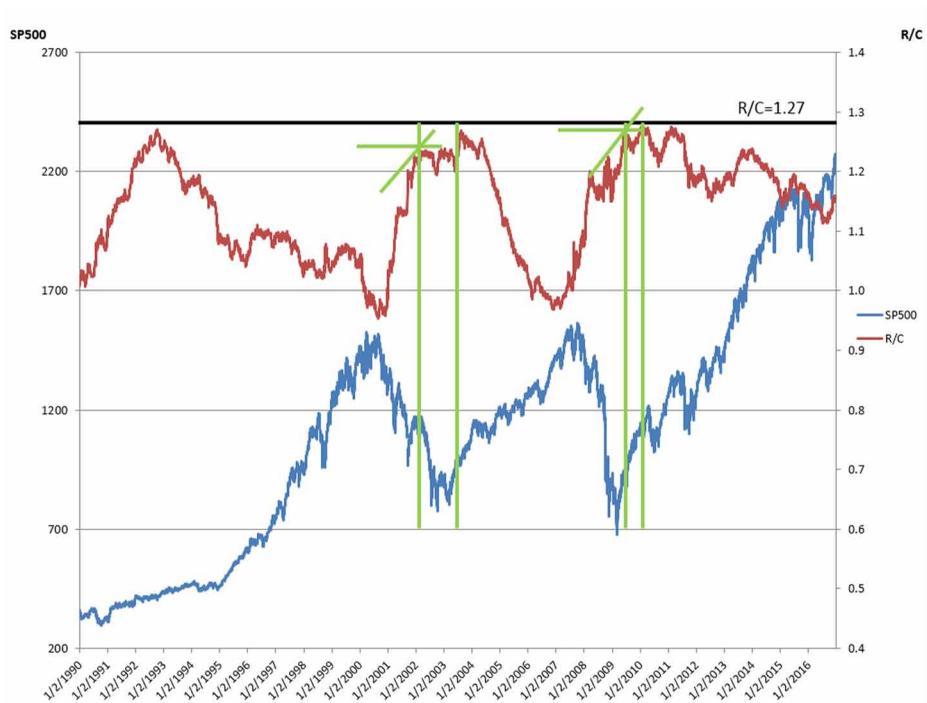


Figure 4. SP500 troughs and R/C



While the R/C gives an indication of whether the economy is near the end of a cycle, it does not give information of how long a cycle might last. Despite this limitation R/C is a new valuable tool for the prediction of market peaks and troughs. R/C may not indicate how long it takes to get to a critical level like $R/C < 1$, but once that level is reached the probability of a near term economic downturn is very high. Conversely when R/C is near 1.27 the probability of an acceleration in the economy is quite high as evidenced by the past two cycles. Additionally, in between these extremes of R/C the probability of a transition between a bear or bull market is relatively low as seen in Figure 4.

THE “CHECK MARK’ YIELD CURVE STRUCTURE AND THE INFORMATION CYCLE

The normal, flat, and inverted orientations of the yield curve are well known. There is however an important yield curve configuration that receives relatively much less attention. This structure is one in which yields first fall and then rise normally as yield increases as seen in Figure 5.

This yield curve structure arises naturally from the information cycle perspective. Additionally, from the entropic perspective this yield curve has an intuitive and important interpretation. The “check mark” structure typically occurs at the start of large downward bear market moves in the SP500. Using daily data from the US treasury department from 1990-2016 we can observe two periods when this configuration appeared. The first occurrence in early 2001 during the so-called dotcom bubble burst as seen in Figure 6. The yield curve data is from 1/16/2001.

The next occurrence of “check mark” configuration occurs at the start of the last bear market. The shape of the yield curve on is depicted in Figure 7 is from 11/19/2007.

Both periods reflect an economy that is in a state of reorganization. Large amounts of information loss and entropy growth occur until the next cycle begins. Once the economy and its resources are effectively reorganized the new cycle can begin. At the beginning of the new cycle information can be efficiently processed, R/C is at its maximum, and the SP500 begins its new ascent.

PARADOX “CHECK MARK” YIELD CURVE AT LOW R/C LEVELS

As stated earlier in the paper high values of R/C can generate this particular structure of the yield curve. However the estimation of R/C in this paper shows empirically when R/C falls to near it’s

Figure 5. “Check mark” yield curve structure

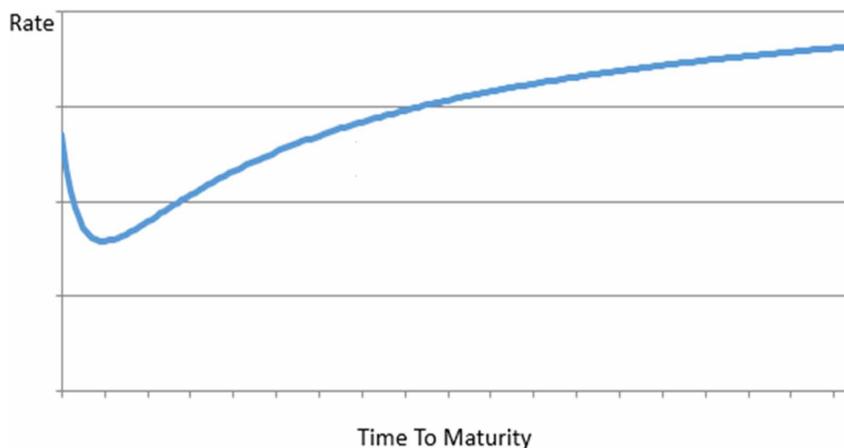


Figure 6. Yield curve shape on 1-16-2001

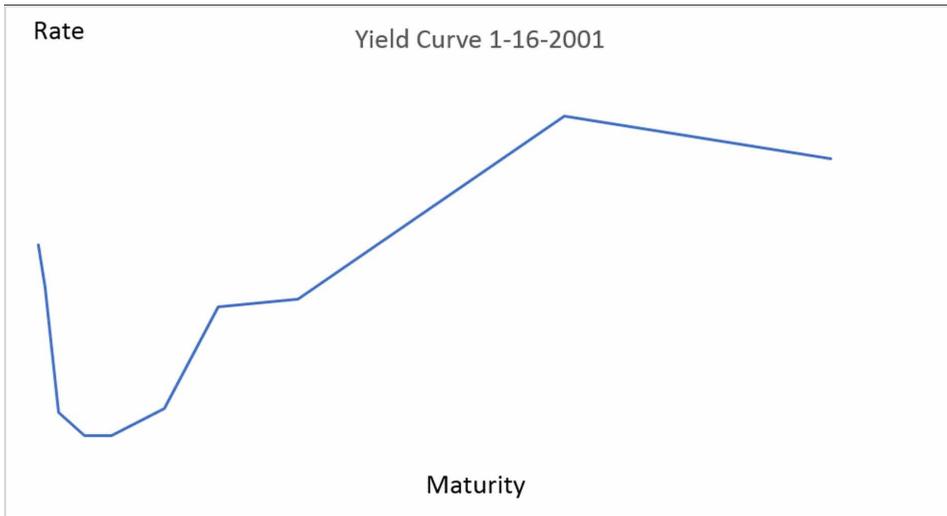
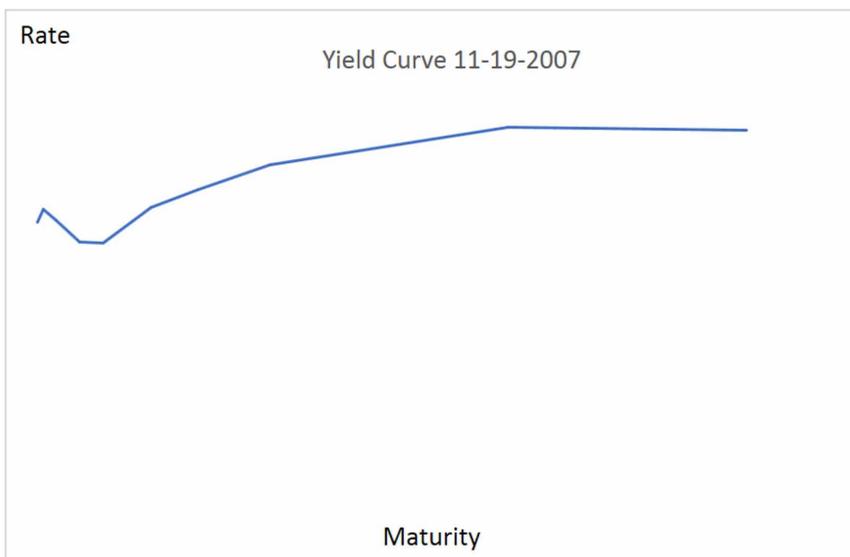


Figure 7. Yield curve shape on 11-19-2007



lowest value when the “check mark” structure of the yield curve appears. The results of the simulation in Figure 5 of Parker (2017) help provide an explanation of this apparent paradox.

Parker (2017) found that the variance of R/C is inversely proportional to its level. In the current paper an additional simulation is performed. This new simulation demonstrates the effect of an increase in the variance of R/C on the yield curve structure. As seen below a random draw of a variable which is $N \sim (0,1)$ is added to R/C. The initial R/C = 1.1:

$$\frac{R}{C}(Total) = \left(\frac{R}{C} + M * N \sim (0,1) \right) \quad (3)$$

$$r(t) = \left(1 - e^{-c_1 \left(1 - \frac{R}{C} (Total) \right)} \right) \frac{\ln(\sqrt{t})}{t} - \left(e^{-c_1 \left(1 - \frac{R}{C} (Total) \right)} \right) \frac{\ln(\sigma)}{t} \quad (4)$$

Initially M is set to zero which results in no random variation in R/C as seen in Figure 8. A normally shaped yield curve with gradually increasing yields with increasing maturities is observed.

Next in Figure 9 M is set to 1. This causes R/C to vary randomly about its mean value 1.1. The resulting yield curve takes on the “check mark” configuration.

Thus the “check mark” configuration can be generated in two ways from the entropic yield curve. The first method was demonstrated in Parker (2017) is with very high values of R/C. The second way that the “check mark” configuration of the yield can be generated is with large values of the variance of R/C. Both of these methods represent an economy in which entropy growth is relatively high compared to a healthy growing economy. If $R/C \gg 1$ then the processing capacity of the economy is greatly exceeded by the information that must be processed. If $VAR(R/C)$ is very large information is also lost as companies cannot efficiently plan and accurately allocate resources to information processing. The rare appearance of this yield configuration is an important sign of the emergence of a bear market. The level and variance of R/C and the emergence of the “check mark” yield curve structure can be combined to help pinpoint the current state of the economy.

CONSISTENCY OF THE R/C MAXIMUM AND MINIMUM VALUES

An obvious question is why are the R/C maximum and minimum values consistently near 1.27 and 0.96 respectively? Is there any significance to these actual values or are they just the quantities that could have taken any maximum or minimum values at the corresponding peaks and troughs? Of course, the estimation method used will affect their calculation as seen in Parker (2017). Below I

Figure 8. Simulated yield curve for the values R/C=1.1, M=0

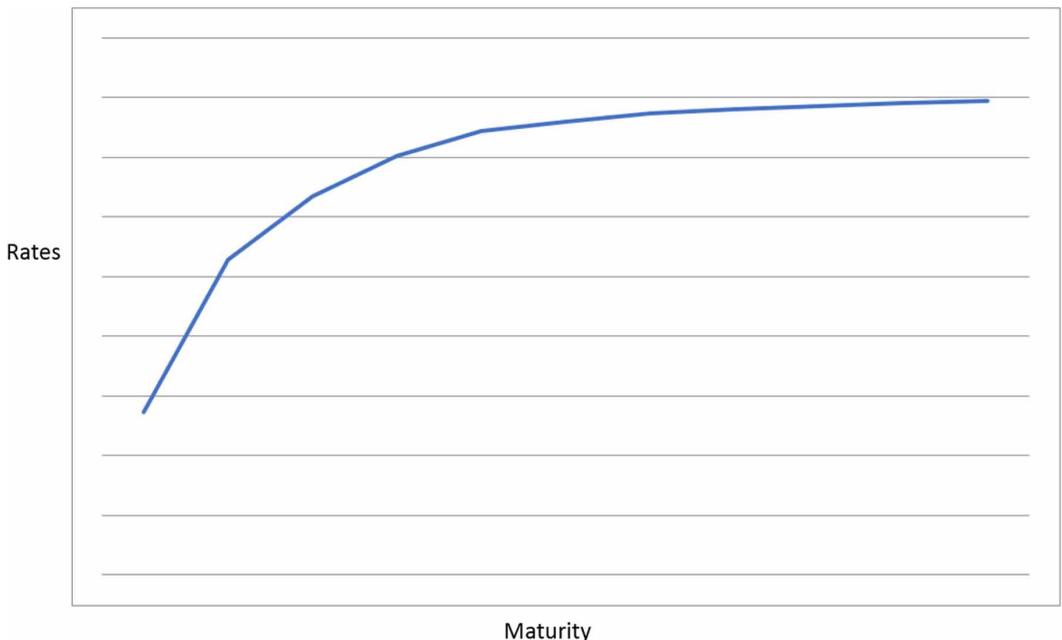
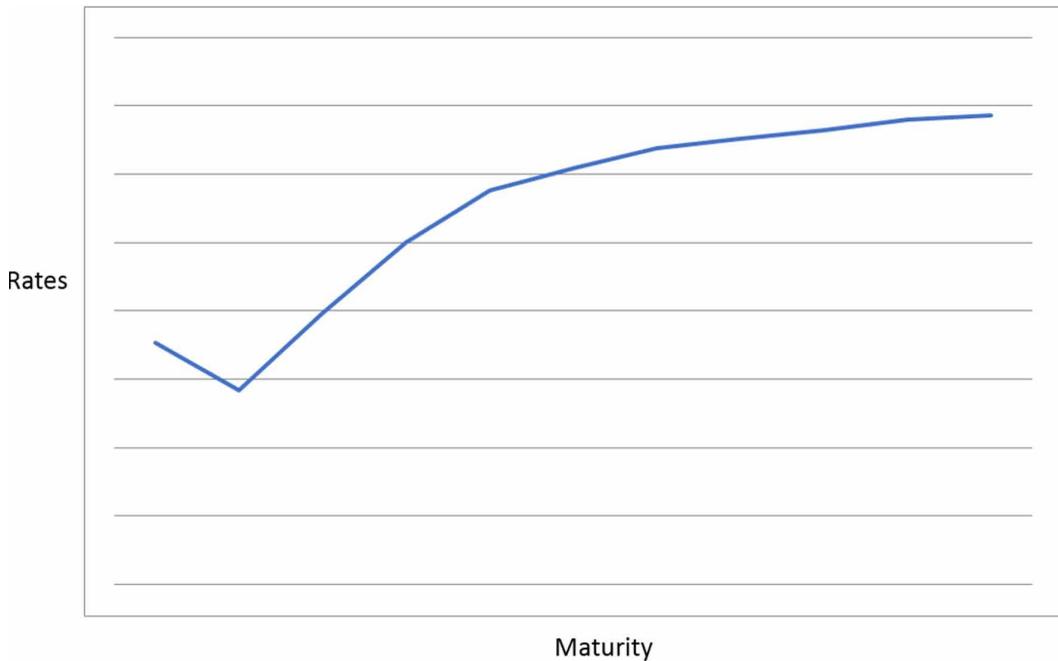


Figure 9. Simulated yield curve for the values R/C=1.1, M=1



offer a geometric interpretation that is consistent with the description of the R/C cycle that intuitively generates these numbers. Finally, an interpretation of their meaning is examined.

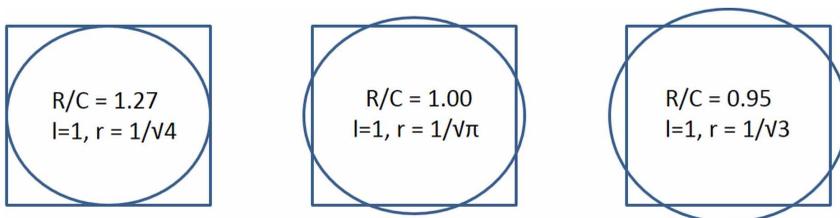
The information available to be processed by the economy will be represented by the area of a unit square. The processing capacity will be visualized as the area of an associated circle. The values seen in the R/C curve can be generated by the ratio of the area of the unit square to the area of the associated circles as seen in Figure 10.

R/C maximum occurs at 1.27:

$$1.27 \approx \text{area of unit square divided by of inscribing circle} = 1/2$$

$$\frac{1}{\pi \frac{1}{2}} = \frac{4}{\pi} = 1.27323\dots$$

Figure 10. Geometric interpretation of R/C



This shows that that 78% Of the area if the square is covered by the circle.

Where $1/R/C = 1/1.27 = 0.7853$. Note that this relationship and value is the maximum circular packing density in a square plane.

R/C minimum occurs at 0.95:

$0.95 \approx \text{area of unit square divided by area of a circle of slightly greater area where } r = 1/\sqrt{3}$

$$\frac{1}{\pi \frac{1}{\sqrt{3}^2}} = \frac{3}{\pi} = 0.95492\dots$$

where $1/R/C = 1/0.95 = 1.04719$. Note that this relationship and value exceeds the circular packing density in a square plane. This means that the condition of non-overlapping circles would be violated in the plane.

At $R/C=0.95$ the nearly circumscribing circle the economy has exhausted the information set from the previous cycle using the existing technology as the processing circle begins to envelope the available information unit square. Any further increase in the processing circle will not significantly increase the information processed. No further efficiencies from information processing growth using current technologies can be attained. At the same time, the SP500 peaks and begins its decline.

A new technology with its associated information set will be utilized to start R/C's climb back up towards its maximum. This is represented by the growth of the information square. The square grows until R/C peaks at 1.27 where finally the processing circle is fully circumscribed by the square. At this point the growth of the processing circle begins to exceed the information square's growth. Simultaneously the SP500 begins its recovery. R/C falls and the SP500 rises until R/C again reaches its minimum at 0.95 and the cycle begins again.

Note that at $R/C = 1.00$ the area of the circle equals that of the square. Presumably this is the level where the processing of information equals the capacity. Why does capacity expansion not halt this level?

The observed R/C minimum of 0.95 can be thought of as a slight overshoot of the level $R/C=1.00$. This can conceptually be explained by the fact that no economic participant knows the economy is at $R/C = 1.00$. The only way to find out is to further increase processing capacity and empirically observe that information processing has not actually increased. Note the overshoot past $R/C = 1.00$ to 0.95 is quite small when looked at from the perspective of the radii of the representative circles. The circles' radii are $r \approx 0.564$ and $r \approx 0.577$ for $R/C = 1.00$ and $R/C = 0.95$ respectively.

Additionally, the greater degree of accuracy in the geometric interpretation at R/C maximum versus the minimum may be also be due to technological competition at cycle start versus a more narrow technology focus at its end. At the beginning of the cycle optimally the best of many competing technologies will be selected. This competition may result in the relatively precise value of 1.27 seen in R/C where the processing circle is fully inscribed. However, it may be more difficult to determine when that previously selected successful technology has run its course. In addition, there is typically some inertia to the adoption of new technologies.

DIGITAL SIGNAL PROCESSING AND ALIASING INTREPRETATION

Some readers familiar with digital signal processing or multidimensional sampling may also note that $R/C = 1.27$ represents the nonoverlapping square lattice circle packing limit. Circles beyond the radius of $r = 1/2$ will begin to overlap. The overlapping circles represent the loss of signal reconstruction

or aliasing. One could argue that the increase in processed information with larger radii will at some point be offset by loss of information with the emergence of aliasing and signal degradation.

THE US ECONOMY'S CURENT STATE OF R/C

With most of the year 2017 complete as of the submission of this draft the question naturally arises, where are we now in the information processing cycle? To answer this question the dataset as described in the Materials and Methods section was expanded to include daily values of the U.S. yield curve rates and the SP500 from January 2017 to date. The values of R/C were estimated from the entropic yield curve in the same fashion as previously described.

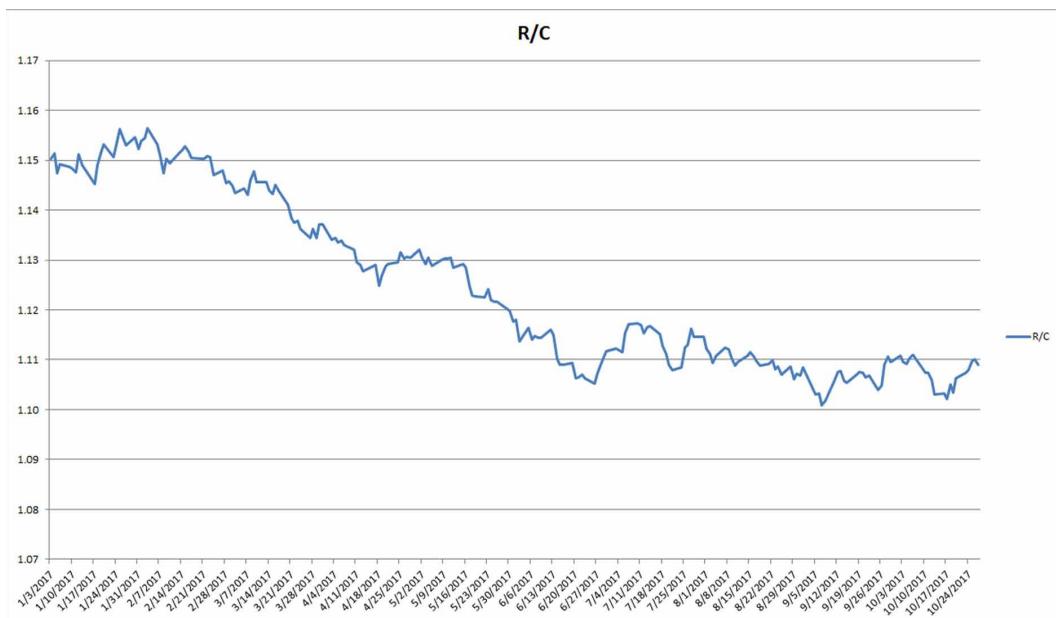
As seen in Figure 11 the US economy started the year at $R/C = 1.15$. Since this time there has been a gradual drop in the level of R/C to about 1.11 at the time of this paper submission. If the cycle is measured from 1.27 through 0.95 this would mean that the US economy is approximately halfway through the current cycle.

R/C may not of course be useful for noneconomic and sudden exogenous shocks such as natural disasters or wars. This is true of any economic indicator. However, it appears that the variable would have been quite informative over the past two cycles. The author maintains daily updates of R/C at the following link to demonstrate in real time the utility of R/C <http://www.relativechannelcapacity.com/>.

CONCLUSION

The entropic yield curve was derived in Parker (2017) by blending concepts from information theory with traditional financial ideas. The implied information processing ratio R/C which can be estimated from the market yield curve rates and the entropic yield curve provides a new means to understand the dynamics of financial markets. This information theoretic perspective can also be used to improve the forecasting of market peaks and troughs. This new variable can be used by economic actors in

Figure 11. The level of R/C 2017 year to date through 10/26/2017



real-time to provide a daily snapshot of the current and probable future states of financial markets. Future research may attempt to measure the relationships between R/C, the entropic yield curve, and other markets, economic variables, and indicators.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Excel file with the macro and data used to estimate R/C is available at: https://drive.google.com/file/d/0B4SScA4-ziPSTkJFVXAAtWU8wU28/view?usp=sharing_eip&ts=59bbfd1c The Treasury Yield Curve data is also freely available at: <https://www.treasury.gov/resource-center/data-chart-center/Pages/index.aspx>. The views expressed in this article are the author's own and do not reflect those of New York Life Insurance Company.

REFERENCES

- Albuquerque, R., Eichenbaum, M., Papanikolaou, D., & Rebelo, S. (2015). Long-run bulls and bears. *Journal of Monetary Economics*, 76(S), 21-36.
- Ben-Naim, A. (2017). Entropy, Shannon's measure of information and Boltzmann's h-theorem. *Entropy* 19(48).
- Burnashev, M. (1976). Data transmission over a discrete channel with feedback, random transmission time. *Problems of Information Transmission*, 12(4), 250–265.
- Jorion, P. (1995). Predicting volatility in the foreign exchange market. *The Journal of Finance*, 50(2), 507–528. doi:10.1111/j.1540-6261.1995.tb04793.x
- Kole, E., & van Dijk, D. (2017). How to identify and forecast bull and bear markets. *Journal of Applied Econometrics*, 32(1), 120–139. doi:10.1002/jae.2511
- Kritzman, M., Li, Y., Page, S., & Rigobon, R. (2011). Principal components as a measure of systemic risk. *Journal of Portfolio Management*, 37(4), 112–126. doi:10.3905/jpm.2011.37.4.112
- Nelson, C., & Siegel, A. (1987). Parsimonious modeling of yield curves. *The Journal of Business*, 60(4), 473–489. doi:10.1086/296409
- Parker, E. (2017). The entropic linkage between equity and bond market dynamics. *Entropy (Basel, Switzerland)*, 19(6), 292. doi:10.3390/e19060292
- Risso, W. (2008). The informational efficiency and the financial crashes. *Research in International Business and Finance*, 22(3), 396–408. doi:10.1016/j.ribaf.2008.02.005
- Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Chicago, IL: The University of Illinois Press.

Edgar Parker graduated from Georgia State University with a BS in Economics, and the University of Maryland University College with an MS in Computer Systems Management. He works for the Information Technology Department of New York Life Insurance.