Global Journal of Business Research
Vol. 10, No. 1, 2016, pp. 65-69
ISSN: 1931-0277 (print)
ISSN: 2157-0191 (online)

IBFR
www.theIBFR.com

SPLICED CORRELATION: THEORY DEVELOPMENT

Jeffry Haber, Iona College

ABSTRACT

Correlation is a common metric used in portfolio management. It describes the relative movement of two streams of data, allowing inference of how one will behave given the movement of the other. Often a significant correlation relationship (whether it be uncorrelated, positively correlated or negatively correlated) in the long-term is not replicated in shorter term periods. Worse, often the short-term correlation is contradictory to the long-term. Utilizing three sets of data, where the streams of two are interchanged to form one stream at varying points of time could allow the long-term correlation to be also replicated in the short-term. There remain various obstacles to overcome, such as scaling, determination of inflection points and the selection of the data streams. Those are left to be solved in future papers - this paper puts forth the theoretical justification for the concept of spliced correlation.

JEL: G10, G11, G12, G23, G170

KEYWORDS: Correlation, Spliced Correlation, Investing, Portfolio Management

INTRODUCTION

Consider this thought experiment – you are walking down the street and you have a feeling that you are being followed. You constantly look back to see if there is a consistent presence of any of the people who are behind you. You cannot detect one. Now think of it from the pursuer's perspective (assuming your paranoia was well-founded). Since you would be able to detect a stable presence behind you, the pursuers alternate who is following you. Every time you look back you see someone different than the previous time. A trained spy might be able to catalog all the people that have recently been behind you and detect the tail, but probably not an ordinary citizen.

Correlation is not that much different from the thought experiment. Consider the following streams of figures (noted "A," "B" and "C"):

A	B	C
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	5	1
7	4	2
8	3	3
9	2	4
10	1	5

The correlation between A and B is 0.0; perfectly uncorrelated. It would seem that streams A and B behave randomly with respect to each other. But break the ten item string into two strings of five items each and a different story emerges – the correlation of the first five items is +1.0, perfectly positively correlated, and the correlation of the last five items is -1.0, perfectly negatively correlated. The two strings of perfect correlation combine to produce a single string of perfect non-correlation.

Now take the correlation between A and C – it is 0.5. If you take the correlation of the first five items you get 1.0, and if you take the correlation of the second five items you also get 1.0. Each of the shorter periods is perfectly correlated, taken as a whole the correlation is halfway between 0.0 and 1.0.

Essentially what is being demonstrated is that the long-term correlation (ten observations) is valid as long as you hold both streams for the full duration. If you enter or leave an investment anywhere within the ten periods your correlation experience might be quite different, and if you held one of the items for half the period your experience might be significantly different than what you would expect, given the correlation for the full period.

To calculate correlation you would traditionally take items two at a time over some period. Correlation is useful when it works, and not useful when it doesn't. A significant correlation figure over longer periods may not be present when taken in shorter intervals, and vice versa. Correlation is not causality, but an indication of what we expect to happen to something by virtue of observation of something else. One of the limitations of correlation has been the need to find two streams of data (in an investment context, generally monthly returns) that are correlated in both the short-term and the long-term. Long-term correlation is desirable because of the reliance that can be placed on it over a sustained duration. But this is only true if the correlation exists when taken in shorter periods as well. It would not be unusual for two streams of data that are highly correlated when calculated over a 15 year period (for example) to have non-correlation over meaningful, but shorter periods. Likewise for non-correlated streams of data over the long-term – it would not be unusual for these streams to be highly correlated in shorter, but significant periods.

So, what is necessary then, is for there to be significant correlation (positive, negative or non-correlation) over the long-term and the short-term. This has typically been hard (if not impossible) to find – compromises in correlation have been accepted, meaning that shorter periods of aberrant behavior have to be tolerated for what is considered the greater good – delivering the long-term correlation. This paper develops a methodology to provide both long- and short-term correlation with no compromise.

LITERATURE REVIEW

The importance of correlation in investing is long held and well-documented. Harry Markowitz's seminal work on portfolio selection placed correlation front and center in the portfolio manager's toolkit.

> "Soon after Harry Markowitz published his landmark 1952 article on portfolio selection, the correlation coefficient assumed vital significance as a measure of diversification and an input to portfolio construction" (Kinlaw and Turkington 2014).

Not only is correlation important when considering equity and debt markets, correlation is also prominent in discussion of any asset that can be included in a portfolio, such as commodities.

> "Indeed, most fund managers have started advising their customers to devote a share of their portfolios to commodity-related products as part of long-term diversification strategy" (Lombardi and Ravazzolo 2013).

> "… from the point of view of investors, the historical performance of collateralized investments in commodity futures suggests that they are an attractive asset class to diversify traditional portfolios of stocks and bonds." (Gorton and Rouwenhorst 2005).

Often, correlation is calculated over longer terms, but its significance and value is not described by the past, but what will happen in the future. A long-term correlation figure often does not hold in the short-term. It would not be unusual to find three consecutive years of highly correlated returns within a longer term (fifteen years) span of uncorrelated returns (this also holds for the various permutations – shorter, but meaningful, periods of highly positive correlation embedded within a long period of either negatively correlated or non-correlated, and so on).

One study (Haber and Braunstein 2008) took 180 observations (representing 15 years of random, monthly returns) and found that all observations had a correlation of .20 or below, representing uncorrelated returns. This result is relevant for the full-duration of the 15 year holding period. The last 12 observations (the final year of the 15 year holding period) provided a starkly different correlation result – 61% of the observations has a correlation of more than .20, 39% had a correlation of .30 or more, 21% had a correlation of .40 or more and 10% had a correlation of .50 or more. The key takeaway of another study was that the nearer and shorter the term, the greater the likelihood that the long-term correlation will fail to hold (Haber and Braunstein 2009).

The variation between long-term and short-term correlation was noted (Preis, Kenett, et al, 2012):

> "… in complex systems, such as financial systems, correlations are not constant but instead vary in time."

The importance of correlation in portfolio development is well-represented in the literature, as is its short-comings. What is missing is a solution to the problem of when a long-term correlation (say a calculation that indicated non-correlation) is contradicted by a short-term correlation (such as the correlation becoming highly positive).

THEORY DEVELOPMENT

Finding correlated streams usually involves selecting a group of variables and testing them two at a time. Spliced correlation starts out the same way, but considers whether various streams can be combined to produce a long-term correlation as well as a similar short-term correlation. Consider a data stream, visually denoted as "X". This stream will be held constant. Two additional streams, represented by "Y" and "Z" are plotted above and below "X." This is represented in Figure 1.

Figure 1: Depiction of When Splicing Would Be Valuable

```
            Y                                   Y
          Y                                   Y
Y Y Y Y Y           Y Y Y Y Y
X X X X X X X X X X X X X X X X X X X X
          Z Z Z Z Z                   Z Z Z Z Z
                    Z
                      Z
```

This figure depicts how one data stream (X) correlates with another stream (Y) for a period of time, and then no longer maintains that correlation. Once Y starts to trail away another stream (Z) correlates with X. This alternating pattern continues.

Stream "Y" is correlated with stream "X" for a period of time, then enters into a period of where it no longer correlates. When "Y" no longer correlates with "X," stream "Z" starts to correlate with "X." Then when "Z" no longer correlates with "X" "Y" resumes its correlation. This alternation occurs on a predictable pattern.

Splicing the two streams together ("Y" and "Z") produces a combined stream of data that can be used in correlation calculations with "X," as depicted in Figure 2.

Figure 2: Depiction of What the Spliced Sequence Would Look Like

X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Y	Y	Y	Y	Y	Z	Z	Z	Z	Z	Y	Y	Y	Y	Y	Z	Z	Z	Z	Z

This figure illustrates what the data streams would like when streams Y and Z are pieced together. Streams Y and Z form a single stream that is used to calculate the correlation with X for any duration.

Correlation calculations are based on streams of numbers, no matter whether they are untouched, adjusted or an artifact of creation. The math will work just as well. What is not trivial is identifying the streams that will be spliced and the inflection points when the splicing takes place.

DEMONSTRATION OF THE THEORY

Investment management is the act of acquiring assets to build a portfolio in order to accomplish a particular goal. The metrics (such as correlation) used to select these assets are typically calculated over the long-term. Previous research has documented that the long-term correlation cannot be relied on in the short-term. This can have significant ramifications for risk management, achieving return goals and diversification. Spliced correlation is a potential methodology that provides a mechanism for the substitution of assets into the portfolio to maintain the long-term correlation over short periods.

To test the theory a narrative needs to be developed about how two data streams relate to each other, and when the circumstances under which they won't. Then an additional stream needs to be identified that will work when the first doesn't, and then the development of a system for understanding when the substitutions should take place.

Because the calculation of correlation is highly sensitive to sign changes, implementing a substituted set of data might introduce sign changes, simply because the figures are based on a different scale (think of using stock quotes for stream "Y" and US Treasury bond yields for stream "Z" – because they are denominated differently, any change from Y to Z will introduce a sizable decrease in data (and likewise there will be a corresponding large increase in going to Z to Y). The relative scaling of the two streams is a consideration.

How to determine the inflection point is another complication. Whether substitutions will be based on an ad-hoc model, where based on some algorithm the system will determine when to substitute or whether the substitution will take place on a scheduled basis (like every 5^{th} day, for instance) will need to be worked out through testing.

CONCLUDING COMMENTS

Correlation is most valuable when it works in the short-term as well as the long-term. Historically, long-term, significant correlations go through extended short-term periods where the correlation is different from the long-term, in a non-trivial manner. Non-correlation can become extreme correlation; highly negative correlation can become highly positive correlation. Since correlation is based on two streams of data, there is an opportunity to combine two (or more) streams into one, for the purposes of calculating correlation. The combination of two (or more) streams into one could allow the long-term correlation to be replicated in any short-term period, given that the additional stream(s) is/are correlated similarly during short periods when the first stream is not, and the model knows when the appropriate inflection points occur so that substitution can be made. There would be further research required to address the implementation issues not discussed in this paper – such as calculating the inflection point when substitution should occur and determining the stream(s) that will be the substitute.

REFERENCES

Gorton, Gary, and K. Geert Rouwenhorst, "Facts and Fantasies about Commodity Futures," February 28, 2005, http://fic.wharton.upenn.edu/fic/papers/06/0607.pdf

Haber, Jeffry "Fooled by Correlation: How Blind Acceptance of Correlation Dogma Destroys Diversification," Journal of Business Diversity, Volume 12(3), 2012, pp 22-25

Haber, Jeffry and Andrew Braunstein, "Examining the Role of Short-Term Correlation in Portfolio Diversification," Graziado Business Report, 2009, Volume 12, Issue 3

Haber, Jeffry and Andrew Braunstein, "Correlation of Uncorrelated Asset Classes," The Journal of International Business and Economy, Volume 9, Issue 2, December 2008, pp 1-12

Kinlaw, Will and David Turkington, "Correlation Surprise," Journal of Asset Management, 2014, pp 385-399

Lombardi, Marco J. and Francesco Ravazzolo, "On the correlation between commodity and equity returns: implications for portfolio allocation," July 11, 2013, http://www.bis.org/publ/work420.pdf

Preis, Tobias, Dror Y. Kenett, H. Eugene Stanley, Dirk Helbing and Eshel Ben-Jacob, "Quantifying the Behavior of Stock Correlations Under Market Stress," October 18, 2012, http://www.nature.com/articles/srep00752

BIOGRAPHY

Jeffry Haber is Controller of the Commonwealth Fund, a private foundation that seeks to move the United States towards a high performance health care system. The Commonwealth Fund has $750 million under investment.

He is also Professor of Accounting at Iona College, teaching undergraduate and graduate courses. He holds a PhD from Rensselaer Polytechnic Institute and MS and BS Degrees from Syracuse University. He holds CPA, CGMA and CAM licenses.

He is the author of the books What if Everything We Knew About Investing Was Wrong? and Accounting Demystified and numerous journal articles such as "Fooled by Correlation: How Blind Acceptance of Correlation Dogma Destroys Diversification," "Do the Unrealized Gains in Alternative Investments Ever Become Realized?" "The Accuracy of the Rates of Return Reported by Fund Managers: A Test in the United States," "Under-Reporting of Returns by Alternative Asset Funds," "Does the Market Go to 1?" "Correlation of Uncorrelated Asset Classes," "Examining the Role of Short-Term Correlation in Portfolio Diversification," "Of "Real" Investment Returns," and "The Failure of Equity Managers to Beat Their Benchmark: Lord is it I (or is it the Benchmark)?"