A New Quantitative Approach to Macro Investing: Using Employment Data to Generate Excess Returns

Matthew Kelly

Apatheia Research Pty Ltd[1]

19 September 2017

Abstract

Macroeconomic data is often noisy, contradictory and lagging. These limitations render the data difficult to integrate into a robust quantitative investment strategy that generates excess returns. This paper outlines a new approach to macro investing that removes these inherent limitations in macroeconomic indicators through an Economic Scoring Model that gives context to the data by utilising non-parametric statistical measures. The resultant model output exhibits a very strong correlation of 0.97 to the S&P 500 stock index as well as superior coefficient of determination when compared to the five employment data series used as model inputs. A backtest covering 50 years of data is performed comparing the Economic Scoring Model Employment Index against an S&P 500 benchmark. The Model demonstrates superior performance against the benchmark with median returns on average 37% higher combined with a higher percentage of winning trades.

[1] http://www.macrowonk.com

Electronic copy available at: https://ssrn.com/abstract=3043677

Introduction

There are countless investment strategies that attempt to achieve returns above a defined benchmark (commonly known as alpha). The most prominent strategies such as those based on factors and momentum attract the most assets and academic attention. However, the performance of these strategies is very much dependent on underlying economic conditions. The S&P 500 experienced drawdowns exceeding 50% during the US recession from 2007 - 2009. Accurately quantifying economic conditions can reduce these recession related drawdowns, accordingly improving returns. The problem is not an absence of data. The issue is that the data is often noisy, contradictory and lagging. Taking employment data as an example, there are multiple indicators yet they rarely move in tandem. Non-farm payrolls may show strong jobs growth in the previous month while the unemployment rate has unexpectedly increased and Average Hourly Earnings suggests wage stagnation.

By adopting a first principles approach and examining the foundations upon which returns are achieved, this paper outlines a quantitative Economic Scoring Model that generates excess returns when used as a filter on passive investments or can be integrated into any active investment strategy. Using the Economic Scoring Model highlights the path of least resistance and improves the probability of generating alpha.[2]

First Principles

Most established investment strategies are based on hypotheses about a company's fundamentals or its price. For example, a value investor bases her investment decisions on the premise that stocks that trade below their intrinsic value tend to outperform the market (on average). Conversely, momentum strategies are based on the premise that markets that are trending up or down will continue to trend beyond the point where one can reasonably argue that they've priced in all available information. There

[2] This paper covers the US equity markets only, however the same principles apply to other markets and asset classes.

Electronic copy available at: https://ssrn.com/abstract=3043677

is another way to think about market returns and it involves a type of reasoning that has been around since the time of Aristotle. By adopting a first principles approach, we strip away all assumptions. Instead, we start at the foundations and build up a hypothesis from there.

A stock market is a diverse collection of companies, yet they all have one thing in common: each company must (eventually) generate revenue and profit to prosper. The higher the revenue and profit, the more valuable a company becomes. As more and more companies become more valuable, the stock market index increases. Company profit is dependent on people buying its goods or services. For people to spend money on goods and services, they must have an income in the first place. They must be employed. There is a circular relationship here in that companies need revenue to hire and revenue is driven by employment. Reasoning up from there, we can then gauge a company's health and consequentially the health of the market as a whole.

Does this mean that employment data is the secret ingredient to alpha? Unfortunately it's not that straight forward. The situation is complicated by a multitude of factors but in general this paper argues that favourable economic conditions create favourable environments for bull markets (and vice versa). Not such an outlandish argument. So there is a relationship between economics and market returns however we need to carefully consider how we accurately quantify economic conditions as there are a number of pitfalls to be aware of. For instance, traditional measures of economic growth bear little correlation to equity returns[3]. If we consider the individual components of most economic growth indicators, it's no surprise that such indicators are poor predictors because they're lagging measures. This whitepaper will outline a model to quantify economic conditions and demonstrate how the model outperforms traditional macroeconomic data in terms of predictive power as well as serving as the foundation of an investment strategy that generates excess returns.

[3] Jay R. Ritter, "Economic growth and equity returns," *Pacific Basin Finance Journal* 13, (2005): 489-503, https://site.warrington.ufl.edu/ritter/files/2015/04/Economic-growth-and-equity-returns-2005.pdf

Data

The data used in the proceeding sections covers the period from January 1967 to December 2016. Aside from being a nice round 50 years, there is no history prior to 1967 for two of the data series used, so this is the maximum sample size we have available. S&P 500 data is sourced from Robert Shiller's Yale website[4] and is dividend adjusted. The data has been converted to monthly resolution by taking the monthly average of the daily closing prices. In addition to the stock market data, there are five employment indicators used:

1. Monthly Total Nonfarm Payrolls

   The number of jobs added or lost in the economy. Nonfarm Payrolls (NFP) incorporates approximately 80% of the workforce who contribute to GDP. Rising NFP is positive.

2. Weekly Initial Claims

   The number of people filing first time claims for unemployment benefits. Declining Initial claims is positive.

3. Weekly Continuing Claims

   As above, except instead of representing the number of initial claims, this release covers the number of people currently receiving benefits. Importantly, it excludes people who have exhausted their benefits. Declining Continuing Claims is positive.

4. Monthly Civilian Labor Force Participation Rate

[4] "Online Data Robert Shiller," Yale University, accessed May 1, 2017, http://www.econ.yale.edu/~shiller/data.htm.

The percentage of working-age people who are currently employed or are actively looking for work. The participation rate represents the active portion of an economy's workforce. Rising Participation Rate is positive.

5. Monthly Unemployment Rate

   A percentage calculated by taking the number of unemployed people and dividing by the number of people in the workforce. Unemployed people are defined as those who are able work and have sought work within the past four weeks. People with temporary or part-time jobs and those who complete at least 15 hours of unpaid domestic work are also considered employed. A declining Unemployment Rate is positive.

The economic data is sourced from the Federal Reserve Bank of St. Louis (FRED) and is seasonally adjusted. It has to be noted that the dataset may contain look-ahead bias in that it contains historical amendments made to the data after it was officially released. Analysis of these amendments shows that they tend to be small and net out over time. It is highly unlikely they would make a statistically significant impact on the final results presented below.

Methodology

We are going to give context to the economic data movements by applying some simple statistical measures to derive a score for each data point. We will then aggregate these scores into an index and show how it improves on the predictive power of the individual economic data series as well as providing predictive power for the S&P 500. When choosing statistical measures, we need to be careful to avoid those that assume normally distributed data.

The following histograms plot the respective economic data series against a normally distributed series containing random numbers with the same mean and standard deviation. All data is from January 1967 to December 2016.

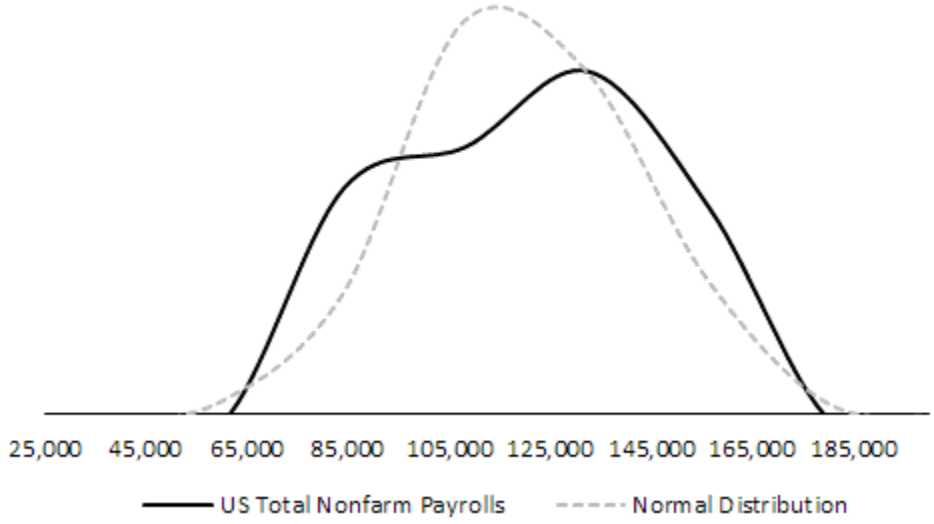


Figure 1: US Total Nonfarm Payrolls.[5] NFP data is negatively skewed (long left tail) and also displays negative kurtosis (flatter peak). The shape suggests the presence of fundamental shifts in the employment landscape over the sample period.

[5] U.S. Bureau of Labor Statistics, All Employees: Total Nonfarm Payrolls [PAYEMS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PAYEMS, May 14, 2017.

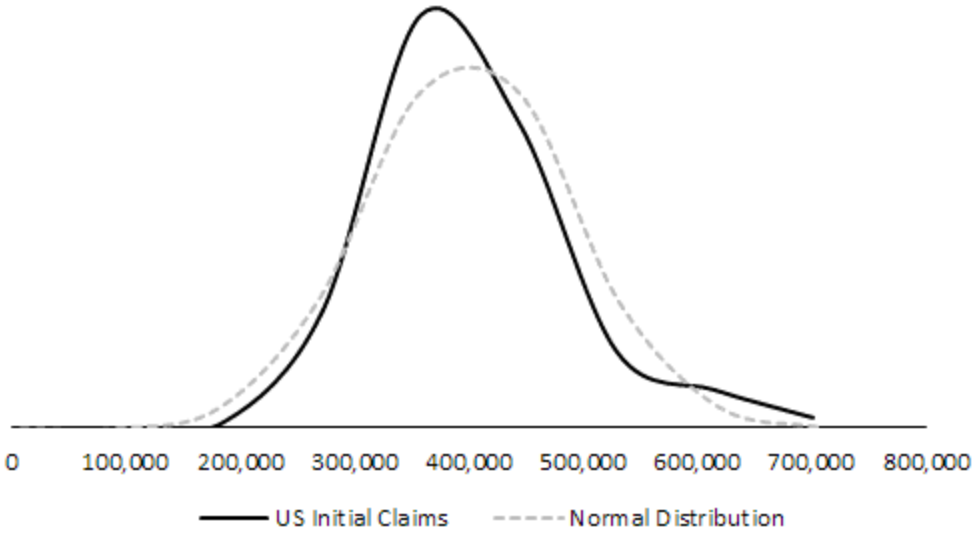
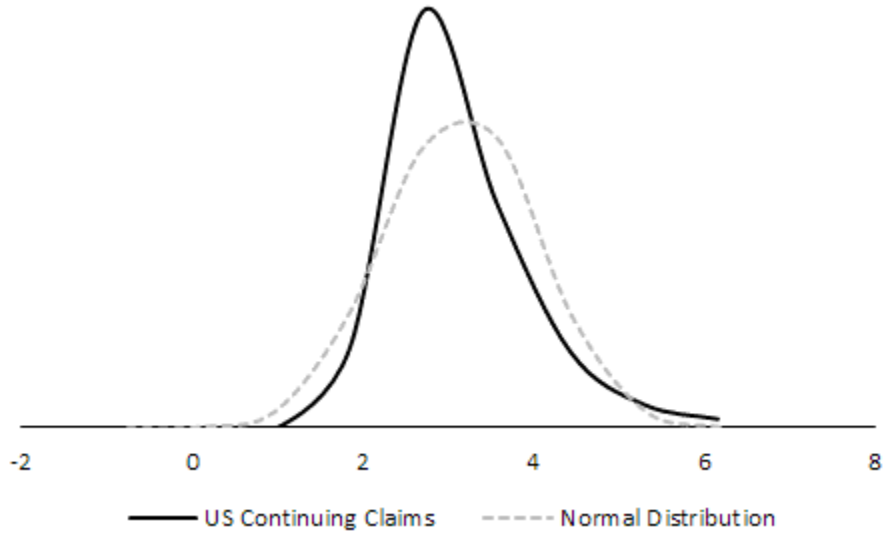


Figure 2: US Initial Claims.[6] This time the data has positive kurtosis (data clustering around the mean) and slight positive skew suggesting periods of high initial claims are more prevalent than periods of low initial claims.



[6] U.S. Employment and Training Administration, Initial Claims [ICSA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/ICSA, May 14, 2017.

Figure 3: US Continuing Claims.[7] This series displays similar traits to Initial Claims in that there is positive kurtosis and a positive skew.

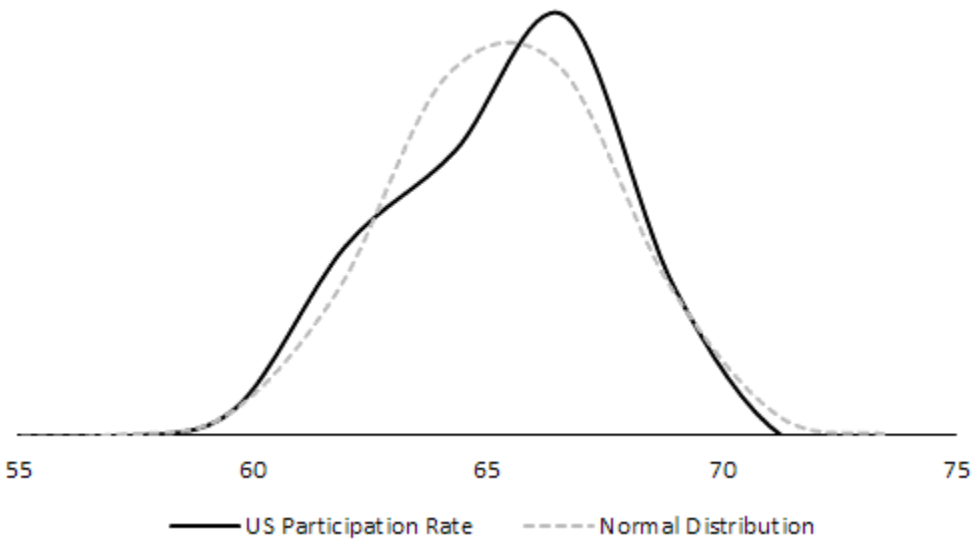


Figure 4: US Participation Rate.[8] The histogram is negatively skewed, suggesting a fundamental shift in the workforce as the number of people actively seeking work declines as the baby boomer generation retires.

[7] U.S. Employment and Training Administration, Continued Claims (Insured Unemployment) [CCSA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CCSA, May 14, 2017.
[8] U.S. Bureau of Labor Statistics, Civilian Labor Force Participation Rate [CIVPART], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CIVPART, May 14, 2017.

Figure 5: US Unemployment Rate.[9] This distribution has a fat right tail, pointing to periods of higher unemployment than a normal distribution (the dotted line) would suggest.

These five histograms have one thing in common; they do not confirm to a normal distribution. Using statistical measures such as mean and standard deviation will not accurately represent variation in the data. Instead, the Economic Scoring Model utilises non-parametric statistical measures as these are not reliant on any kind of distribution shape. For each economic data series, we calculate the following statistics over 30 periods of history:

1. 10th Percentile (P10)

   10% of the sample is found below this value.

2. 25th Percentile (P25)

   25% of the sample is found below this value.

3. 50th Percentile (P50)

[9] U.S. Bureau of Labor Statistics, Civilian Unemployment Rate [UNRATE], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/UNRATE, May 14, 2017.

50% of the sample is found below this value. Also known as median, presents a more robust measure of central tendency than mean as it is less likely to be skewed by large or small numbers.

4. 75th Percentile (P75)

   75% of the sample is found below this value.

5. 90th Percentile (P90)

   90% of the sample is found below this value.

6. Interquartile Range (IQR)

   The difference between P75 and P25.

7. Low Outlier

   P25 - (IQR × 1.5). Data below this level is regarded as an extreme low deviation.

8. High Outlier

   P75 + (IQR × 1.5). Data above this level is regarded as an extreme high deviation.

Nonfarm Payrolls, Participation Rate and Unemployment Rate data is monthly therefore the statistics cover 30 months of data. Initial and Continuing Claims data is weekly so the statistics covers 7 months of data. The data sample is a rolling 30 periods so that each time a new release is made, it is incorporated into the sample and the data point from 30 periods ago is dropped from the sample.

Once we have these 8 measures, we can calculate a score for each release that provides an accurate measure of the impact of the release as it relates to the historical sample. The scores have a range from -1 to +1 and are calculated according to the following ranges:

Range	Score
Current > Previous and <= P50	0.2
Current > Previous and >= P50 and <= P75	0.4
Current > Previous and > P75 and <= P90	0.6
Current > Previous and > P90 and <= High Outlier	0.8
Current > Previous and > High Outlier	1.0
Current < Previous and >= P50	-0.2
Current < Previous and <= P50 and >= P25	-0.4
Current < Previous and < P25 and >= P10	-0.6
Current < Previous and < P10 and >= Low Outlier	-0.8
Current < Previous and < Low Outlier	-1.0
Current = Previous	0

Account needs to be taken of the direction of the economic indicator and its meaning for the economy. An increase is positive for the economy when looking at Nonfarm Payrolls and the Participation Rate. With Initial Claims, Continuing Claims and the Unemployment Rate, increases are negative for the economy. For these indicators, the score is reversed (a positive score using the above ranges is converted into a negative score and vice versa).

Now that we have a score for each data release, they are aggregated into monthly totals before being compiled into an Employment Index. The index is at monthly resolution which is adequate for the purposes of this paper. In practice, an index that is updated immediately after the data release is preferable as it enables more timely investment decisions.

The Employment Index starts at 100 in January 1967. For subsequent months, index values are simply the previous index value plus the sum of all 5 economic indicator scores. Recall that two of the economic data series are at weekly resolution. For these, we sum the weekly scores to arrive at a monthly score. The resultant Employment Index is below:

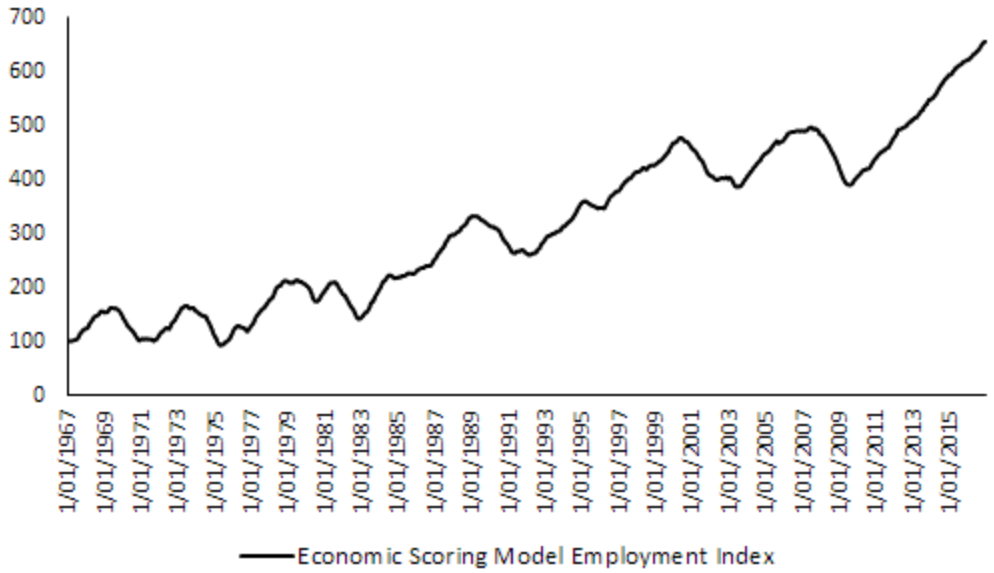


Figure 6: Economic Scoring Model Employment Index

In the following sections we will delve into the predictive power of this index and explore some practical applications.

Results

For the Employment Index to be of any use, it needs to improve upon any predictive power that exists in the five separate employment data series. For this purpose, we will first examine the correlations between the Employment Index and the S&P 500 before analysing the coefficient of determination (R^2) to determine how well the Employment Index (the independent variable) predicts changes in the S&P 500 (the dependent variable).

Correlations between January 1967 and December 2016:

	Employment Index	Nonfarm Payrolls	Initial Claims	Continuing Claims	Participation Rate	Unemployment Rate
S&P 500	0.97	0.88	-0.18	0.10	0.30	-0.18

Correlation results show a very strong relationship between the Employment Index and the S&P 500 with a correlation coefficient of 0.97. This is better than all of the individual employment data series.

The following charts detail the R^2 for the same data.

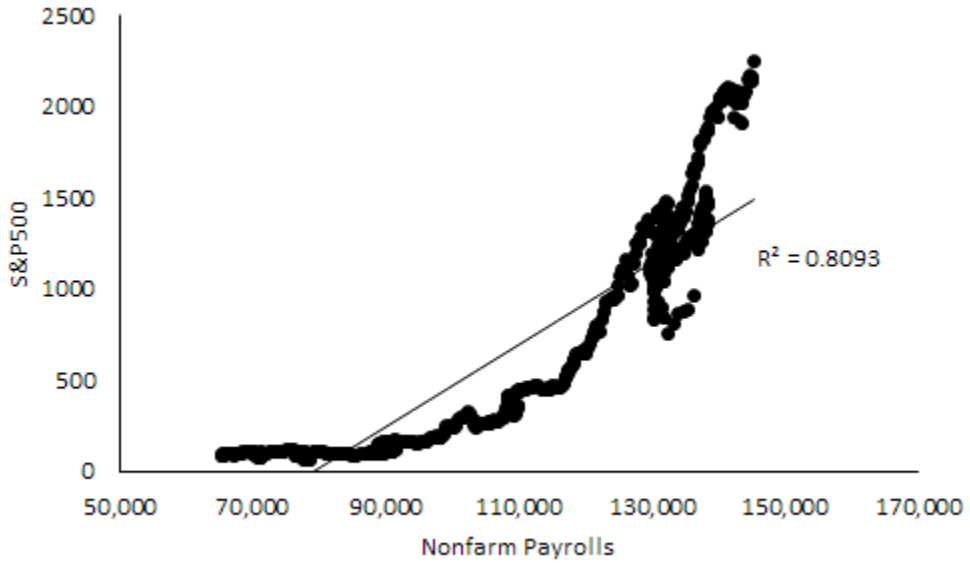


Figure 7: Nonfarm Payrolls / S&P500. With an R^2 of 0.809, NFP has a good linear relationship with the S&P500. There are however outliers in the data and several regime shifts, making it less than ideal to use as a predictive tool.

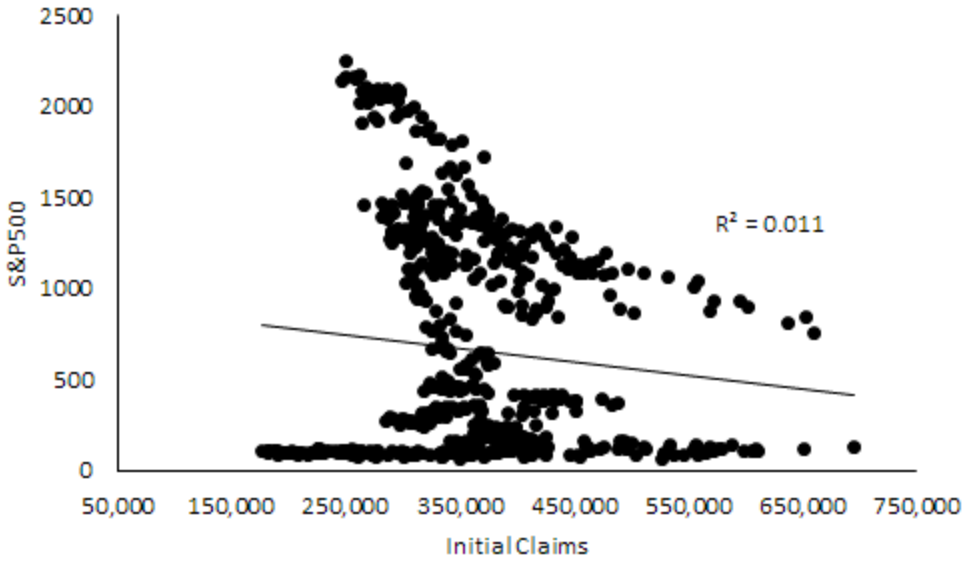


Figure 8: Initial Claims / S&P500. Initial Claims has a poor fit to the S&P 500 with an R^2 of 0.011. The scatter plot shows several clusters of data where initial claims actually increases yet the S&P500 is range bound.

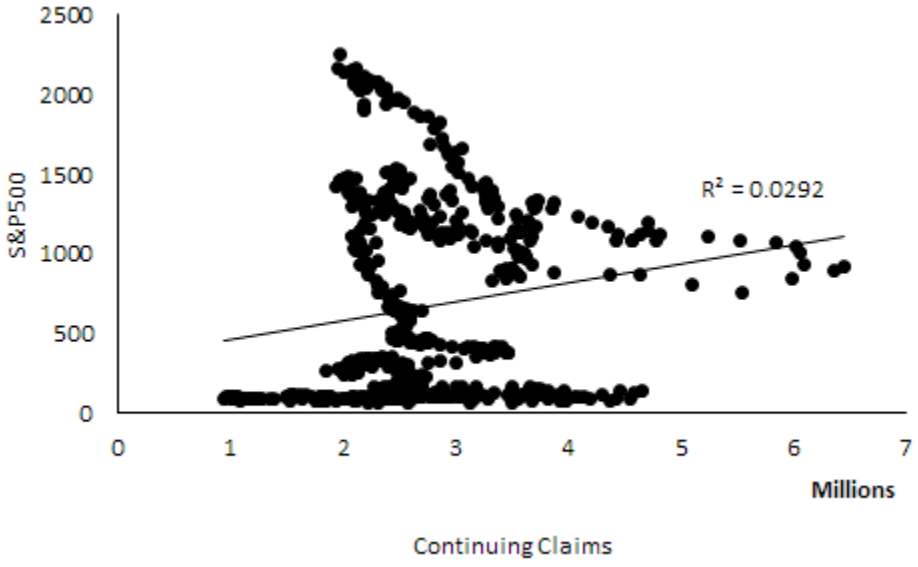


Figure 9: Continuing Claims / S&P500. Similar patterns to Initial Claims, which is to be expected. Again, there are clusters of data where continuing claims increases yet the S&P500 is range bound.

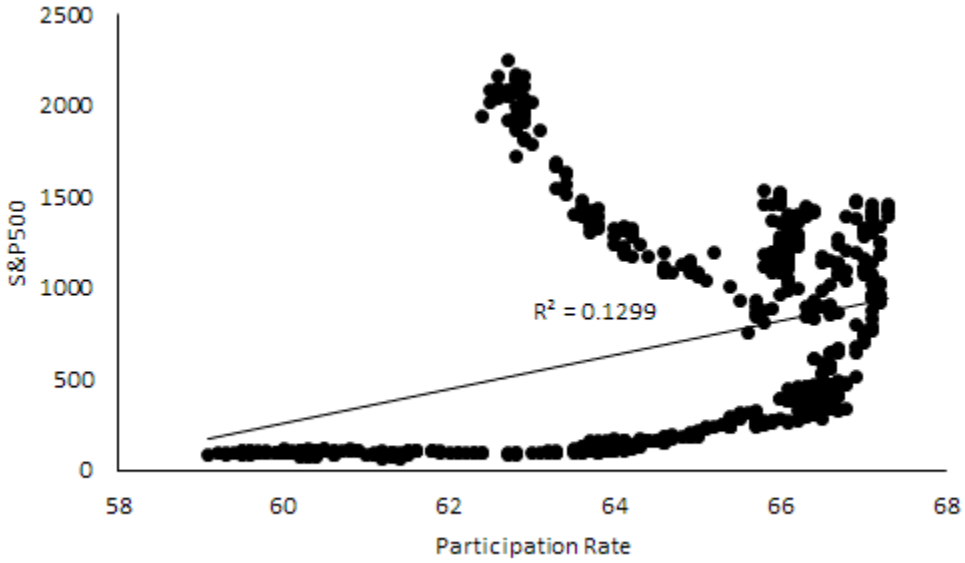


Figure 10: Participation Rate / S&P500. Two distinct regimes are discernable here. Upper data points display a relationship where the S&P500 declines with rising Participation Rates which is contrary to expectations.

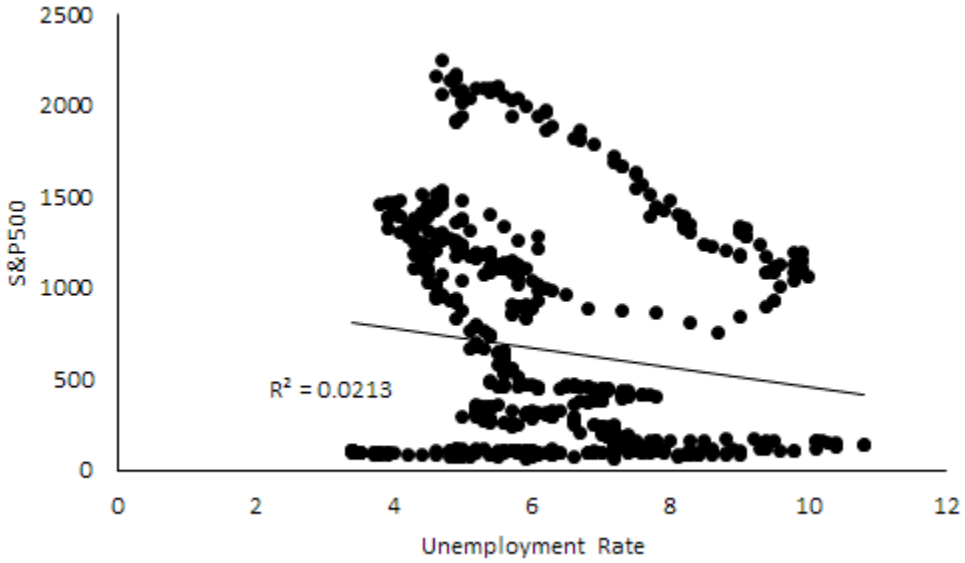


Figure 11: Unemployment Rate / S&P500. Displays several regimes where the S&P500 declines with rising unemployment rate (as expected). Also displays patterns where the unemployment rates rises yet S&P500 remains range bound, contrary to expectations.

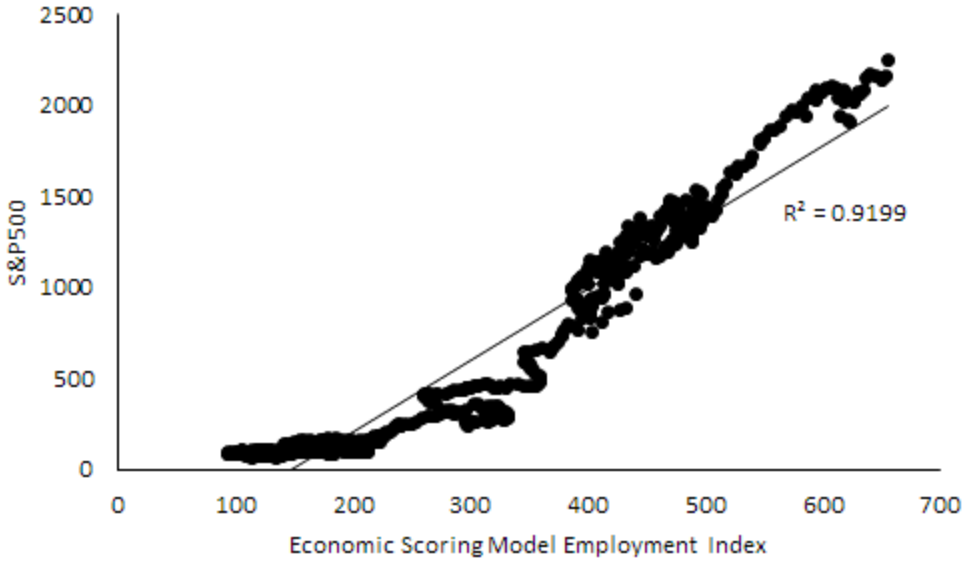


Figure 12: Employment Index / S&P500. With an R^2 of 0.9199, the Employment Index has the tightest range of data points and the best R^2, suggesting that the index has predictive power in relation to the S&P 500.

The statistical tests clearly demonstrate that the Employment Index exhibits a strong fit to the S&P 500 when compared to the five economic data series. Statistical tests do not always transfer into real world applications however. In the following section we will conduct a backtest using a simple breakout system.

Practical Applications

Instead of the Employment Index value, the direction of the Employment Index is far more important. Any index is dependent upon the chosen start date, which is artificial and skews the resultant values. For example, a start date of January 2007 produces an Employment Index that goes to zero 30 months later:

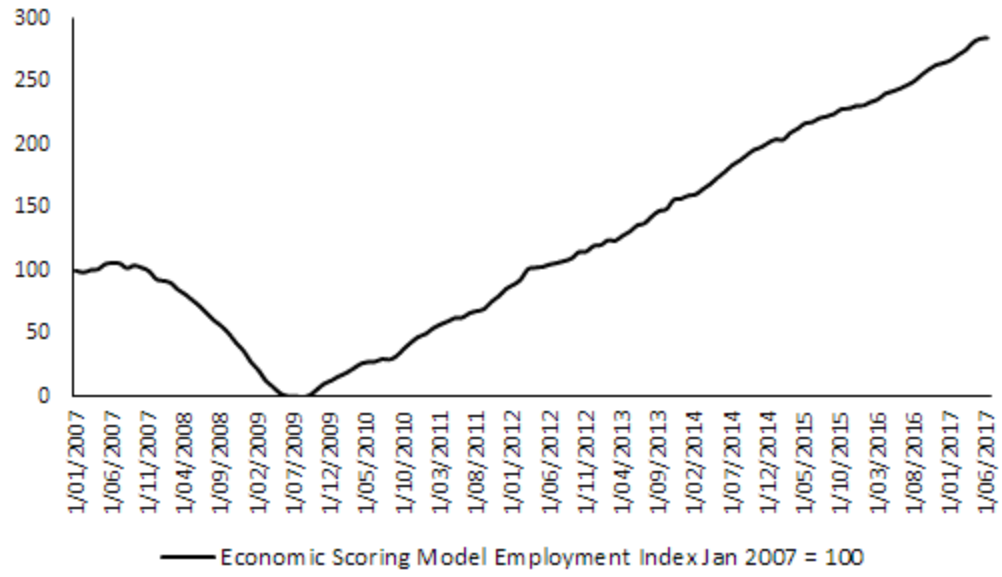


Figure 13: Economic Scoring Model Index, Jan 2007 = 100

There are several ways to quantify direction. We will test a 10 month breakout, which would be equivalent to approximately 200 days in a model based on daily data.

Entries are determined by a new 10 month high in the Employment Index. Entry price is the following month's average price to prevent look-back bias in the results. The benchmark follows the same entry rule except instead of breakouts in the Employment Index, entry is determined by breakouts in the S&P 500. Results are analysed over a 12 month period following the entry to determine the permanence of the signal.

Note that prices are monthly averages for the S&P 500. Differences between the hypothetical execution price and the actual price achieved in a real world application are bound to happen, however these should average out over time (less than favourable execution prices should average out instances of favourable prices). This approach is intended to give a good indication of the overall expectancy of the method. It is not intended to deliver a complete trading system. Likewise, we are using the full data set here and any robust backtest should integrate out-of-sample testing.

The following table and chart details the median return between the Employment Index breakout and the S&P 500 breakout.

Months	S&P 500 Median Return	Employment Index Median Return	Difference (%)
1	0.50%	1.04%	52%
2	1.04%	1.88%	45%
3	1.46%	2.93%	50%
4	2.29%	3.52%	35%
5	3.42%	4.55%	25%
6	4.05%	6.20%	35%
7	4.00%	6.09%	34%
8	4.21%	6.68%	37%
9	5.34%	6.58%	19%
10	6.04%	10.14%	40%
11	6.71%	11.10%	40%
12	7.91%	11.67%	32%

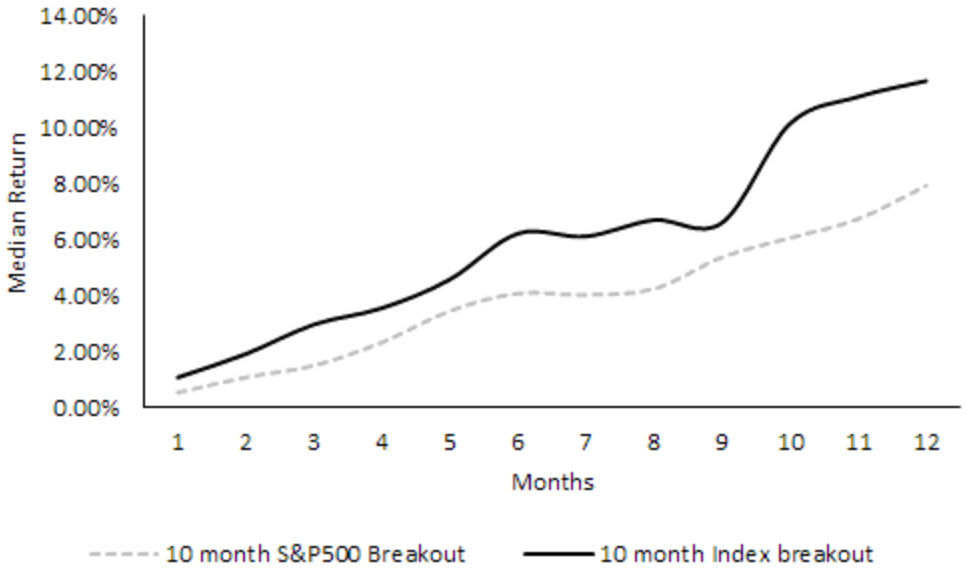


Figure 14: Median Return of Employment Index Breakout vs S&P 500 Breakout

Median return for the Employment Index breakout is on average 37% higher than a 10 month price breakout in the S&P 500. The trend in returns over the 12 months further suggests that the predictive value of the index breakout extends far beyond the initial signal.

The following chart examines the percentage of winners for the Employment Index breakout and S&P 500 breakout:

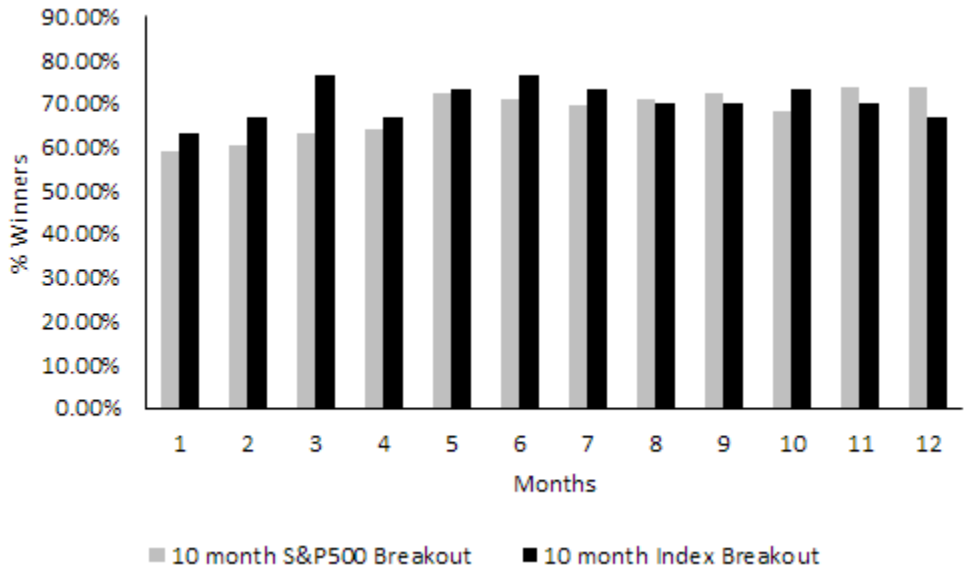


Figure 15: Percentage Winners - Employment Index breakout vs S&P 500 breakout

Here again the Employment Index breakout demonstrates superior performance over the S&P 500 breakout with 8 out of 12 months displaying a higher percentage of winning trades. The average percentage of winners over the full 12 months is approximately 70%, demonstrating that the Employment Index breakout exhibits a higher probability of winning trades over what would be expected according to chance.

Conclusion

Adopting a first principles approach allows us to question the source of market returns, yielding a new model that generates alpha. Raw macroeconomic data is often noisy, contradictory and non-contextual. The Economic Scoring Model outlined in this paper gives context to economic data movements. The resultant Employment Index exhibits a very strong correlation of 0.97 to the S&P 500 stock index as well as superior R^2 when compared to the five individual employment data series. When performing a simple

10 month breakout backtest, the Economic Scoring Model Employment Index again demonstrated superior performance against the S&P 500 benchmark with median returns on average 37% higher than benchmark and a higher percentage of winning trades. The Economic Scoring Model provides a powerful way to evaluate economic data, strip it of noise and output a reliable signal with strong predictive power.