**Characterization of Factor Performance by Economic Regime**

**Applied Finance Project (MFE18)**

**Advisor**

*Professor Ronald Kahn*

**Students**

*Kelvin Cheng, James Cho, Michael Tang, Zibo Zhao*

March 2018

# 

**I. Abstract**

Risk premium across factors and asset classes have varying dynamics in different business cycles and economic conditions. Based on the expectation of the nonlinearity between economic cyclicality and factor performance, we break down the workflow into a two-stage model in hope of gaining new insight. The first stage involves a economic regime prediction model. As our baseline, we use publicly available leading economic indicators to classify forward economic states into four regimes. The second stage of the model assesses factor performance in each regime using standard Fama-French factors in the US equity market. Utilizing historical factor performance in each regime, we construct a factor-weighting strategy that attempts to outperform a benchmark. We further expand our model employing statistical methods for economic state identification; we construct a transition-probability matrix for the economic states and subsequently utilize the probabilities in factor weighting to compare performance with our baseline model.

FuckTRUM

**II. Introduction**

Factor models are widely used both for institutional and individual stock investors. For most equity investors, stock selection, or, in the factor model framework, factor selection, contribute most of the return in their portfolio. Some studies states that the factor returns are generated by linear processes therefore not affected by other issues. However, many empirical studies point out that stock returns follow complicated patterns with multiple ‘regimes’. i.e., it’s widely believed that factor performance is related to the overall macroeconomics environment. For example, momentum factor usually performs well, however, in hash period it could suffer from huge drawdown (i.e, 50% within 2 months).

In this paper, we adopt a two-stage model to explore the relationship. The first part is a economic status identification and prediction model. And the second part investigates factor performance under different market regimes. Different weighting schemes for factor signals are tried and compared.

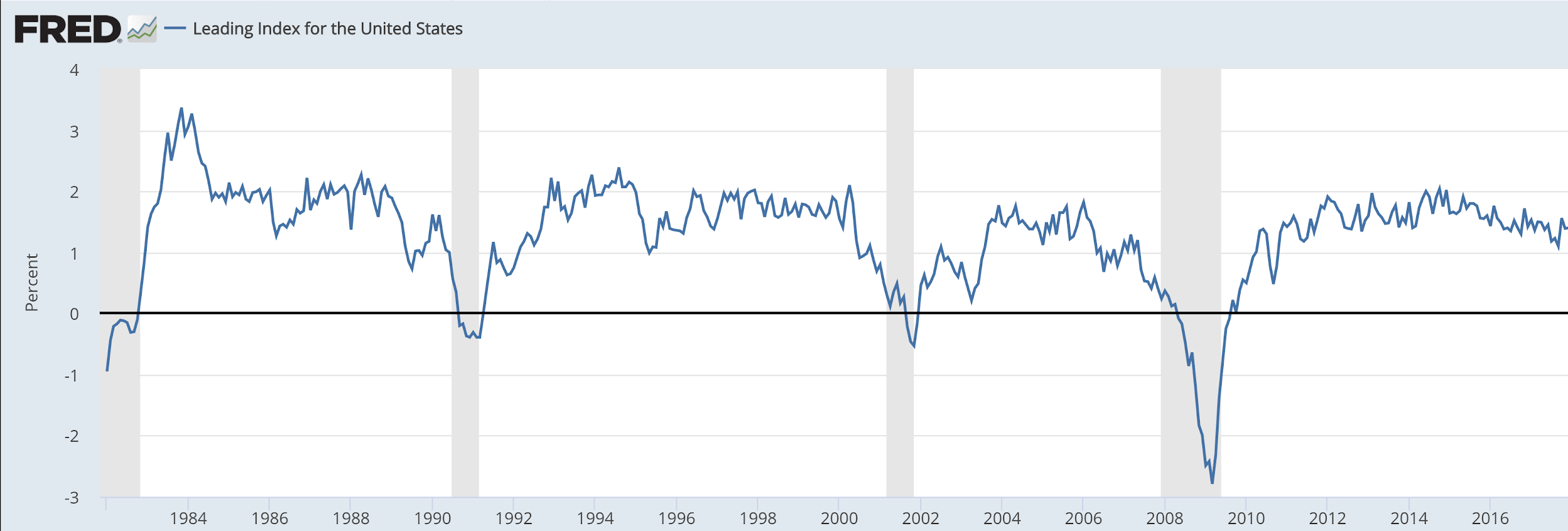
After the two stage process, we also employ a one-stage machine learning model that directly links the economic variable into factor performance. This attempt is to achieve higher precision in the expense of economic intuition.

**III. Methodology**

**Economic State Identification model**

We utilize Philadelphia FED’s Leading Index (LEI for short) for the United States, which is a combination of its underlying state level leading indices. The index is released in a monthly basis and tracks the six-month growth rate of the state's coincident index.

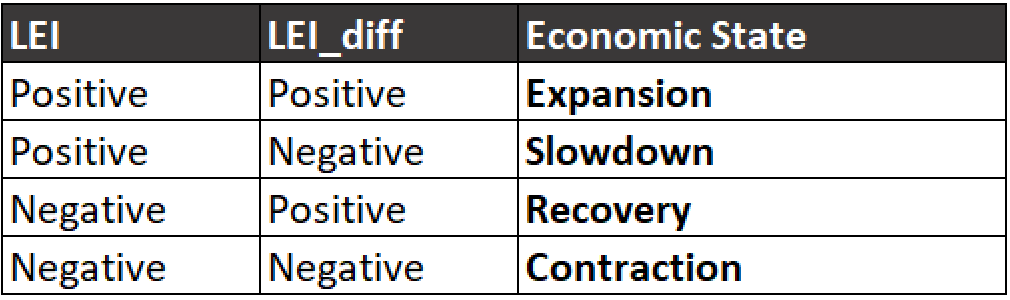
*Leading Index for the United States - Philadelphia FED*



The leading index includes variables that lead the economy: unemployment insurance claims, housing permits, delivery times from the ISM manufacturing survey, and the 10year/3month Treasury spread. According to Philadelphia FED, the LEI tracks the 6 month forward Coincident Index of the US economy.

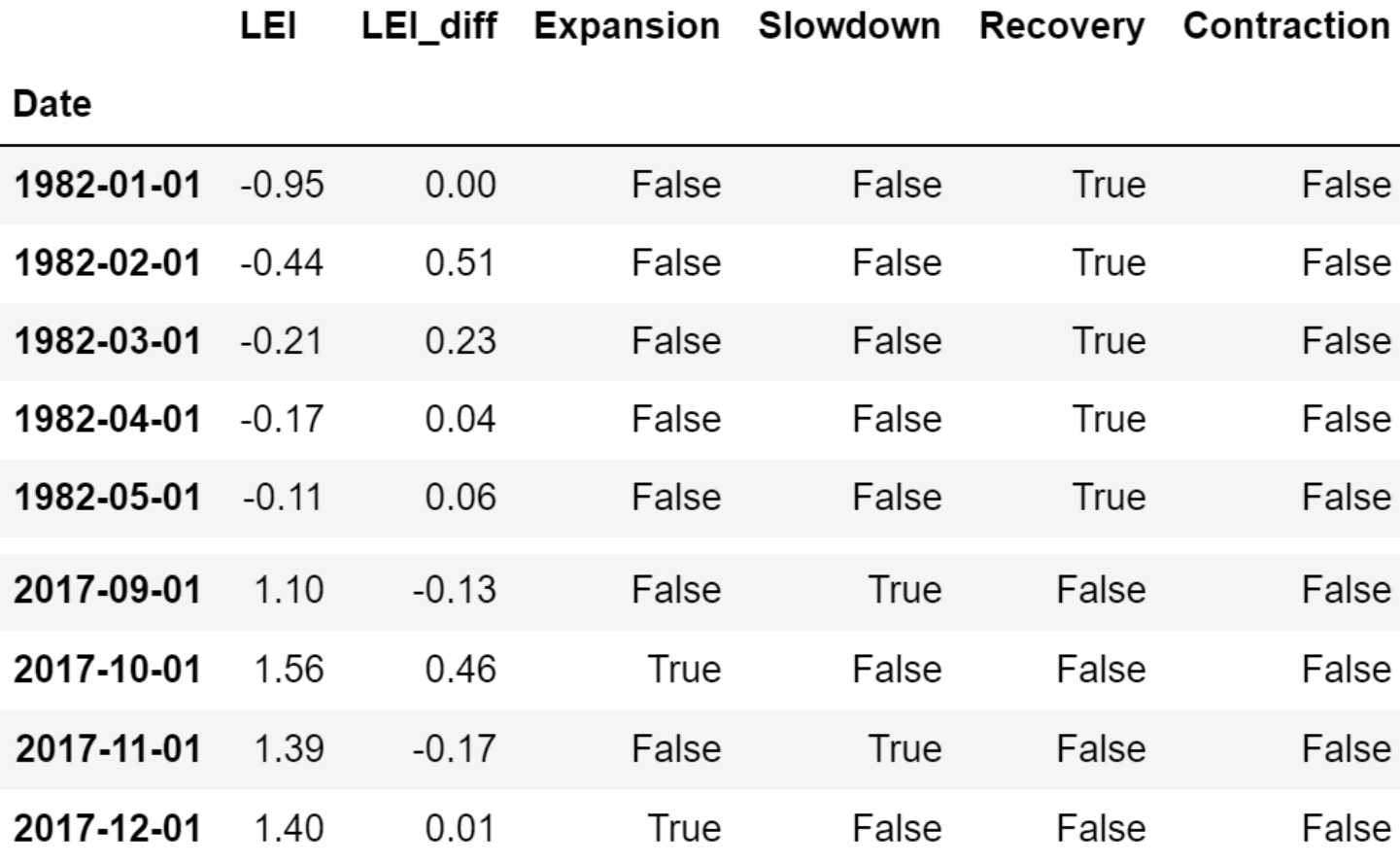
From the first (“LEI” in the table below) and second order growth rates (“LEI\_diff”) of the LEI we extract the 6-month forward economic state. Positive first and second order LEI growth correspond to Expansion, positive and negative first and second order growth correspond to slowdown, negative and positive growth to Recovery, and negative first and second order growths to Contraction. A summary table is attached below.

*Economic State Classification*



For a initial take we have classified 1982-2017 data into four economic regimes. As shown below, changing levels of the LEI result in dynamic switching between regimes. However, the current baseline model shows frequent switching between regimes as shown in the table below. To smooth out the switching frequencies, various smoothing methodologies such as the Hodrick–Prescott filter.

*Economic State Classification between 1982 - 2017*



Beyond manual feature engineering, we also adopted a unsupervised learning algorithm to identify the economic states. The major method for this job is PCA and t-SNE. Since we believe the connection between economic variables and factor performance is intrinsically nonlinear, we prefer the nonlinear dimension reduction algorithm t-SNE (t-distributed stochastic neighbor embedding). t-SNE is an ML algorithm to perform dimension reduction. Compared to PCA, it’s a nonlinear dimensionality reduction algorithm used for mapping high-dimensional raw data into lower dimensions (usually 2D or 3D for data visualization purpose). After that, the ‘similar’ data (judged by model itself) are mapped into near space while the dissimilar ones mapped into distant one.

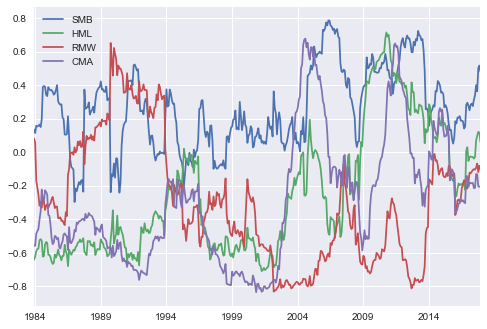
### 

### Factor Model

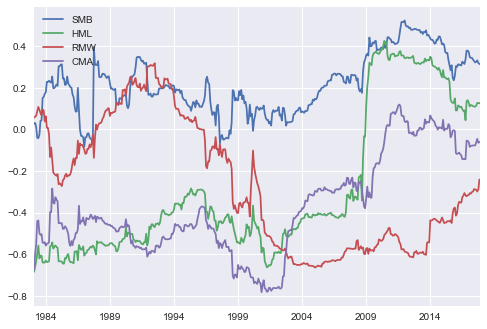
The factor model used the Fama-French 5-factors returns data. The data is sourced from the Kenneth R. French data library, which used the 2017 December Center for Research in Security Prices (CRSP) database. The frequency of the returns data is in daily that covered period between July 1963 to December 2017. According to the data library, the Fama-French 5-factors portfolio was constructed by using 6 valued-weight portfolios that were formed by size, book-to-market, operating profitability and investment. The 5-factors are market excess return, SMB (small minus big), HML (high minus low), RWA (robust minus weak) and CMA (conversative minus aggressive).

Here, we also plot the correlation graph (the 2-year rolling correlation for these factor returns with market excess return). We found some significant pattern in this graph. RMW correlation is positive before 90s and turns into negative after that. SBM correlation is positive and higher when the market is in good condition, and vice versa. HML correlation is not that stable compare to the other ones. CMA correlation is generally not high, but high correlation usually imply bad market condition. The exponentially-weighted rolling curve is also displayed. (analysis TBD)

*2-Year Rolling Correlation*



*2-Year Exponential-weighted Rolling Correlation of Factor*



Beyond the Fama-French factors, there have been numerous developments into the design of various factors that capture different sources of return. For example, MSCI provide Barra Factor Models for almost all equity markets in the world. Due to limitation on the data source, we constructed a long-short market neutral strategy by using the Fama-French 5-factors return data because these data are well published from trusted source.

Prior to building the proposed regime identified factor model, we built a universal equal-weighted factor model to serve as the baseline. Both proposed and baseline models exclude the market excess return factor as we do not want to have any exposures to the market and keep it market-neutral. The 5-factors model, thus, reduced to 4-factors (i.e., SMB, HML, RWA and CMA), which will be described in the data section.

The baseline model approach is described as the following. The first step is to calculate the factor weights in a monthly basis using a one-year rolling daily factor returns data. For example, the first set of weight falls on July 1984 because in the dataset, we used one-year daily factor returns data that started on June 1983 and ended on June 1984. The factor weights can take on only four numbers, which are 1, 0.5, -0.5 and -1. The sum of the four weights are zero to fulfil the market-neutral requirement. The assignment of the weights to the 4-factors are followed by computing the raw factor weights and ranking them in a descending order. As such, the factor that receives largest raw weight would take on the aforementioned factor weight of 1, and the second largest raw weight factor would be assigned 0.5 and so on. The reason of picking a one-year time window is to calculate the information ratio and correlations among the 4-factors, which are the ingredients for raw factor weights computation. The formula for raw factor weights is given as

where time t is in month and constant is set to be 1. The correlation threshold is 0.3. The implication is to filter out the low and negatively correlated factor pairs and only account for the information ratio in such case as a measurement for raw weight. Once the raw weights are computed, we multiply the aforementioned corresponding factor weights to the factor returns to obtain the portfolio return for time t.

where r\_t is the portfolio return at time t and f is the corresponding factor that is selected based upon the time t raw weights.

### Proposed regime identified factor model

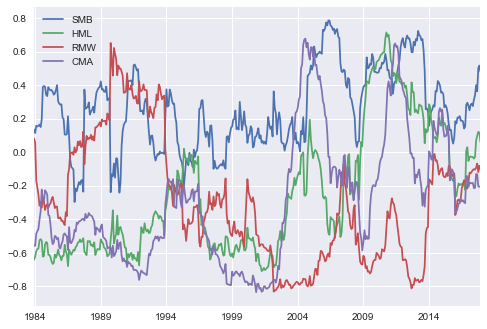
TBU - To be updated in further iterations.

### Fama-French Factor Returns Data

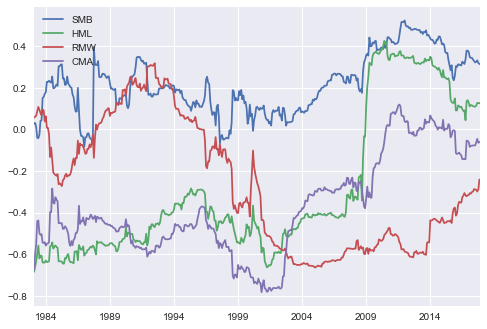
The factor model used the Fama-French 5-factors returns data. The data is sourced from the Kenneth R. French data library, which used the 2017 December Center for Research in Security Prices (CRSP) database. The frequency of the returns data is in daily that covered period between July 1963 to December 2017. According to the data library, the Fama-French 5-factors portfolio was constructed by using 6 valued-weight portfolios that were formed by size, book-to-market, operating profitability and investment. The 5-factors are market excess return, SMB (small minus big), HML (high minus low), RWA (robust minus weak) and CMA (conversative minus aggressive).

Here, we also plot the correlation graph (the 2-year rolling correlation for these factor returns with market excess return). We found some significant pattern in this graph. RMW correlation is positive before 90s and turns into negative after that. SBM correlation is positive and higher when the market is in good condition, and vice versa. HML correlation is not that stable compare to the other ones. CMA correlation is generally not high, but high correlation usually imply bad market condition. The exponentially-weighted rolling curve is also displayed. (analysis TBD)

*2-Year Rolling Correlation*



*2-Year Exponential-weighted Rolling Correlation of Factor*

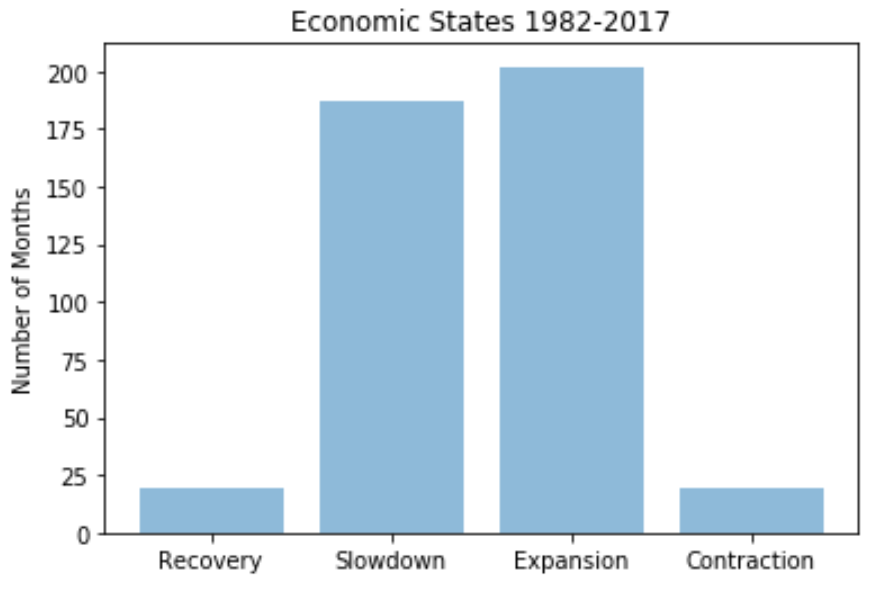


**IV. Results**

**Preliminary Results**

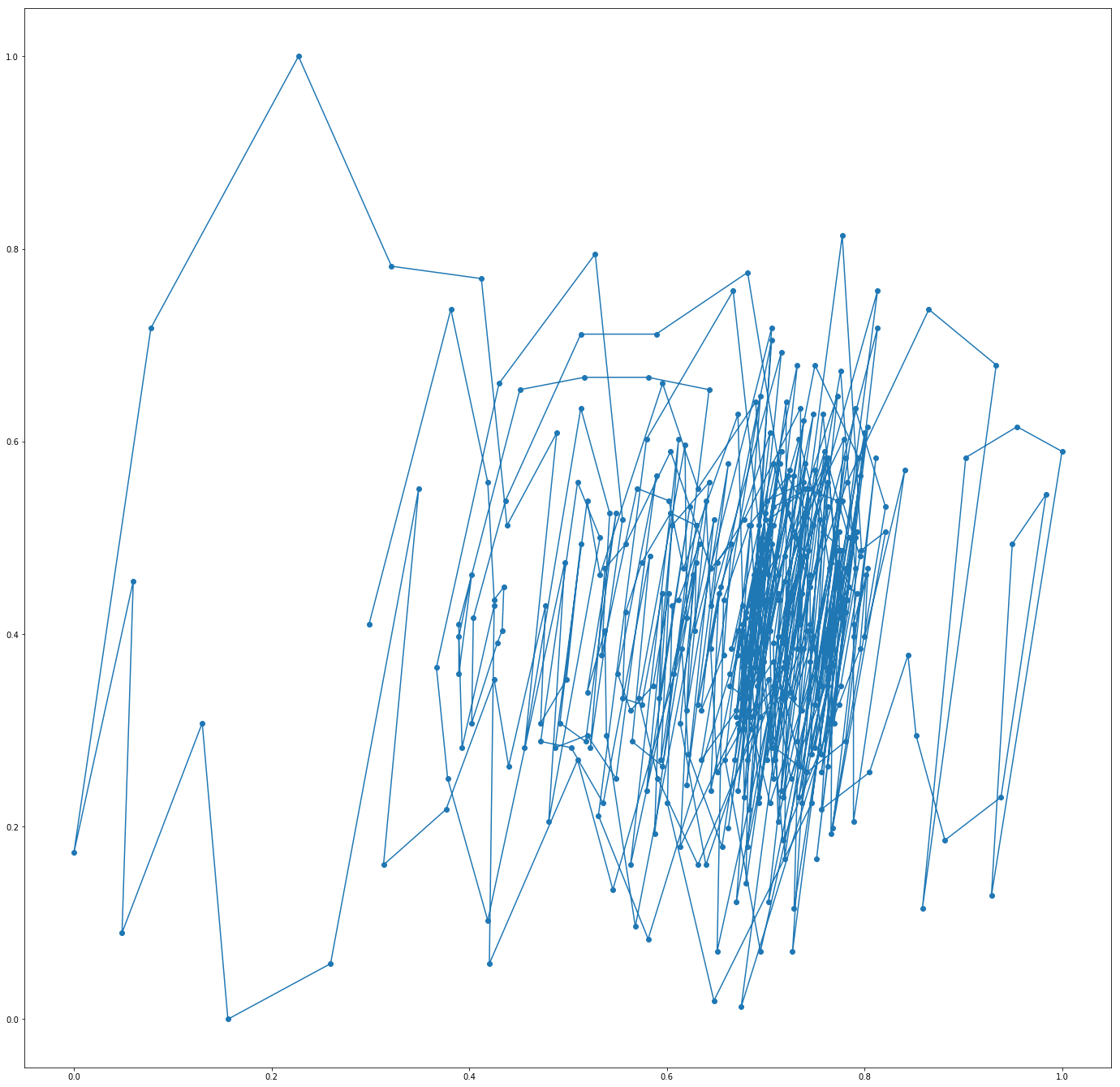
From the economic state transition model we found that the economy is under ‘Expansion’ and ‘Slowdown’ for the majority of the sample since the LEI itself is seldom negative (which makes intuitive, since most time the US economy is growing between 1980 and 2017).

*Frequency of Economic States*



Then we also plot the path of the Leading Index. As seen from the graph below (X-axis for the 1st order growth, Y-axis for the 2nd order growth). We find that a large part of the variance comes from the change across the Y-axis; 2nd-order growth of the LEI accounts for the changes in the states more than the 1st-order growth does.

*Path of Leading Index: 1st order growth on X-axis, 2nd order growth on Y-axis*



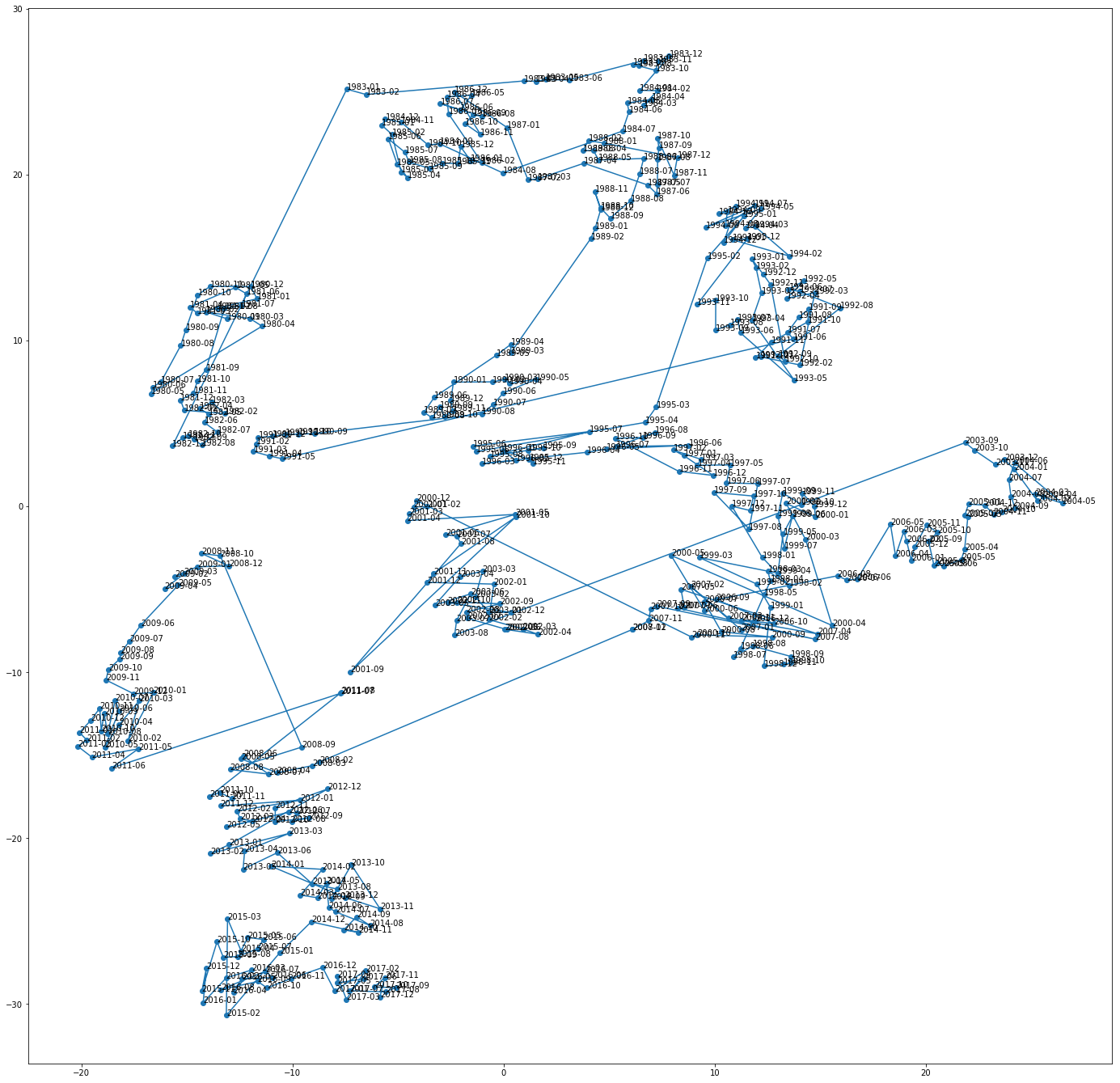
Then we apply t-SNE using 8 economic variables summarized below.

Economic Variables - Inputs for t-SNE

|  |  |
| --- | --- |
| **Variable** | **Description** |
| RATEINF/INFLATION\_USA | Inflation rate, YoY, with seasonal adjustment |
| ISM/MAN\_PMI | Manufacture PMI |
| ISM/AVG\_HR | Average working hours |
| FRED/M2SL | M2 increase rate, YoY |
| FRED/T10YFFM | 10-year rate premium |
| FRED/PERMIT | Newly-issued housing permits |
| FRED/PAYEMS | Nonfarm payroll data |
| FRED/ICSA | Monthly initial claims (unemployment) |

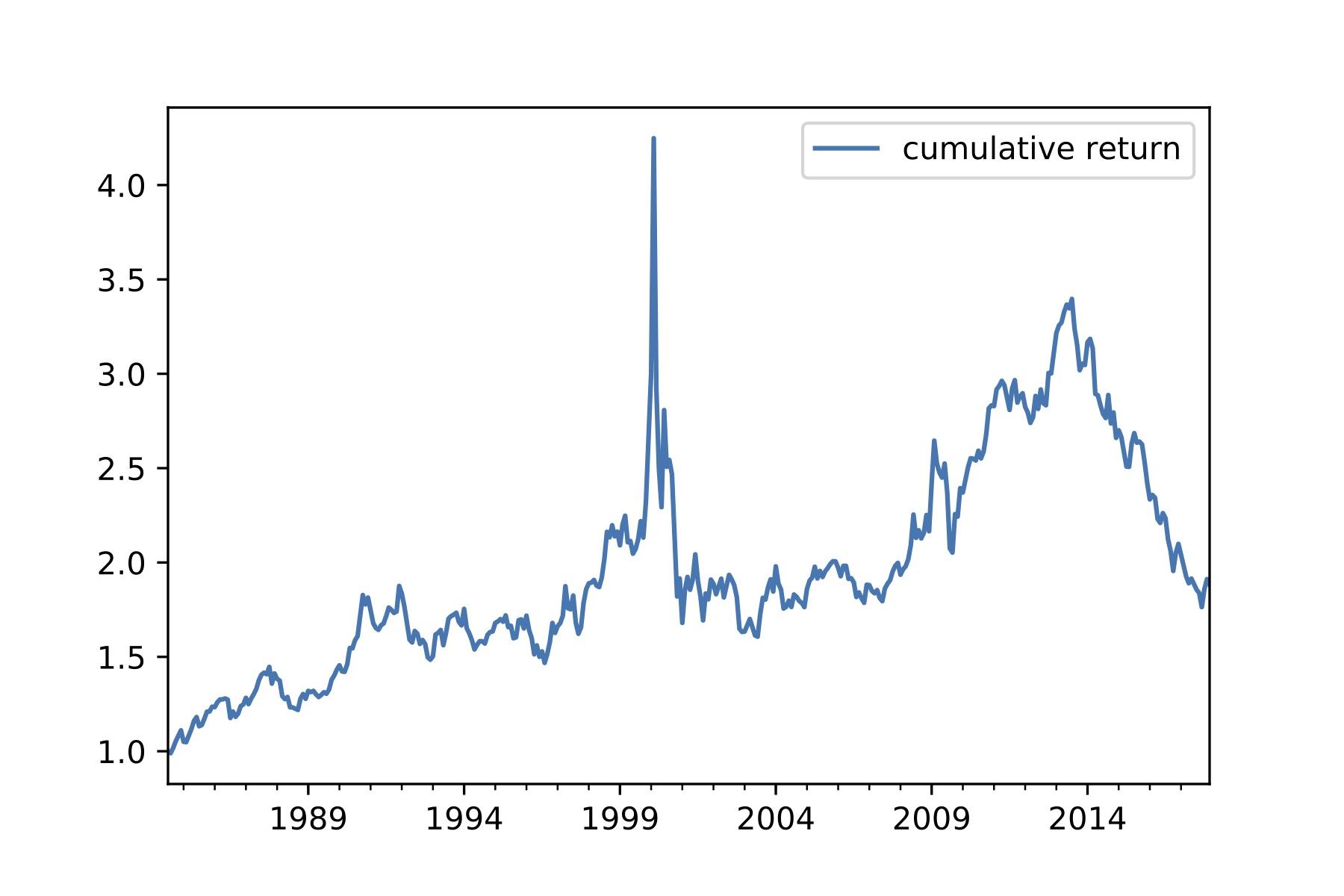
The graph below shows the result from t-SNE analysis. On this graph the two axes are nonlinear functions of all 8 variables. We can observe that for consecutive months their position is generally clustered together. However, there are months where jumps do occur. Our next step is to find out whether they indicate real regime changes and predict structural shifts in factor returns.

T-SNE Dimensionality Reduction

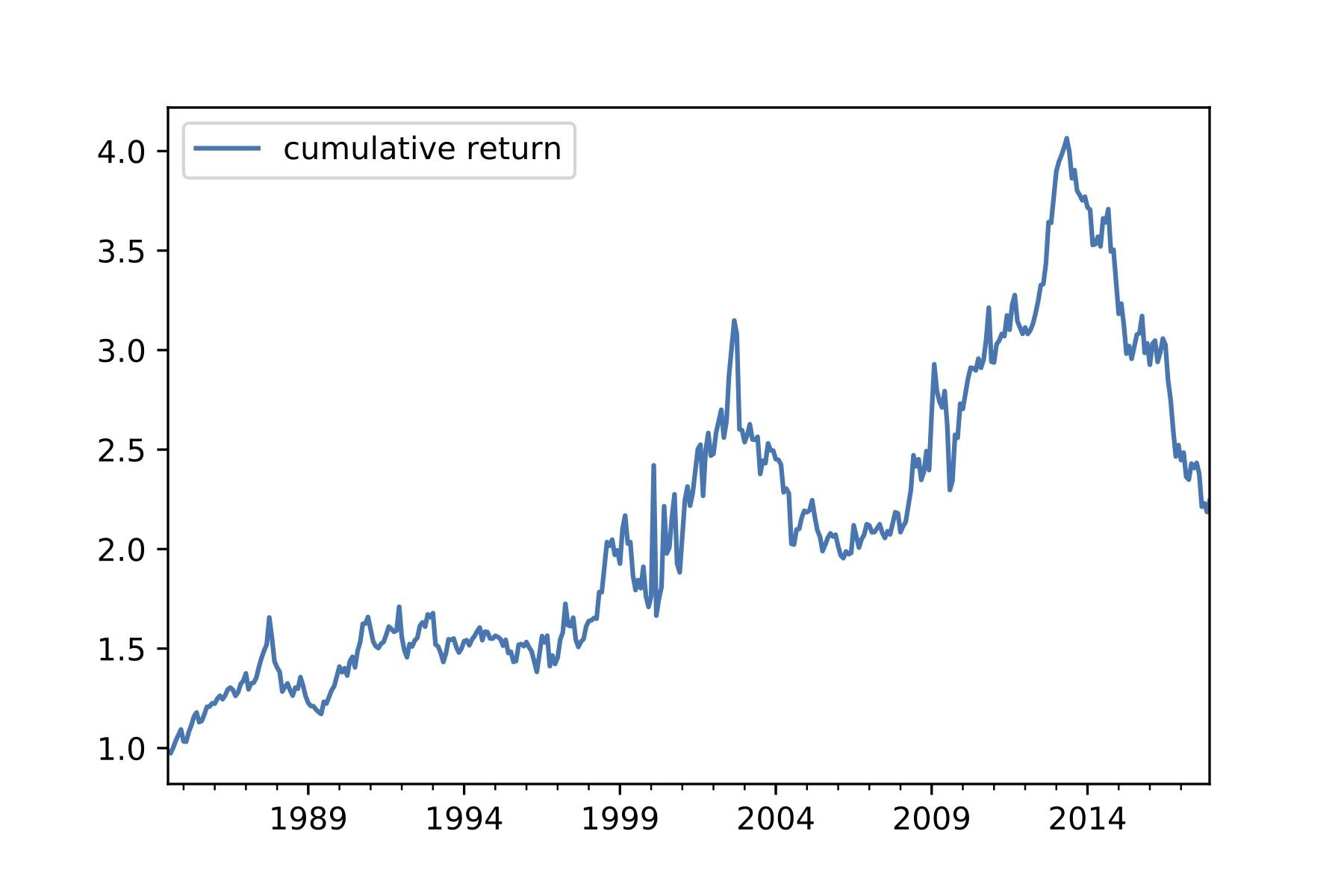


The baseline factor model is an equal weighted long-short portfolio of four Fama-French factors, holding 2 long positions for outperforming factors and 2 short positions for underperformers for each economic state. Shown below is the cumulative return of the long-short benchmark portfolio.

*Cumulative Return of the Benchmark Portfolio*



*Cumulative Return of the Proposed Regime identified Portfolio*



**Main Result**

TBU - To be updated in further iterations.

**Conclusion**

TBU - To be updated in further iterations.

# **Appendices**

TBU - To be updated in further iterations.

# **Reference**

TBU - To be updated in further iterations.