Determinants of and Prediction of Investor Exuberance Using the K-METHOD: Reduction-Regression-Correlation-Regression (RRCR).

By

Karina Krishnan
Beachwood, OH, 44122

The Research Question:

Purpose

The objective of this study is to introduce a new statistical analysis where there are many explanatory variables, and we need to know which specific explanatory variables we need to focus on. To accomplish this, I developed a simple Reduction-Regression-Correlation-Regression (RRCR) analysis that I call the "K-Method" (Karina Method), which is easily programmable in simple steps and can adapt to new data.

The Application

I use the K-Method to explain "cyclically-adjusted price-to-earnings" (CAPE) ratio levels of a stock market, which is one standard metric used to evaluate whether a market is overvalued, undervalued, or fairly valued. This metric was developed by Nobel Prize winner, *Professor Robert Shiller* of Yale University, which examines the times when the market is "irrationally exuberant". These are the times when the stock or index price is trading at a high multiple compared to the cyclically adjusted (inflation, etc.) earnings of the company or an index. So, investors are willing to pay much more per share, as compared to the firm's average earnings.

Robert Shiller

The CAPE is constructed as follows, by Professor Shiller. The real (inflation-adjusted) S&P composite price is calculated using the Consumer Price Index every month. This is divided by the 10-year monthly real S&P Composite Earnings (inflation-adjusted using the Consumer Price Index) of the immediate past 120 months. Thus, the inflation-adjusted stock price is divided by

the average of the last 10-year inflation-adjusted earnings to determine CAPE for each month. This measures investor optimism or pessimism.

The CAPE can be influenced by many factors. The current prices of the stocks and stock indices, domestic and worldwide, reflect on the CAPE. In addition, economy-wide variables, such as GDP and inflation, affect stock prices. Inflation increases costliness, which decreases buying power to a degree, eventually affecting stock prices. When GDP is higher, there is optimism about economic output, which tends to help stock prices. Fixed-income securities prices can affect stock prices by affecting the cost of funds for firms. Money is more expensive to borrow when interest rates are higher, consequently, depleting company profit margins. Commodity prices can affect stock prices via the cost of inputs for firms. Both fixed-income securities and commodities also affect stock attractiveness in terms of being alternative investment opportunities for investors.

Hypothesis

Among the various economy-wide variables, fixed-income variables, stock market variables, and commodity variables, only certain variables are most correlated with the price-to-earnings multiple at which the market trades, and the same (few) variables can also be used to predict out-of-sample CAPE.

Research Question

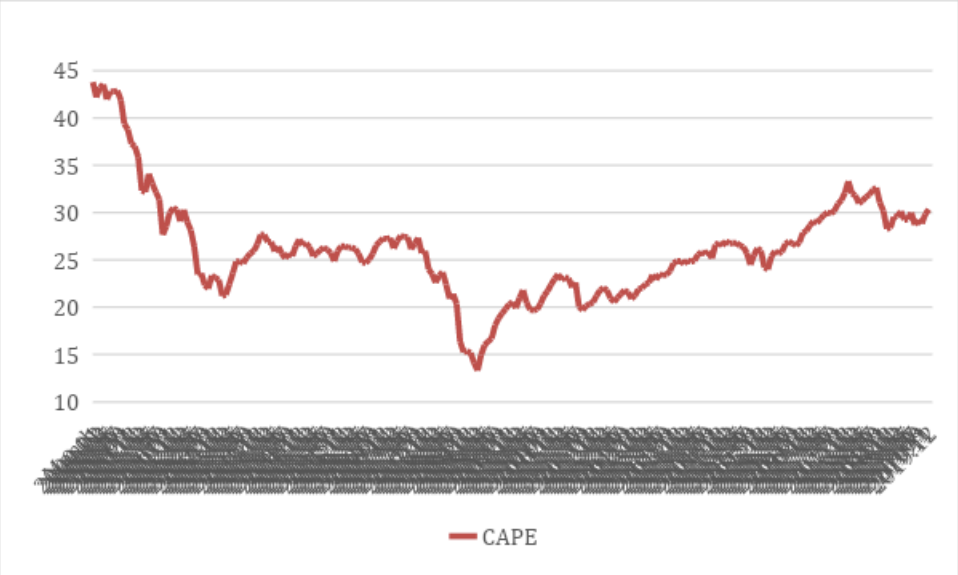
Using a simple programmable algorithm, which variables influence investor exuberance in the stock market, at any point in time, and can be used to predict out-of-sample CAPE in the near future?

Methodology:

Data

The "Cyclically Adjusted Price-to-Earnings ratio" CAPE data comes from using a market index that was used in the book/paper, and website of *Professor Robert Shiller*. I use monthly data from Jan 2000 to Dec 2019 (240 months), and match with economy-wide data for the same months (8 variables) taken from MacroTrends website, 9 Fixed-Income variables taken from MacroTrends website, 28 Stock index and stock market prices variables taken from MacroTrends website and Center for Research in Securities Prices database, and 12 commodities prices variables taken from MacroTrends website (a total of 57 possible explanatory variables), as compared to 240 data points.

I divide the data into 2 equal sub-samples: Jan 2000 through Dec 2009 is the training sample, and Jan 2010 through Dec 2019 is the sample for "out-of-sample" analysis.



CAPE Plot by month: the measure of investor exuberance

The explanatory variables are:

Economy wide	Fixed Income	Stock Market	Commodities
Housing Starts	1-month LIBOR Rates (%)	NASDAQ Composite	Tobacco
U.S. National Unemployment Rate (%)	3-month LIBOR Rates (%)	S&P 500 Index	Crude Oil Brent
Initial Jobless Claims	6-month LIBOR Rates (%)	Dow to GDP Ratio	Natural Gas
Real Retail Sales ($)	1-year LIBOR Rates (%)	Hang Seng Composite Index	Copper
Auto and Light Truck Sales(monthly)	TED Spread (%)	DAX 30 Index	Soybeans
Debt to GDP Ratio (%)	Federal Funds Rate (%)	Nikkei 225 Index	Corn
Industrial Production level	10-year Treasury Rate (%)	Microsoft Stock Price	Cotton
Capacity Utilization Rate	1-year Treasury Rate (%)	Coke Stock Price	Wheat
	5-year Treasury Rate (%)	IBM Stock Price	Coffee
		Chevron Stock Price	Oats
		Apple Stock Price	Sugar
		PG Stock Price	Soybean Oil
		CAT Stock Price	
		BA Stock Price	
		JNJ Stock Price	
		3M Stock Price	
		MERCK Stock Price	
		DISNEY Stock Price	
		McDonald Stock Price	
		JPM Stock Price	
		Walmart Stock Price	
		Nike Stock Price	
		AXP Stock Price	
		Intel Stock Price	
		Citigroup Stock Price	
		Cisco Stock Price	
		GS Stock Price	
		United Health Stock Price	

Analysis

Step 1: REDUCTION:
Identifying 2 Principal Components from each group = 8 explanatory variables, whose explanatory power for the group of explanatory variables is shown below.

The PCs are identified using the Eigenvectors output of running the Principal Components Analysis (PCA) using Stata software, after computing the *Z scores* of all explanatory variables (subtract mean and normalize by the standard deviation), to bring all variables to eliminate inter-variable value differences. Their explanatory powers are shown below.

The *pca* command in STATA outputs Eignvalues (showing the proportion explained by each PC), and the Eigen vectors that need to be matrix multiplied to get the PC's

Table 1: Explanatory Power of Principal Component 1 and Principal Component 2 (orthogonal to PC 1)

Step 1: PCA	CAPE
Average of CAPE	26.11
PC	8 (top 2 of each category of explanatory variables)
Economy	86.22%
Fixed Income	94.46%
Stocks	73.79%
Commodities	82.80%
Average	84.32%

Step 2: REGRESSION (1st stage):
Regression on Cyclically-Adjusted Price-to-Earnings Ratio (CAPE) in the training period using the 8 PC's identified above as explanatory variables.

The "*Regress"* command in STATA outputs the regression coefficients of each variable and its significance "t statistic."

CAPE is regressed on the 8 PCs in the training period. The significant ones are shown below:

Table 2: Significance of the significant 2nd Stage PC that explains CAPE:

N =	Regression on CAPE 120 Significant Explanatory Variable
PC1 of Fixed Income Variables	t statistic = 8.01
PC1 of Stock Market variables	t statistic = 13.02

Step 3: CORRELATION:
We first check the loading of the 2 PC1 variables (above) that were significant in explaining CAPE, on the 8 PCs.

Table 3: Correlations of PC1s and original variables:

	PC1 of Stock Market Variables	PC1 of Fixed Income Variables
10-year Treasury Rate (%) Fixed Income Variable	80.99%	77.32%
5-year Treasury Rate (%) Fixed Income Variable	82.00%	88.73%
NASDAQ Composite Index Stock Market variable	87.67%	74.55%
S&P 500 Index Stock Market variable	83.91%	82.75%
Dow to GDP Ratio Stock Market variable	90.71%	71.03%

Step 4: REGRESSION (2nd stage):
CAPE in the training period is now regressed on the original variables (the explanatory data) above. The significant ones are shown below.

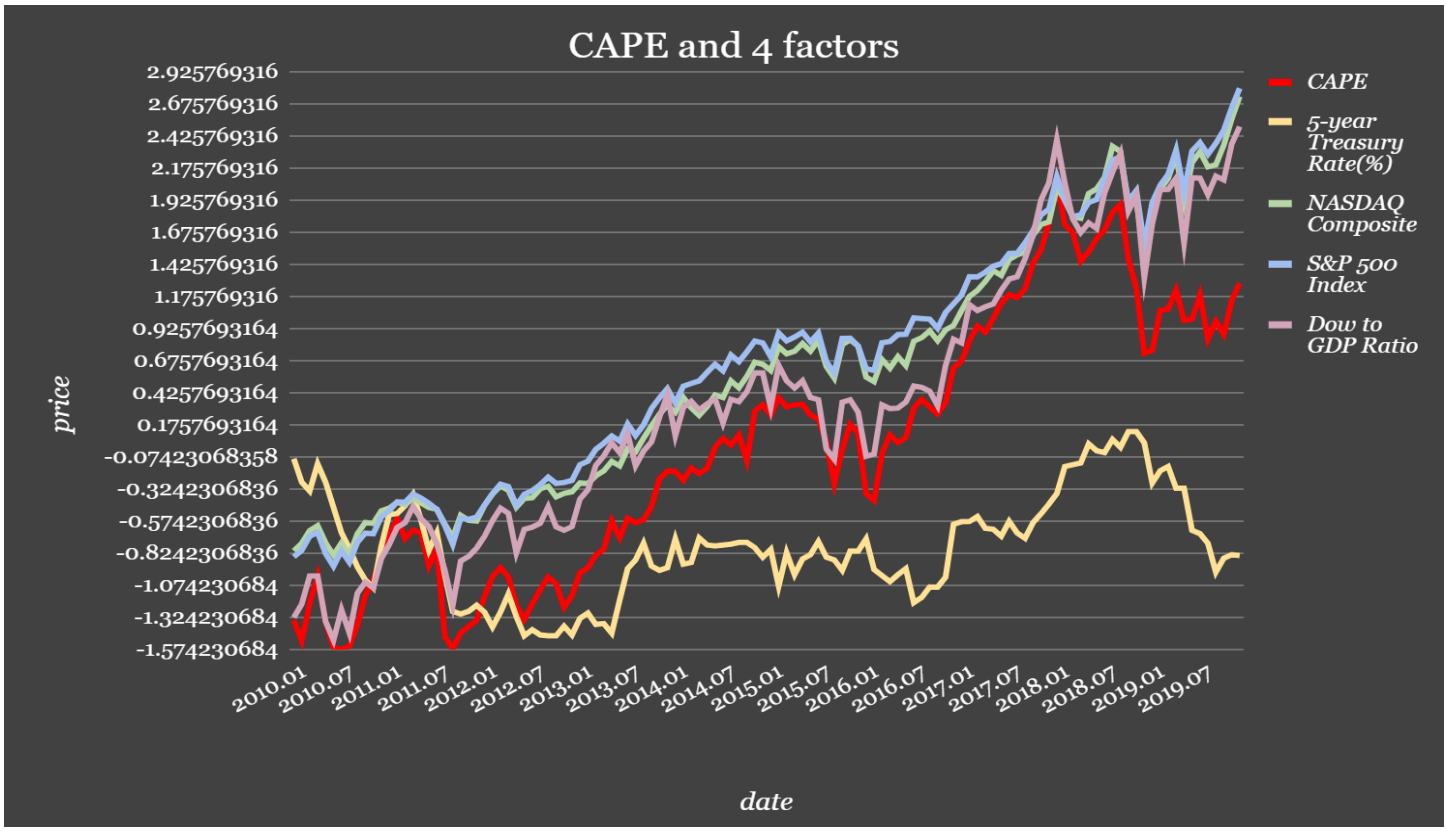
Table 4: Significance of CAPE on original explanatory variables. The significant ones are:

	Regression on CAPE
N (training period) =	120
	Significant Explanatory Variable
5-year Treasury Rate (%) Fixed Income Variable	t statistic = 12.41
NASDAQ Composite Index Stock Market variable	t statistic = 17.97
S&P 500 Index Stock Market variable	t statistic = 18.34
Dow to GDP Ratio Stock Market variable	t statistic = 22.51

These are the final remaining explanatory variables that can also be used to determine out-of-sample CAPE.

Out of Sample Correlation

In the out-of-sample period (2010-2019) that was not used above, the plot of CAPE and the 4 explanatory variables that determine CAPE in the training period is shown below. We find that these variables strongly correlate with CAPE even in the out-of-sample period, and can thus continue to determine CAPE in the near future.



Analysis and Discussion:

My "K-Method" involves 4 steps. In the first step, a Principal Components (PC) Analysis is conducted to identify the PCs with the most explanatory power over sets of explanatory variables, which are numerous as compared to data points (in the training period). 2 PCs are then identified for each explanatory variable set that, on average, have about 84% explanatory power of all original variables. Then these PCs are regressed on CAPE to determine which ones are significantly correlated with CAPE. Then the correlations of these significant PCs with the original large number of explanatory variables are examined. The original explanatory variables with the highest correlations with all significant PCs are selected. Then, a second step regression is performed of these variables on CAPE. The significant ones are the ones that significantly affect the training period's CAPE and may be used to determine out-of-sample CAPE in the near future.

I found that these are one fixed income and 3 stock market index variables. I then check the plots of these explanatory variables with the out-of-sample CAPE. These explanatory variables may change as the data changes, but the method will not change.

Conclusions:

My "K-Method" involves 4 steps: Reduction-Regression-Correlation-Regression (RRCR). This simple method is easily programmable and adaptable for data analytics when the number of explanatory variables is large compared to the sample size. I used Stata programming combined with Excel, but the Python program can also be easily written. In this research, I examine the main determinants of the Cyclically-adjusted Price-to-earnings ratio (or *Shiller's* CAPE), which is a measure of investor exuberance in the stock markets, whether or not irrational. I find that for the first 2 decades of this millennium, certain interest rates and stock index prices matter to determine CAPE levels and may predict CAPE levels.

Investor exuberance in stock markets depends on the cost of funds for companies (the interest rate) and the general stock market optimism (as measured by stock market indices).

References:

Irrational Exuberance | Princeton University Press

http://www.econ.yale.edu/~shiller/data.htm