*Journal of Applied Finance & Banking, vol. 7, no. 6, 2017, 85-97*
*ISSN: 1792-6580 (print version), 1792-6599 (online)*
*Scienpress Ltd, 2017*

# Capital Asset Pricing Model – investigation and Testing

**Huang Xian Yu**[1]

**Abstract**

This paper aims to develop testing model based on logistic regression with three factors to investigate the equity premium portion of CAPM model. It includes (1) literature review on equity premium of CAPM (Capital Asset Pricing Model) model; (2) Set up logistic regression model; (3) Data collection from Datastream; (4) Use of Matlab in regression; (5) Data input in logistic regression; (6) Create a homemade model to prove the nonexistence of equity premium puzzle. Together with investigating the proper definition of risk-free rate, this paper investigates and tests the basic model of CAPM.

**JEL classification numbers:** G1
**Keywords:** CAPM model, risk-free rate, risk premium, logistic regression, volatility index.

## 1 Introduction

This paper investigates the proxy for risk-free rate used in past researches and argues that the proxy for risk-free rate used in the past researches is underestimated. Historical return has shown abnormally high returns on S&P 500 over that of U.S. government bond, which is generally accepted as risk-free. Gold has been considered as risk-free theoretically, this risk-free rate proxy should be the larger of Treasury yield or return on gold.

---

[1] Department of Finance, Chu Hai College of Higher Education, Hong Kong

Article Info: *Received* : August 7, 2017. *Revised* : August 30, 2017.
*Published online* : November 1, 2017

This paper also doubts that risk premium might be wrongly estimated in the past and went through the historical data related to the risk premium and made experiment on two main variables based on the basic formula of Capital Asset Pricing Model. We investigate the historical data related to the equity premium puzzle and made experiment on two main variables considered based on the basic formula of Capital Asset Pricing Model (CAPM), including the estimation of risk premium using other 4 factors and the selection of appropriate risk free rate adopting the rates of return from gold. Also, those factors on risk premium will also be considered separately in different low-high situation of the observed risk-free rate.

## 2 Literature Review

Debates on the equity premium puzzle, the unexplained return from risky security in excess of the returns from risk-free security, has been for more than three decades since 1985 by Mehra and Prescott. In the past, US government has been regarded and accepted as risk-free and risk premium of a securities' return is measured as any excess return of the security over the US government bond. This paper argues that the proxy for risk-free rate used in past researches is underestimated and more appropriate proxy for risk-free rate should also take return on gold into account.

In Mehra and Prescott (1985), it was illustrated that using classical theory, returns on stocks should only be 1% higher than that of US government bonds. Given that the average return of S&P500 was 7% (over 1889 – 1978) was too substantial, given that the short term virtually risk-free debt was below 1%. The study covered the S&P performance over 1889 and 1978. The paper leads to debates on the existence of excessiveness of equity premium leading to the challenge of CAPM model.

Siegel (1992) expanded the study to year between 1802 and 1990 and concluded that the risk premium over a longer time period is relative smaller. It was concluded that real return on stock remained stable, while real return on short-term riskless debts fall sharply.

Campbell and Cochrane (1999) modified individual preference to derive a consumption based model in an attempt to solve equity premium puzzle and therefore sustainability and reliability of CAPM. It assumes that utility is not only affected by current consumption, but also by a historical level known as habitat level, which slowly and nonlinearly to historical level. The model was able to break the link between intertemporal substitution, risk aversion and precautionary savings present in standard power utility model. The

model is consistent with both consumption and asset market data. However, Mehra (2003) questioned whether investors actually have the huge time varying counter-cyclical variations in risk aversion postulated in the model. He concluded that the doubt of equity premium does not only exist in U.S., but probably over the world such as Japan, France, UK and Germany. The above countries plus U.S. accounted for 85% of global equity value. His study shows that equity premium puzzle is, very likely, a worldwide phenomenon.

# 3 Methodology

1. To check the existence of equity premium puzzle, equity premium is divided into different categories, and four categories are set to very low (less than -7%), low (-7% to 2%), high (2% to 5%), and very high (more than 5%). They are formed so as to find the effect of X factors on different categories of equity premium.
2. Three factors affecting risk premium including VIX, Production Manager Index, and Industrial Production Index, were used in this research project to explain the sustainability and reliability of CAPM.
3. Under the ordered logistic regression, two models are created with two different risk-free rate - treasury yield and the larger of treasury yield and gold return. Under each model, it is also created two calculation methods for the beta of the independent variables. One method remains all beta of the independent variables and the other one only remains the significant beta of independent variables and sets the insignificant beta of independent variables to zero. Use a logistic regression to determine factors selection, test models on the reliability of the logistics regression using (a) treasury yield and (b) larger of treasury yield and gold return.
4. Compare results of two models and develop a homemade model to prove the non-existence of excessive risk premium.

# 4 Data source and development tools

Data would be collected from the DataStream of Thomson Reuters on the US stock market between year March 1990 and January 2015 contained in datastream. With risk premium being the dependent variable, Standard and Poor's 500 Composite was used to calculate the market return (Rm) and 3 months US treasury bill is used as risk-free rate ($R_f$). Another risk-free rate calculated by gold price referred to NYMEX gold 3-month futures. The data of gold future, NYMEX Gold Futures #1 (GC1) was collected from Quandl, an online paid database. The data of volatility index (VIX) was collected from the Federal Reserve Bank of St. Louis. MATLAB 2014b is used for statistical calculation

and presentation.

Several independent economic variables are collected, which are filtered the insignificant factors to the risk premium from those variables. Seven independent variables will be chosen from the database. These variables included consumer confidence index, CPI for all items in all urban, industrial production of manufacturing (INDUSTPRO), money supply in definition 1 (M1), money supply in definition 2 (M2), personal consumption expenditures (PCE), purchasing manager index (PMI), Volatility index (VIX) and unemployment rate. Data abstracted from the datastream is divided into different combinations as factors of the ordered logistics regression.

## 5 Collinearity

Problem of multicollinearity by calculating the correlation between the factors are also performed and highly correlated factors are excluded. An ordered logistic regression model with the three economic factors is hypothesized to provide a nonlinear relationship between economic factors and the categories of equity premium and a result of predicted equity premium in terms of probability.

## 6 Determination of risk free rate

The risk-free rate of gold is derived from the cost of carry model, which expresses the future price as a function of the spot price and the cost of carry. The model specifies:

$$F = Se^{(r+s-c)t}$$

where F= Futures gold price
S = Spot gold price
C= storage cost
r = risk-free rate

By inserting values of F, S and C, risk free rate r is attained.

## 7 Independent variables

- Purchasing Manger Index (PMI) – an index indicating the overall economic health of manufacturing sectors by considering new orders, inventory levels, production, supplier deliveries and the employment environment. It is used to indicate the overall performance of manufacturing.

- US Consumer Price Index (CPI) in all items – is one of the most important indicator of inflation in many countries that does reflect the purchasing power for local residents.
- Money Supply M1, M2 – measure the most and second most liquid component of the money supply. They are related to the monetary policy that US government adopted and will affect the economic of US and may indirectly have an impact to the equity market.
- Personal Consumption Expenditure (PCE) – is chain type price index that reflects consumption behaviors on product and service. It reflects the reality of the economy of US in term of price level.
- Unemployment rate – reflects the productivity of an economy.
- Volatility index (VIX) – is an indicator describing the overall environment of the market and atmosphere on investors.

To determine whether the independent factors are significant or not, it is implemented the ordered logistic regression with the factor independently by Matlab.

**Model 1**

Assumption: **Treasury yield is the risk free rate**

$H_0$: Beta = 0

Ha: Beta is not equal to zero

The results are as follows:

| Risk Premium | | VERY LOW | LOW | HIGH | VERY HIGH |
|---|---|---|---|---|---|
| **VIX** | Intercept | -5.063909621 | -4.137948046 | -0.562080658 | 1.213533432 |
| | VIX | 3.93047749 | 2.956327904 | 1.080278949 | -1.301651863 |
| | | | | | |
| CPI | Intercept | -1.021081706 | -1.20735859 | 1.466023445 | 0.171607147 |
| | CPI | 0.13538424 | 0.583323566 | -1.267437503 | -0.186515392 |
| | | | | | |
| PCE | Intercept | -5.562812773 | -3.861491601 | 1.320081465 | 0.069261648 |
| | PCE | -1.939474941 | -0.84778517 | -0.007125931 | -0.450458925 |
| | | | | | |
| M2 | Intercept | -5.481678159 | -4.562159354 | 0.921299571 | 0.081953397 |
| | M2 | 2.492324131 | 2.211398151 | 0.847664699 | -0.454133883 |
| | | | | | |
| M1 | Intercept | -5.739773878 | -4.302858067 | 1.478561112 | -0.115490561 |
| | M1 | 0.633598719 | 1.01525587 | -0.819732156 | 0.198624188 |
| | | | | | |
| PMI | Intercept | -5.317458201 | -4.114634359 | 1.591158467 | 0.005940749 |
| | PMI | -3.416903685 | -2.498216306 | -2.028519574 | -2.713970097 |
| | | | | | |
| INDUSTPRO | Intercept | -5.852446035 | -4.305935675 | 1.319491014 | -0.075024802 |
| | INDUSTPRO | 0.499368658 | -1.058135092 | -0.676238044 | 0.703050664 |

For 95% Confidence Interval,
t-test>1.96, t-test<-1.96, reject the null
1.96>t-test>-1.96, cannot reject the null (shown in red)

The results above show that, under 95% confidence interval, coefficients of only VIX, M2 and PMI fall outside 1.96 and -1.96 showing they have significant beta with beta of all other factors being equal to zero.

Using VIX, M2 and PMI to rerun the logistic regression with results show as follows:

| Combination | Very low | Low | High | Very high |
|---|---|---|---|---|
| intercept | -4.724869254 | -3.904649278 | -0.491526251 | 1.110446574 |
| VIX | 2.973530191 | 2.378661509 | 0.941774105 | -1.133121667 |
| PMI | -2.080115092 | -2.207293922 | -2.073623761 | -2.754065891 |
| M2 | 0.702582267 | 1.176636666 | 0.567587995 | -0.015823331 |

Results shows that M2 cannot be rejected. Rerun VIX and PMI as factors by Matlab:

Ho: Beta = 0
$H_1$: Beta is not equal to zero

The result is as follows:

| Combination | Very low | Low | High | Very high |
|---|---|---|---|---|
| intercept | -4.799679045 | -4.054963956 | -0.555844267 | 1.151889777 |
| VIX | 3.347860061 | 2.888325077 | 1.137138309 | -1.210640969 |
| PMI | -2.182461983 | -2.298989371 | -2.070732834 | -2.76775613 |

Most of the betas of the factors are significant; VIX and PMI can be used as factors to predict the categories of equity premium and a result of predicted equity premium in terms of probability. Also, the problem of multicollinearity should be checked and the result is as follow:

**Correlation test**

Correlation test is performed between VIX and PMI as follows:

| Correlation test | VIX | PMI |
|---|---|---|
| VIX | 1 | -0.117514898 |
| PMI | -0.117514898 | 1 |

The correlation of VIX and PMI are small that the effect of multicollinearity is little. Therefore, It is conducted a 2-factor model under the model concerning Treasury yield as risk-free rate.

**Model 2**

Assumption: Risk free rate is the larger Treasury yield and gold return

$H_0$: Beta = 0

Ha: Beta is not equal to zero

| | | VERY LOW | LOW | HIGH | VERY HIGH |
|---|---|---|---|---|---|
| VIX | Intercept | -4.863877601 | -1.332991149 | 1.114249127 | 1.816643545 |
| | VIX | 4.162393574 | 1.628720066 | 0.07299488 | -1.742726413 |
| | | | | | |
| CPI | Intercept | -1.87298788 | -1.757885225 | -0.293659178 | -0.969903271 |
| | CPI | 1.393717467 | 1.879882892 | 0.873537784 | 1.057729831 |
| | | | | | |
| PCE | Intercept | -5.811601101 | -3.400144044 | 1.373331982 | 0.097388261 |
| | PCE | -5.502366945 | -4.922667678 | -3.883601124 | -2.199756485 |
| | | | | | |
| M2 | Intercept | -5.694071024 | -3.288938212 | 1.38552144 | 0.050306482 |
| | M2 | -5.263181985 | -4.765064249 | -3.633570149 | -1.830466795 |
| | | | | | |
| M1 | Intercept | -5.679609608 | -3.542720335 | 1.19458484 | 0.23069386 |
| | M1 | -5.847807182 | -5.536663192 | -3.944492794 | -1.088942992 |
| | | | | | |
| PMI | Intercept | -4.864463955 | -1.481761301 | 2.826211194 | 0.507937942 |
| | PMI | -6.5516857 | -5.875361815 | -4.378175692 | -2.78773347 |
| | | | | | |
| INDUSTPRO | Intercept | -5.922951331 | -3.741972865 | 1.152150102 | 0.019160237 |
| | INDUSTPRO | -5.67785482 | -5.068753289 | -3.63203221 | -1.463069255 |

Similarly, it is implemented the ordered logistic regression with all combinations of the above significant factors again and again, and then, it is found that the best combination of factors that provides the most significant beta is VIX, PMI and Industrial Production Index. The result is as follows:

| VIX, PMI and Industpro | Very low | Low | High | Very high |
|---|---|---|---|---|
| intercept | -5.734098153 | -3.186610329 | 0.004569239 | 1.538421398 |
| VIX | 3.94445389 | 2.000172114 | 0.465572465 | -1.517891181 |
| PMI | -2.469749752 | -2.944803539 | -3.095733403 | -2.527187498 |
| Industpro | -4.408104269 | -4.002611774 | -2.580310519 | -0.615873076 |

Also, the problem of multicollinearity should be checked and the result is as follow:

| Correlation | VIX | PMI | Industpro |
|---|---|---|---|
| VIX | 1 | -0.117514898 | -0.197754379 |
| PMI | -0.117514898 | 1 | 0.152188293 |
| Industpro | -0.197754379 | 0.152188293 | 1 |

The correlation of VIX and PMI are small that the effect of multicollinearity is little.
For 95% Confidence Interval,

t-test >1.96, t-test<-1.96, reject the null
1.96 >t-test > -1.96, cannot reject the null

The result shows that most of the beta of these 3 factors are significant in 95% Confidence Interval. In other words, it can be said that VIX, PMI and Industrial Production Index can be used as factors to predict the categories of equity premium and a result of predicted equity premium in terms of probability. Therefore, It is conducted a 3-factor model under the model concerning the large of Treasury yield and gold return as risk-free rate.

# 8 Original model

From the original regression models using treasury yield as the risk-free rate to calculate the risk premium, it was found that the logistic regression formulas are as follow.

Logistic regression remaining all coefficients:

$$\ln\left(\frac{\pi(\text{very low})}{\pi(\text{factors})}\right)$$

$$= -5.9276 + 0.1413\beta(\text{VIX}) - 22.0309\beta(\text{PMI})$$

$$\ln\left(\frac{\pi\ \text{low})}{\pi(\text{factors})}\right)$$

$$= -2.7212 + 0.0836\beta(\text{VIX}) - 14.6917\beta(\text{PMI})$$

$$\ln\left(\frac{\pi(\text{high})}{\pi(\text{factors})}\right)$$

$$= -2.2726 + 0.0270\beta(\text{VIX}) - 9.7563\beta(\text{PMI})$$

$$\ln\left(\frac{\pi(\text{very high})}{\pi(\text{factors})}\right)$$

$$= -0.6261 + 0.0341\beta(\text{VIX}) - 14.2607\beta(\text{PMI})$$

**T-Test**

| Risk premium | Very low | Low | High | Very high |
|---|---|---|---|---|
| Intercept | -4.799679045 | -4.054963956 | -0.555844267 | 1.151889777 |
| VIX | 3.347860061 | 2.888325077 | 1.137138309 | -1.210640969 |
| PMI | -2.182461983 | -2.298989371 | -2.070732834 | -2.76775613 |

By the T test regression, intercept and VIX will not be significant when risk premium is high and very high. By taking away the insignificant factors, we derive at the following model:

$$\ln\left(\frac{\pi(\text{very low})}{\pi(\text{factors})}\right)$$

$$= -5.9276 + 0.1413\beta(\text{VIX}) - 22.0309\beta(\text{PMI})$$

$$\ln\left(\frac{\pi\ \text{low})}{\pi(\text{factors})}\right)$$

$$= -2.7212 + 0.0836\beta(\text{VIX}) - 14.6917\beta(\text{PMI})$$

$$\ln\left(\frac{\pi(\text{high})}{\pi(\text{factors})}\right)$$

$$= -9.7563\beta(\text{PMI})$$

$$\ln\left(\frac{\pi(\text{very high})}{\pi(\text{factors})}\right)$$

$$= -14.2607\beta(\text{PMI})$$

# 9 Modified model

From the modified regression models using the large of treasury yield and gold return as the risk-free rate to calculate the risk premium, it was found that the logistic regression formulas are as follow.

Remaining all coefficients:

$$\ln\left(\frac{\pi(\text{very low})}{\pi(\text{factors})}\right)$$

$$= -6.7939 + 0.1644\beta(\text{VIX}) - 22.3.99\beta(\text{PMI}) - 88.8645\beta(\text{INDUSTPRO})$$

$$\ln\left(\frac{\pi\ \text{low})}{\pi(\text{factors})}\right)$$

= -2.4247 + 0.0684β(VIX) –20.4192β(PMI) – 74.0667β(INDUSTPRO)

$$\ln\left(\frac{\pi(\text{high})}{\pi(\text{factors})}\right)$$

= 0.0028 + 0.0141β(VIX) –18.3863β(PMI) – 46.0584β(INDUSTPRO)

$$\ln\left(\frac{\pi(\text{very high})}{\pi(\text{factors})}\right)$$

= 1.0631 -0.0555β(VIX) – 16.4492β(PMI) – 12.2279β(INDUSTPRO)

## 10 T test

| | Very low | Low | High | Very high |
|---|---|---|---|---|
| intercept | 5.734098153 | -3.186610329 | 0.004569239 | 1.538421398 |
| VIX | 3.94445389 | 2.000172114 | 0.465572465 | -1.517891181 |
| PMI | -2.469749752 | -2.944803539 | -3.095733403 | -2.527187498 |
| INDUSTPRO | -4.408104269 | -4.002611774 | -2.580310519 | -0.615873076 |

From the T-test for this regression, it reflects that VIX and intercept, when the risk premium is high and very high, will not be significant that it is not as reliable as when the risk premium is low and very low. Also, the industrial production should be rejected when the risk premium is very high. Therefore, we come up with the following regression model that the insignificant factors are taken away from the previous model.

Without insignificant coefficient:

$$\ln\left(\frac{\pi(\text{very low})}{\pi(\text{factors})}\right)$$

= -6.7939 + 0.1644β(VIX) – 22.3.99β(PMI) – 88.8645β(INDUSTPRO)

$$\ln\left(\frac{\pi\ \text{low})}{\pi(\text{factors})}\right)$$

= -2.4247 + 0.0684β(VIX) –20.4192β(PMI) – 74.0667β(INDUSTPRO)

$$\ln\left(\frac{\pi(\text{high})}{\pi(\text{factors})}\right)$$

= –18.3863β(PMI) – 46.0584β(INDUSTPRO)

$$\ln\left(\frac{\pi(\text{very high})}{\pi(\text{factors})}\right)$$

$=-16.4492\beta(\text{PMI})$

## 11 Test on fitness of the 3 Factor model with the use of classic CAPM

The expected value of equity premium calculated by our model is also the market premium. With this in mind, it is put into the CAPM model (rstock = rf + betastock (market premium)) to calculate the expected return on stocks. Specifically, two stocks, Bank of America and General Electric are used to test the robustness of our model. After expected returns on stocks are calculated, we have expected returns of stocks and actual returns of the two stocks. Actual returns are regressed on expected returns. The regression model used is a linear regression model.

Result summary:

| Stock stat | Bank of America (Gold) | Bank of America (Gold)* | Bank of America | Bank of America* | General Electric (Gold) | General Electric (Gold)* | General Electric | General Electric* |
|---|---|---|---|---|---|---|---|---|
| Slope coefficient | 1.2806 | 1.0212 | 0.48853 | 0.53064 | 0.92543 | 0.55166 | 0.47207 | 0.50835 |
| T-stat | 4.5935 | 3.4696 | 1.879 | 2.088 | 3.0955 | 1.754 | 1.7259 | 1.9008 |
| R-squared | 0.0735 | 0.0433 | 0.0131 | 0.0161 | 0.0348 | 0.0114 | 0.0111 | 0.0134 |
| Significant (a=5%) | Yes | Yes | No | Yes | Yes | No | No | No |

(Gold): Model with gold return considered

Models with gold return considered for the proxy of risk-free rate have significant slope coefficients mostly. Among them, models that have no special treatment for insignificant x factors are taken as 0 have significant slope coefficients for both

stocks.

For the two models under model with consideration for gold return and no special treatment, Bank of America has the slope of 1.28, while General Electric 0.93, which are pretty close to 1. The slope coefficient value of 1 means our predicted and actual returns increase by the same amount, which would, in turn, prove our 3-factor model with consideration for gold return is robust in calculating risk premium.

## 12 Compare average risk premium of both models

In the model with treasury yield as risk-free rate, the average risk premium is calculated as 2.27% per month, the existence of equity premium puzzle. However, in the model with larger of treasury yield and gold return as risk-free rate, the average risk premium is calculated as 0.36%. The result shows that an equity premium may not exist when concerning the larger of treasury yield and gold return as risk free rate.

## 13 Further Developments and conclusion

In order to enhance the significance of our homemade model, the first way is considering data with longer periods of time for more than 50 years.

With our model, equity premium puzzle could be explained away by using a new proxy for risk free rate, which instead of only considering the government bond yield, return on gold is also considered. More specifically, the larger of government bond yield or return on gold is taken as the risk-free rate in our model.

## References

[1] Wikipedia, "Equity Premium Puzzle", 30 May 2015. http://en.wikipedia.org/wiki/Equity_premium_puzzle
[2] Wikipedia, "Industrial production Index", 2 August 2013 http://en.wikipedia.org/wiki/Industrial_production_index
[3] Economic Research, "CBOE Volatility Index: VIX" https://research.stlouisfed.org/fred2/series/VIXCLS/downloaddata
[4] Andrew B. Abel, "The Equity Premium Puzzle", Business Review – Federal Reserve Bank of Philadelpha, 19
[5] Campbell, John Y. and Cochrane, John H, 1995, "By force of habit: a consumption-based explanation of aggregate stock market behavior," NBER

Working Paper 4995, National Bureau of Economic Research, Cambridge, MA.

[6] Daniel Mostovoy, "The Equity Market Premium Puzzle – CAPM and Minimum Variance Portfolios", Northfield Seminar, April 24, 2008.

[7] Campbell, John Y. and Cochrane, John H, 1995, "By force of habit: a consumption-based explanation of aggregate stock market behavior," NBER Working Paper 4995, National Bureau of Economic Research, Cambridge, MA.

[8] John E. Parsons, "The Equity Risk Premium and the Cost of Capital", CEEPR Workshop Cambridge, MA, Center for Energy and Environmental Policy Research, May 2006.

[9] Mehra, Rajnish, 2003. "The Equity Premium: Why Is It a Puzzle?", Financial Analysts Journal, vol. 59, no. 1 (January/February): 54-69.

[10] Mehra, Rajnish and Prescott, Edward C., 1985. "The Equity Premium: A Puzzle.", Journal of Monetary Economics, vol. 15, no. 1 (March): 145-162.

[11] Mehra, Rajnish and Prescott, Edward C. 2003. "The Equity Premium in Retrospect," NBER Working Papers 9525, National Bureau of Economic Research, Inc.

[12] Pable Fernadez, Javier Aguirremalloa, Heinrich Liechtenstein, "The Equity Premium Puzzle: High Required Equity Premium, Undervaluation and Self Fulfilling Prophecy", IESE Business School – University of Narvarra, 2009.

[13] Siegel, Jeremy J., 1992. "The Equity Premium: Stock and Bond Returns Since 1802", Financial Analyst Journal, vol. 48, no. 1 (January/February): 28-39.