A VAR Analysis of the Causal Relations among
Stock Returns, Inflation and Real Activities

Student Name: Jiayi Hong
Student ID: 1151833
1. Introduction

Today, stock investment has become a very popular method for investors. With the high inflation rate today, it is important for those investors to consider how the market reacts to inflation. Fortunately, the predictability of the return on the stock market has been profoundly discussed for a long time and the relations among inflation, real activities and stock prices have been investigated by many economists. For example, in 1981, American economist Eugene F. Fama published an article ‘stock returns, real activity, inflation, and money’ on ‘The American Economic Review’ to explain the anomalous stock return- inflation relations, he demonstrated that there is a negative relation between the inflation rate and the stock return since the stock return is more fundamental determinants of equity values. This paper is attempt to test the negative relation hypothesis among inflation, real activities and the stock return by using the vector- autoregresstion (VAR) approach and then compared the results of this approach to the random walk approach hypothesis.

The structure of this paper is: in section 2, some of the famous existing literature of the topic related area will be introduced in order to explain and reinforce the main idea of this paper; then in sections 3, it is a data and sample description section, a full description of the data that used in this paper will be presented; section 4 will describe the methodology which will be used to the
data analysis; afterwards, Section 5 will present the main result of the previous analysis and a critical discussion will also be made based on the result; continually, the last section will make a conclusion of this paper will some suggestions.

2. Literature Review

There are massive evidences shows that since World War II, stock return have done relatively poor when inflation occurs. In addition, the returns of growth stocks have been more negatively correlated with unexpected inflation. The Economist like Fama (1981), Mark Skousen(2006),Ram and Spencer(1983), and James, Koreisha, and Partch(1985) all attempt to demonstrate the negative association between stock returns and inflation rate. Among these, the theoretical supports behind them are mainly based on two point of views. First, “if inflation is a growing problem, investment analysts become suspicious of high economic growth or good job reports, because they fear that it reflects an inflationary boom, an artificial recovery generated by “easy credit” by the government. Thus under inflationary conditions, analysts do not think strong job creation and economic growth are sustainable, and stock market falls in price because they think that Fed will need to tighten in the future” (Mark Skousen, 2006). Second, it is the famous proxy hypothesis raised by Fama(1981), he states that high inflation rates anticipate low growth rate of
aggregate economic activity: “as economic activity growth is expected to slow down, the growth rate of the demand for real cash balances is also expected to decrease, leading to an increase in inflation. In addition, high stock returns anticipate high growth rates of aggregate economic activity. Thus, inflation and stock returns are driven in opposite directions by anticipated business fluctuations, and correlate negatively.”

Although the estimates of the above causality model seem reasonable, their methodology does not have the power to identify the accurate direction of causality between stock returns and changes in inflationary expectations. In 1985, James, Koreisha and Partch used the vector autoregressive moving average (VARMA) model to investigate the relations among the stock returns, real activity, the money supply, and inflation. As a result, they found the evidence of linkages between these variables.

Besides, there still some economists hold a totally different opinion towards it. Like the random walk hypothesis Jules Regnault made on 1863. He stated that stock market prices evolve according to a random walk and thus the prices of the stock market cannot be predicted.

The aim of this paper is to use the data set capturing the post-war stock market returns and other useful economic factors to test these famous hypotheses.
above by using a multivariate vector-autoregression (VAR) approach and
investigate the linkages between the stock returns, inflation and real activities.

3. Data

The data we used to examine the causal relations between stock returns and
inflation are come from Jeffrey Wooldridge’s textbook ‘Introductory
Econometrics: A Modern Approach.’

First, as a measure of common stock returns, the monthly return on Standard
& Poor’s 500 index starting from January 1947 to June 1993 will be used. This
variable is calculated as percentage change of sp500 plus the dividend yield
rate. As a proxy for the change in expected inflation, the change of the three
month Treasury bill rate (ci3) will be used, which is calculated as i3-i3[n-1].
And the real activities will be measured by the variable ‘ip’ which represent
the index of industrial production. Finally the variable ‘pcip’ means the growth
rate of the index of industrial production. To be notice that, all the growth rate
in this paper will be calculated as annul rate since the monthly changes in
industrial production are highly fluctuated by seasonal reasons.

Similar to James, Koreisha and Partch’s paper in 1985, the inflation rate is
represented by the interest rate in this paper. This is also account with
Fisher’s point of view. In Fisher’s book ‘The Theory of Interest’ he found the
evidence that the current interest rate is influenced by the past inflation. It infers the last period’s interest rate can also indicate the expected inflation rate. This is the reason why the difference of interest rates in current and last period will be taken as the change in the expected inflation.

First, we use ordinary least-squares (OLS) to estimate the relation between the return on stock (‘rsp500’) and both expected inflation (‘i3’) and changes in expected inflation (‘ci3’). The result is presented as below:

\[
\text{Rsp500} = 18.5066 - 1.25778\text{i3} - 11.6027\text{ci3}
\]

\[
(3.1573) \quad (0.5330) \quad (3.4352)
\]

\[R^2 = 0.0317 \quad F = 9.07\]

(See Appendix 1 Table 1)

It is obvious that there is a negative relation between stock returns and both expected inflation and changes in expected inflation, and this is also statistically significant. The result above is in line with the results from Fama and Schwert(1981) and Geske and Roll(1983).

Then, we estimated the relation between the return on stock (‘rsp500’), expected inflation (‘i3’), growth in real activities (‘pcip’), and changes in expected inflation (‘ci3’). The result is presented as below:
The results above are similar to those found by Fama, that is there is no significant relation between expected inflation and returns on stock when growth in real activities variable is added to the model, but, the negative relations between changes in expected inflation and stock returns is still significant. Sum up, the phenomena that Fama and Geske and Roll try to demonstrate are also found in our sample.

It should be noticed that there are also drawbacks in the data. First, the data is from 1947 to 1993 as mentioned before. In the beginning of this period, the Treasury bills rate might not be a proper proxy for the expected inflation rate since the Treasury bills rate had been pegged by the government before 1951. Second, using the change in Treasury bill returns as a measure of expected inflation ignores the possibility that a change in the real interest rate will cause the change in nominal interest rate. However, Fama and Gibbons find that “real rate changes follow a random walk and are a small component of changes in Treasury bill returns. Moreover, since the reversed causality model proposed by Geske and Roll is formulated in terms of stock returns.
and nominal interest rate changes” (Geske, 1983), we still use Treasury bill returns for the analysis.

4. Methodology

First, an Augmented Dickey-Fuller test will be used to test the null hypothesis that the there is a unit root in ‘rsp500’, ‘i3’ and ‘ip’. Since if there is a unit root, the level will not predict the direction of the next step.

A VAR is a model in which K variables are specified as a linear function of p of their own lags, p lags of the other K-1 variables, and the additional exogenous variables. Thus, a VAR process could be written as:

\[ \beta(L)Y_t = \varepsilon_t \]

\( \beta(L) \) is a \( k \times n \) matrix with coefficients where \( \beta(L)=I-\beta_1L-\beta_2L^2-\beta_3L^3-\ldots-\beta_kL^k \) for K-order VAR; \( Y_t \) is a \( n \times 1 \) matrix which contains \( n \) endogenous variables; and \( \varepsilon_t \) is a \( K \times 1 \) vector of random shocks which are independently, identically and normally distributed with mean zero and covariance \( \Sigma \). In this paper, \( Y_t \) includes monthly stock market returns (‘rsp500’), changes in the expected inflation rate (‘ci3’) and real activities (‘pcip’).
In the time series model, serial correlation is always a big issue that frequently observed. That is the error terms $\epsilon_t$ and $\epsilon_s$, for $t \neq s$, are correlated. With time series issue in the model, OLS estimator will have low power and the efficiency of the test will also be affected. To test for autocorrelation in the residuals of VAR models, a LM (langrange-multiplier) test will be used. The formula of this test shows at below:

$$LMS = (T - d - 0.5) \ln \left( \frac{|\Sigma|}{|\Sigma_s|} \right)$$

In the formula, $T$ is the number of observations; $d$ is the number of parameters estimated in the augmented VAR; $|\Sigma|$ is the maximum likelihood estimate of the variance-covariance matrix of the residuals from the augmented VAR. In our model, there are three equations in the VAR. The augmented VAR includes s lags of three new variables in the original VAR. The asymptotic distribution of LM is chi-square with $K^2$ degrees of freedom, and $K$ is 3 in our case.

The utilization of a VAR process is also an appropriate way to investigate the dynamic relationships between variables. And the impulse response functions (IRFs) are a useful tool for characterizing the dynamic responses implied by the estimated VARs such as the persistence of the effects of shocks to one variable on another variable.
The first-order VAR for the n-vector $y_t$ is:

$$y_t = \mu + Ay_{t-1} + \varepsilon_t.$$ \hspace{1cm} (1)

where $\mu$ is the vector of intercepts to allow the means of $y_t$ to be non-zero.

The IRF of a shock to variable $j$ on variables $l$ after $k$ periods is:

$$IRF(i, j, k) = E[y_{i,t} \mid \varepsilon_{j,t-k} = 1, \varepsilon_{\setminus j,t-k} = 0, \varepsilon_{t-k+1} = 0, \ldots, \varepsilon_t = 0] - E[y_{i,t} \mid \varepsilon_{j,t-k} = 0, \varepsilon_{\setminus j,t-k} = 0, \varepsilon_{t-k+1} = 0, \ldots, \varepsilon_t = 0].$$ \hspace{1cm} (2)

By backward substitution on (1)

$$y_t = \mu + Ay_{t-1} + \varepsilon_t$$

$$= \mu + A\mu + A^2y_{t-2} + \varepsilon_t + A\varepsilon_{t-1}$$

$$= \sum_{s=0}^{t-1} A^s\mu + A^t y_0 + \sum_{s=0}^{t-1} A^s \varepsilon_{t-s}$$ \hspace{1cm} (3)

$$E[y_t \mid \varepsilon_{j,t-k} = 1, \varepsilon_{\setminus j,t-k} = 0, \varepsilon_{t-k+1} = 0, \ldots, \varepsilon_t = 0]$$

$$= \sum_{s=0}^{t-1} A^s\mu + A^t y_0 + \sum_{s=k+1}^{t-1} C_s \varepsilon_{t-s} + C_k \varepsilon_{t-k}$$

$$E[y_t \mid \varepsilon_{j,t-k} = 0, \varepsilon_{\setminus j,t-k} = 0, \varepsilon_{t-k+1} = 0, \ldots, \varepsilon_t = 0]$$

$$= \sum_{s=0}^{t-1} A^s\mu + A^t y_0 + \sum_{s=k+1}^{t-1} C_s \varepsilon_{t-s}$$ \hspace{1cm} (4)

Then, substitute (4) into (2):

$$IRF = A^k \varepsilon_{t-k}$$ \hspace{1cm} (5)
$\epsilon_{t,k}$ is a vector of zeros but with 1 as the $j^{th}$ element. $A^k \cdot \epsilon_{t,k}$ is the $j^{th}$ column of $A_k$.

IRFs have certain properties for certain models. For the linear model, the response is linear in the size of the shock and the response to positive and negative shocks is symmetric. However, it is criticized because the IRF considers a shock to $\epsilon_i$ is isolated, but when $\epsilon_t \sim \text{IN}(0, \Sigma)$, unless $\Sigma$ is diagonal, the shocks are correlated. But in reality, $\Sigma$ is seldom to be diagonal. So historically, a shock to $\epsilon_i$ would on average be associated with shocks of given magnitudes and sign to other residuals. And this problem is talked by orthogonalizing the innovations. Since the covariance matrix $\Sigma$ is positive definite. Thus there exists a non-singular lower triangular matrix, $P$, such that $P \Sigma P' = I$, so that defining $\eta_t = P \epsilon_t$, where $\eta_t$ are the orthogonalized innovations, thus $E[\eta_t \eta_t'] = E[P \epsilon_t \epsilon_t' P'] = I$. This is called Choleski decomposition. Since the decomposition method needs further investigation, in this paper, the shock is just consider in isolation.

This paper has two main aims. One is to test whether stock returns, changes in expected inflation, and real activities in the past and current period can forecast each other when their lags are included in the model. Another aim is
to find out whether using VAR approach can get a more accurate forecast than using Random Walk approach in explaining the stock market returns. The out-of-sample forecast approach can be used to predict the returns on the stock market. Because the amount of observations needs to large enough, a way for the purpose of ensuring that could be to have the first 400 observations as our sample running a VAR process, and subsequently initiating the next period’s prediction about return, in addition, the recursive way will also be used. That means, each time when predict a value, the original 400 values will be used together to run the VAR to forecast the 402nd one until the sample be run out. The forecast errors can be obtained by comparing the predicted values with original observations, and its mean squared value is then calculated too. The way of getting the mean squared forecast error is also appropriate for the random walk approach. Theil’s U statistics will be used to compare the forecast accuracy for these two approaches. Theil’s $\text{stat} = \frac{\text{RMSE(VAR)}}{\text{RMSE(RW)}}$, where RMSE is the root of the MSFE. The U-stat result will be compare with 1, if it is smaller than one, it means VAR approach provide a better forecasting ability.

5. **Empirical results**

As the data section shows, by using OLS, the simple regression on stock returns and change in expected inflation as well as the regression on stock
returns, inflation rate, growth in real activities, and expected change in inflation rate both clarify a significant negative relations between stock returns and inflation rate.

Then, according to the ADF test results, the null hypothesis of there is a unit root in ‘rsp500’ has been rejected. However, in the test for ‘i3’ and ‘ip’, the null hypothesis has failed to reject. Thus, the first difference of ‘i3’and ‘ip’ will be used as proxies for changes in expected inflation and growth in real activities. By retesting the ‘ci3’ and ‘pcip’, there are no unit roots in these variables. Hence, the establishment of the three-variable system VAR approach may be carried out, and with respect to our variables of interest, there is no requirement for co-integration.(See Appendix 2)

By taking the Lagrange-multiplier test, we also reject the null hypothesis that there is no autocorrelation in the residuals. Therefore, the hypothesis of the disturbances is not autocorrelated is true, and model misspecification does not appear to be the case (See Appendix 3).

With accordance to the impulse response function graph (See Graph 1 below). It is obvious that a significant response movement on stock returns will be induced by an innovation in changes in expected inflation just like what it does to real activities.
The 3rd picture in first line shows an innovation shows the IRF of an innovation in ‘ci3’ on ‘rsp500’. It shows that stock returns would be affected by expected inflation, and if a positive shock on it appears, a reduction in stock return can be foreseen. There is also a period of fluctuation associated with this reduction, and may last till approximately 8 periods later. In a certain extent, this phenomenon tells that the relation between stock returns and expected inflation is negative. However, the 2nd the 3rd lines show that there seems no obvious changes for other variables on stock return.

(Graph 1)
At last, the MSFE results are 1804.147 and 2699.876 respectively by using VAR approach and Random Walk Approach. Thus the Theil’s sstat is:

\[
\text{Theil's sstat} = \frac{\text{RMSE(VAR)}}{\text{RMSE(RW)}} = \sqrt{\frac{1804.147}{2699.876}} = 0.817
\]

The result is less than 1. This is the evidence of the VAR approach provides a better forecasting ability.

6. Conclusions

To conclude, the causal relations among stock market returns, change in expected inflation and real activities is observed using a VAR approach established based on data from 1947 to 1993. Annualized rate was applied to all the relative data. A negative relation between stock returns and expected inflation is discovered as a significant finding, and this matches what Fama (1981) suggested. Besides, the VAR approach also provides a refined explanation on stock returns than the Random Walk approach does. This phenomenon is impressive; however it is still a hypothesis which should be proved further with alternative methods. Our model does have some constraints, and the one to be mentioned here is the ignorance of the economic events during the large horizon period, such as government policies, natural disasters. It is recommended that further work should be carried out with smaller interval between returns because the stock market is turbulent,
daily returns should be applied instead of monthly data cause it may have even smaller gap.
Reference


### Appendix

(Appendix 1)

**. reg resp500 i3 ci3**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 557</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F( 2, 554) = 9.07</td>
</tr>
<tr>
<td>Model</td>
<td>28612.8864</td>
<td>2</td>
<td>14306.4432</td>
<td>Prob &gt; F = 0.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>874287.803</td>
<td>554</td>
<td>1578.13683</td>
<td>R-squared = 0.0317</td>
</tr>
<tr>
<td>Total</td>
<td>902900.69</td>
<td>556</td>
<td>1623.9221</td>
<td>Adj R-squared = 0.0282</td>
</tr>
</tbody>
</table>

|       | Coef.  | Std. Err. | t     | P> |t| [95% Conf. Interval] |
|-------|--------|-----------|-------|----|----------------------|
| i3    | -1.257782 | .5330316 | -2.36  | 0.019 | -2.304792 | -2.2107715 |
| ci3   | -11.60267 | 3.435248 | -3.38  | 0.001 | -18.35037 | -4.854965  |
| _cons | 18.50695 | 3.15726  | 5.86   | 0.000 | 12.30489 | 24.70821   |

**. reg resp500 i3 pcp ci3**

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 557</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F( 3, 553) = 6.42</td>
</tr>
<tr>
<td>Model</td>
<td>30394.7139</td>
<td>3</td>
<td>10131.5713</td>
<td>Prob &gt; F = 0.0003</td>
</tr>
<tr>
<td>Residual</td>
<td>872505.976</td>
<td>553</td>
<td>1577.76849</td>
<td>R-squared = 0.0337</td>
</tr>
<tr>
<td>Total</td>
<td>902900.69</td>
<td>556</td>
<td>1623.9221</td>
<td>Adj R-squared = 0.0284</td>
</tr>
</tbody>
</table>

|       | Coef.  | Std. Err. | t     | P>|t| | [95% Conf. Interval] |
|-------|--------|-----------|-------|-----|----------------------|
| i3    | -1.182881 | .5376095 | -2.20  | 0.028 | -2.238888 | -1.1268748 |
| pcp   | .1396101 | .1313729 | 1.06   | 0.288 | -.1184408 | .397661   |
| ci3   | -12.43588 | 3.523195 | -3.53  | 0.000 | -19.35636 | -5.515395  |
| _cons | 17.84533 | 3.259239 | 5.41   | 0.000 | 11.24338 | 24.04739   |
(Appendix 2)

```
dfuller rsp500, lags(12)

Augmented Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-6.624</td>
<td>-3.430</td>
<td>-2.860</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.0000

.dfuller i3, lags(12)

Augmented Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-1.918</td>
<td>-3.430</td>
<td>-2.860</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.3238

.dfuller ip, lags(12)

Augmented Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>0.045</td>
<td>-3.430</td>
<td>-2.860</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.9621

.dfuller pcip, lags(12)

Augmented Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-7.648</td>
<td>-3.430</td>
<td>-2.860</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.0000
```
. dfuller c13, lags(12)

Augmented Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-6.558</td>
<td>-3.430</td>
<td>-2.860</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.0000
. varlmar

Lagrange-multiplier test

<table>
<thead>
<tr>
<th>lag</th>
<th>chi2</th>
<th>df</th>
<th>Prob &gt; chi2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.5210</td>
<td>9</td>
<td>0.87391</td>
</tr>
<tr>
<td>2</td>
<td>11.2002</td>
<td>9</td>
<td>0.26223</td>
</tr>
</tbody>
</table>

H0: no autocorrelation at lag order