# Forecasting the price of gold: An error correction approach

Gangopadhyay, Kausik and Jangir, Abhishek and Sensarma, Rudra

February 2015

# Forecasting the price of gold: An error correction approach

**Short title: Forecasting the gold price**

Kausik Gangopadhyay[1]
Indian Institute of Management Kozhikode
IIMK Campus P.O., Kozhikode, India 673570
Phone: 91-495-2809118; Fax: 91-495-2803010
E-mail: kausik@iimk.ac.in

Abhishek Jangir
NAV Capital LLP,
A-402, Phase 4, Lake Homes,
Chandivali, Powai,
Mumbai, India  400076
Phone: 91-22-28581911
E-mail: abhijangir@gmail.com

Rudra Sensarma
Indian Institute of Management Kozhikode
IIMK Campus P.O., Kozhikode, India 673570
Phone: 91-495-2809118; Fax: 91-495-2803010
E-mail: rsensarma@iimk.ac.in

---

[1] Corresponding Author

# Forecasting the price of gold: An error correction approach

**Short title: Forecasting the gold price**

**ABSTRACT**

Gold prices in Indian market may be influenced by a multitude of factors such as investment decision, inflation hedge and consumption motives. Gold prices are modelled using a vector error correction model. We identify investment decision and inflation hedge as prime movers of the data. We also present out-of-sample forecasts of our model and the related properties.

**Keywords:** Gold price; cointegration; vector error correction model; inflation hedge.

## EXTENDED SUMMARY

India is one of the major gold consuming countries and demand from India is acknowledged to be a major factor in determining international gold prices. This paper is the first of its kind to develop a model for explaining and forecasting gold prices in India.

Gold is regarded as a good hedge both against inflation as well as against fall in value of other assets. Some of the factors that influence gold prices include inflation, exchange rate, bond prices, market performance, seasonality, income, oil prices, and business cycles. While these factors are expected to work in India as in other countries, there is an additional role of gold that may not be relevant elsewhere and has been hitherto ignored in this literature. Indians buy gold not just for investments but also for personal reasons to be used as a luxury. If this reason to buy gold is significant, then higher affordability should lead to increased demand and therefore higher price for gold. We capture this wealth effect through the stock market index.

Our time series (April 1990 to August 2013) variables are non-stationary. We use a cointegrating framework that is vector error correction approach to model and forecast the price of gold. The stock market index has a negative relationship with gold price contradicting the argument for gold being a luxury good but supporting the role of gold as a hedge. Oil price and exchange rate have negative relationships with gold price implying that gold is a good hedge against oil and the dollar as an investment. The CPI has a positive relationship with gold indicating that gold is a good inflation hedge.

We affirm robustness of the results of our exercise. The predictive capacity of our error correction model beats alternative specifications such as the random walk using different sub-periods and forecasting horizons.

# INTRODUCTION

India is one of the major gold consuming countries and high import demand from India is often cited as the primary reason for the persistent current account deficit. This high demand from India is acknowledged to be a major factor in determining international gold prices. However not much is known about what determines gold prices in India. Understanding the determinants of gold price will help in developing a predictive model for forecasting future prices. This can be useful for the purpose of portfolio decision of investors as well as a critical input for policy making.

This paper is the first of its kind to develop a model for explaining and forecasting gold prices in India. We estimate the nature of relationship of gold price in India with key determinants such as the stock market index, oil prices, exchange rate, interest rate and consumer price index. We find that gold is useful as a portfolio hedge as well as a hedge against inflation. Our model is able to predict future gold prices with reasonable levels of accuracy.

# 1 RESEARCH BACKGROUND

Among all precious metals, gold might be the most popular investment. It has stood the test of time, performed well during crisis situations like market decline, currency failure, inflation, war etc. It is regarded as a good hedge both against inflation as well as against fall in value of other assets. Gold's usefulness as an inflation hedge would imply that when general prices are high, gold prices will also be high so that the asset can be sold in order to finance general spending activity. However gold's role as a hedge against other assets (such as stocks, bonds, foreign currency) would mean that when the prices of other assets fall, price of gold rises such that the resulting portfolio is diversified.

Many studies have looked into pattern of gold prices (see e.g. Capie, Mills and Wood, 2005; Worthington and Pahlavani, 2007; Baur and Lucey, 2010) to identify the factors that influence gold prices. Some of the factors that influence gold prices include inflation, exchange rate, bond prices, market performance, seasonality, income, oil prices, and business cycles. However there is no work that has been done to examine gold prices in India.

We carry out an analysis to study the factors influencing gold prices in India by collecting monthly data on gold prices and other factors over a long time period. While the hedge factors are expected to work in India as in other countries, there is an additional role of gold that may not be relevant elsewhere and has been hitherto ignored in this literature. Indians buy gold not just for investments but also for personal reasons to be used as a luxury good (to wear as jewellery, to gift in weddings, for religious reasons etc.). If this reason to buy gold is significant, then higher affordability should lead to increased demand and therefore higher price for gold. We capture the wealth effect through the stock market index.

The time series variables that we study are, largely, non-stationary variables. Therefore, we need to analyse them in a cointegrating framework. We use a vector error correction

approach to model and forecast the price of gold. Our benchmark estimates are for the period from April 1990 to August 2013.

We find that gold price has a cointegrating relationship with the stock market index, exchange rate, CPI, US bond rates and oil price. The stock market index has a negative relationship with gold price contradicting the argument for gold being a luxury good but supporting the role of gold as a hedge. This is consistent with Baur and McDermott (2010). The exchange rate has a negative relationship with gold price implying that a stronger rupee is associated with costlier gold. Our finding demonstrates that gold is a good hedge against dollar from the point of view of domestic investors which is also the case for developed countries (Reboredo, 2013).

Oil price has a negative relationship implying that gold is a good hedge against oil as an investment, in contrast to existing evidences from developed countries (see for example, Reboredo, 2013). The CPI has a positive relationship with gold indicating that gold is a good inflation hedge, a result that has been previously obtained for developed countries (Worthington and Pahlavani, 2007; Ghosh et al., 2004). Finally US bond rates are negatively related to gold price, indicating that when returns from international investments fall, investors may switch to gold.

We tested for robustness of the results of our exercise. We have taken some transformations of the variables, e.g. logarithmic. We have added difference polynomials of independent variables. Our findings are quite robust to these alternative specifications. The relationship established by us provides interesting insights into the role of gold in portfolio diversification and as a hedge against inflation in the Indian context. The predictive capacity of our error correction model beats alternative specifications such as the random walk using different sub-periods and forecasting horizons.

## 2 DATA AND METHODOLOGY

### 2.1 Data Source

The gold price data is obtained from the Reserve Bank of India's website. It is taken in real terms by deflating it using the consumer price index (CPI). The CPI data we use is for Urban Non-Manual Employees and later Industrial Workers maintained by Labour Bureau, Government of India[2]. We have taken the equity market index Sensex as a proxy for the stock market. Whenever Sensex suffers a decline, the loss stricken investors may move towards gold which increases its demand which in turn increases its price. On the other hand if Sensex represents the wealth of the people, then a higher value of the Sensex may indicate that the purchasing power of people increases, so they may be able to afford more gold which increases its price. Sensex data is obtained from the website of the Bombay Stock Exchange[3].

When exchange rate increases, it makes gold imports more expensive leading to an increase in the domestic price of gold. The USD-INR exchange rate is collected from Indexmundi website[4] and Bloomberg. However if gold is a good hedge against the exchange rate then we would expect gold prices to be negatively related with the

---

[2] http://labourbureau.nic.in/indexes.htm
[33] http://www.bseindia.com
[44] http://www.indexmundi.com

exchange rate. This would mean that a fall in the dollar would induce investors would move towards gold thereby leading to higher gold prices.

When oil prices increase, then the cost of production increases which reduces the profits of investors who then switch to gold for safety. Therefore, oil prices are expected to have a positive effect on gold prices. But people don't buy oil simply as a factor of production, many trade it as a commodity for capital gains. So an increase in oil prices would be beneficial for such investors and they won't invest in gold. Again, this means a negative relationship between gold prices and oil prices. Oil prices are obtained from the Indexmundi website.

Indian bond market is still in its nascent stage. A select group of authorised domestic financial institutions are the only players in the secondary market and liquidity is limited across maturities. So the effect of the bond rates was not analysed, also because of lack of data which was available only from 2004 after the setting up of CCIL. However we used interest rates on US bonds to control for international investment prospects and the data is obtained from US treasury website. We use monthly data on the above variables between April 1990 and August 2013.

## 2.2 Data Transformations

We consider real price of gold (GLP) that is free from the influence of general price movements. We have normalised gold price by dividing the nominal value by the consumer price index. For example, if the gold price in a particular month is INR XXX and CPI is YYY, then GLP is calculated as INR XXX/YYY. The Bombay Stock Exchange Sensitive Index or Sensex (SNX) is taken in a logarithmic scale to represent stock prices. Other determinants of gold price that we consider are US Dollar- Indian Rupee exchange rate in logarithmic scale (denoted by EXR). We considered CPI in a logarithmic scale as well as the oil price (OIL). Finally we consider US bond rates (INT) to capture returns from international investments.

## 2.3 Pre-testing Time-series Properties of the Data

We start by performing unit root tests for all our time series. There are several specifications of a unit root process: Random walk, Random walk with drift, Random Walk with linear trend and drift etc. The cookbook procedure for carrying out unit root test is schematically shown in Figure 1. This procedure allows one not only to test for potential non-stationarity of the process but also to categorise the extent of the random walk process. An important practical issue for the implementation of the ADF test is the specification of the lag length p. If p is too small then the remaining serial correlation in the errors will bias the test. If p is too large then the power of the test will suffer.

For an optimal selection of lag length, we follow the procedure suggested by Ng and Perron (1995). We set an upper bound for p and estimate the ADF test regression. If the absolute value of the t-statistic for testing the significance of the last lagged difference is not significant then we decrease p by one and repeat the same process; else we stop at that p. Schwert (1989) suggested the rule of thumb for determining the maximum lag length which is the highest integer contained in $12 \times (T/100)^{0.25}$.
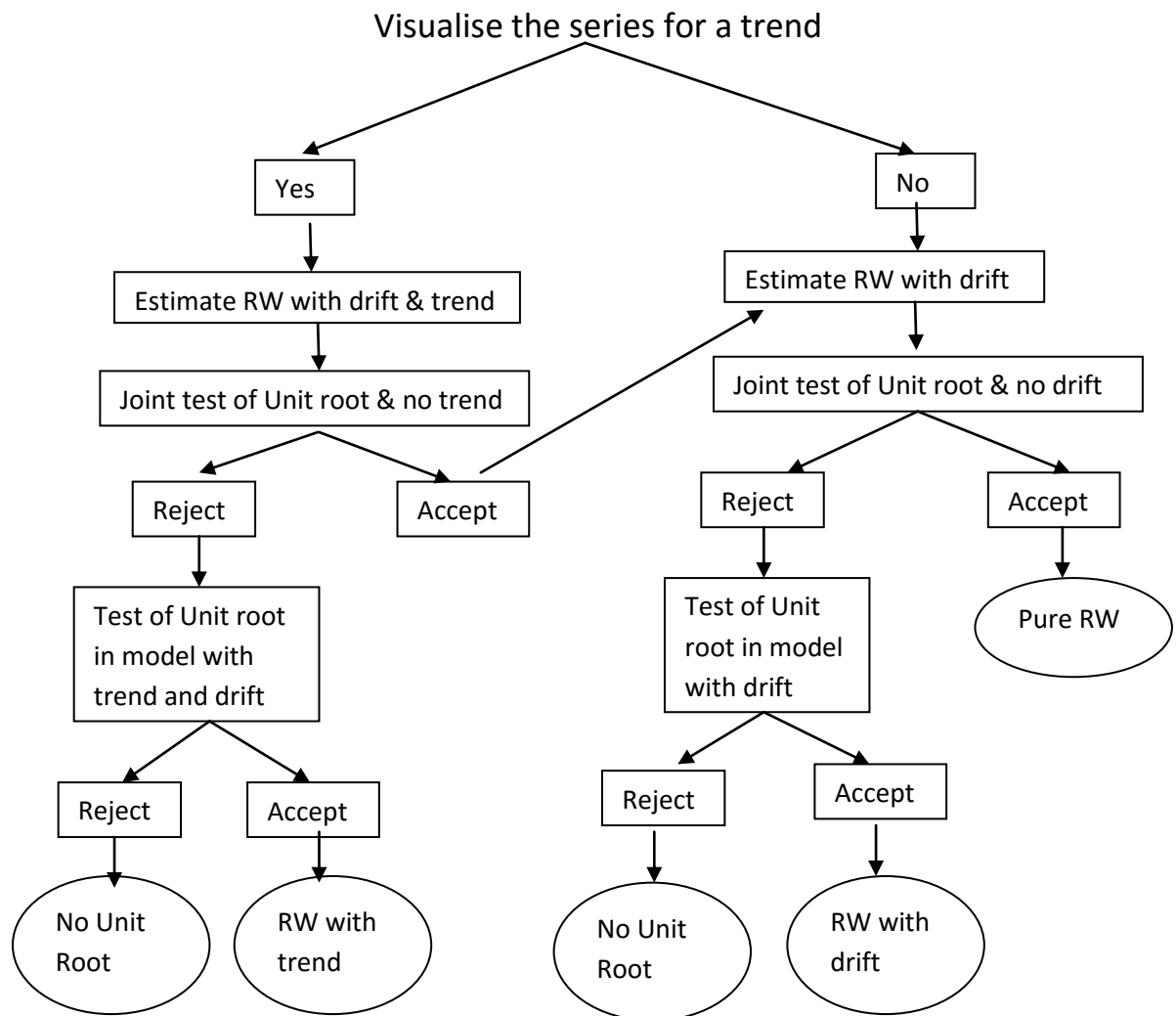
## Visualise the series for a trend

```
                    Visualise the series for a trend
                         /                    \
                      Yes                       No
                       |                         |
          Estimate RW with drift & trend   Estimate RW with drift
                       |                         |
          Joint test of Unit root & no trend  Joint test of Unit root & no drift
                    /        \                  /         \
                Reject      Accept          Reject       Accept
                  |                           |             |
          Test of Unit root          Test of Unit        Pure RW
          in model with              root in model
          trend and drift            with drift
               /    \                    /     \
          Reject   Accept           Reject    Accept
            |        |                |          |
         No Unit   RW with        No Unit     RW with
         Root      trend          Root        drift
```

*Figure 1. Schematic diagram for a cookbook procedure on testing for unit root*

Once we have found some variables as non-stationary, we can find the long run relationship between detection of cointegrating relationship. Johansen's method of Vector Error Correction Model (VECM) is appropriate in this connection.

## 3  ECONOMETRIC METHODOLOGY

### 3.1  Cointegration

A set of variables is cointegrated when there exists a stable long run relationship between them. While the original test of cointegration was provided by Engle and Granger (1987), due to the well-known deficiencies of this simple approach (Enders, 2004), we follow the approach subsequently provided by Johansen and Juselius (Johansen, 1988, 1991; Johansen and Juselius, 1990). Formally put, let $y_1$, $y_2$, ... ,$y_k$ be a set of variables which we are interested in. Suppose that each variable is integrated of order one, viz. I(1) and therefore needs differencing in order to attain stationarity. If

there exists linear combination(s) of the variables which is (are) I(0), then the variables are said to be cointegrated, i.e. they have a stable long run relationship. Then the cointegrating vector can be estimated which quantifies the relationship between the concerned variables.

## 3.2 Vector Error Correction Model

The Vector Error Correction Model (VECM) involves expressing an nx1 vector of stationary time series (say $y_t$) in terms of a constant, lagged values of itself and an error correction term. The standard VECM (p) model can be represented as,

$$\Delta y_t = c + \varphi_1 \Delta y_{t-1} + \varphi_2 \Delta y_{t-2} + \cdots + \varphi_p \Delta y_{t-p} + ECT_t + \varepsilon_t \qquad (1)$$

where ECT refers to the Error Correction Term that is a product of an adjustment factor (α) and the cointegrating vector (β). The cointegrating vector shows the long term equilibrium relationship between the concerned variables while the adjustment factors shows the speed of adjustment towards equilibrium in case there is any deviation.

## 4 FINDINGS

### 4.1 Summary of Data

Figure 2 plots all the relevant variables over time. The results of unit root tests are summarised in Table 1. The unit root tests clearly indicate that the relevant variables are integrated of order 1, which is indicative of most elementary degree of non-stationarity.
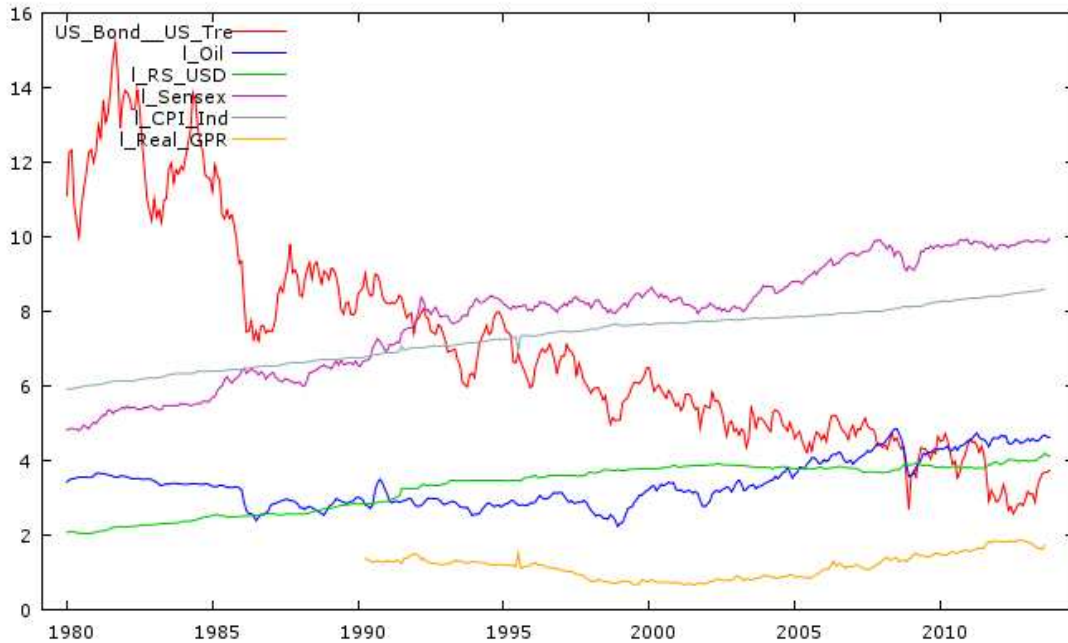


Figure 1. Variables considered in our analysis are plotted against time

*Table 1.* Unit Root Tests

| Variable | ADF Test | | | | KPSS test | | | Overall |
|---|---|---|---|---|---|---|---|---|
| | Lag | Trend | p-value | Verdict | Lag | Trend | Verdict | Conclusion |
| GLP | 1 | No | 0.693 | Accept | 17 | No | Reject at 1% | I(1) |
| ΔGLP | 0 | No | 0.000 | Reject | 17 | No | Accept at 1%; Reject at 5% | I(0) |
| SNX | 16 | No | 0.482 | Accept | 16 | No | Reject at 1% | I(1) |
| ΔSNX | 15 | No | 0.000 | Reject | 16 | No | Accept at 10% | I(0) |
| EXR | 7 | No | 0.843 | Accept | 7 | No | Reject at 1% | I(1) |
| ΔEXR | 1 | No | 0.000 | Reject | 1 | No | Accept at 10% | I(0) |
| INT | 15 | No | 0.739 | Accept | 15 | No | Reject at 1% | I(1) |
| ΔINT | 14 | No | 0.000 | Reject | 14 | No | Accept at 10% | I(0) |
| CPI | 15 | No | 0.861 | Accept | 15 | No | Reject at 1% | I(1) |
| ΔCPI | 14 | No | 0.000 | Reject | 14 | No | Accept at 10% | I(0) |
| OIL | 13 | No | 0.909 | Accept | 13 | No | Reject at 1% | I(1) |
| ΔOIL | 14 | No | 0.000 | Reject | 14 | No | Accept at 10% | I(0) |

## 4.2  Results from Modelling Gold Price

The results from the test of cointegration are reported in Table 2. The trace test suggests that ranks of 0 and 1 are rejected at 5% level of significance. However rank of 2 cannot be rejected at the 5% level. In other words we can conclude that the variables have one cointegrating relationship among them.

*Table 2.* Johansen test of cointegration

| Rank | Eigenvalue | Trace Test Statistics | Trace Test p-value | Lmax Test Statistics | Lmax Test p-value |
|---|---|---|---|---|---|
| 0 | 0.22104 | 148.49 | 0.0000 | 67.194 | 0.0000 |
| 1 | 0.12026 | 81.292 | 0.0039 | 34.468 | 0.0385 |
| 2 | 0.094388 | 46.825 | 0.0608 | 26.67 | 0.0626 |
| 3 | 0.051379 | 20.155 | 0.4233 | 14.189 | 0.3635 |

| | | | | | |
|---|---|---|---|---|---|
| 4 | 0.021655 | 5.9661 | 0.7024 | 5.8893 | 0.6330 |
| 5 | 0.00028544 | 0.076793 | 0.7817 | 0.076793 | 0.7817 |

Once the presence of cointegration is established we move to estimation of the cointegrating vectors and the VECM. Table 3 shows the cointegrating vector along with the standard errors of the estimates in parentheses. The coefficients suggest a relationship of the following nature:

$$GLP = -1.9362*SNX - 17.300*EXR - 1.4296*INT - 0.6073*OIL + 6.8974*CPI$$

(2)

The above equation indicates that gold prices and the stock market move in the opposite direction in the long run. Unlike what is expected of a luxury good, the wealth effect does not seem to dominate in the sense that higher wealth (captured by a rise in the Sensex) does not get reflected in increased demand and price of gold. However the role of gold as a hedge clearly dominates. As gold price moves in a different direction to that of the stock market it may be inferred that the role of gold as a portfolio hedge dominates its use as a luxury good in India.

This result may also imply that gold is a safe haven asset – an asset where investors can move into in times of high volatility. The difference between a hedge and a safe haven lies in the holding of the alternative asset under extreme market conditions. Under extreme market fluctuations investors tend to hold a safe haven asset whereas a hedge is a substitute for another asset when the latter is not performing as well.

Exchange rate is negatively related to gold prices in spite of the fact that bulk of the gold consumed in India is imported. Therefore a weaker exchange rate should translate into higher cost of imported gold which would also make domestic price higher. However our result seems to indicate that gold is a hedge against the dollar as well as could be a safe haven. US bond rates have a positive relationship with gold which suggests that when returns from investing outside the country are high, demand for gold in India may fall and therefore its price. Oil price has a positive relationship implying that gold may act as a good hedge against prices of commodities such as oil that are held by investors in their portfolio.

*Table 3.* Vector Error Correction Model

VECM system, lag order 12

Maximum likelihood estimates, observations 1991:04-2013:08 (T = 269)

Cointegration rank = 1

| | β (cointegrating vectors, standard errors in parentheses) | α (adjustment vectors) |
|---|---|---|
| GLP | 1.0000 (0.0000) | 0.00015683 |

| | | |
|---|---|---|
| SNX | 1.9362 (0.54604) | -0.021286 |
| OIL | 0.60729 (0.68443) | 0.00072074 |
| EXR | 17.300 (2.2887) | -0.0077827 |
| INT | 1.4296 (0.47508) | 0.018677 |
| CPI | -6.8974  (1.9428) | -0.0054060 |

Log-likelihood = 2652.2341

Determinant of covariance matrix = 1.0995337e-016

AIC = -16.4627; BIC = -10.6096; HQC = -14.1121

Mean dependent var   0.001679   S.D. dependent var   0.048958

Sum squared resid   0.394416   S.E. of regression   0.044298

R-squared         0.386001   Adjusted R-squared   0.181335

rho          -0.010599   Durbin-Watson       2.006257

## 4.3  Forecasting the Price of Gold

To test our model's out-of-sample properties, we re-estimate our model restricting our data to August 2012. We use our model to forecast one year of data (Figure 2) and find that the average error is about 3.47%. (See Table 4). The Root Mean Squared Error of 0.11 is much less than the equivalent measure for a random walk model which is 0.75.
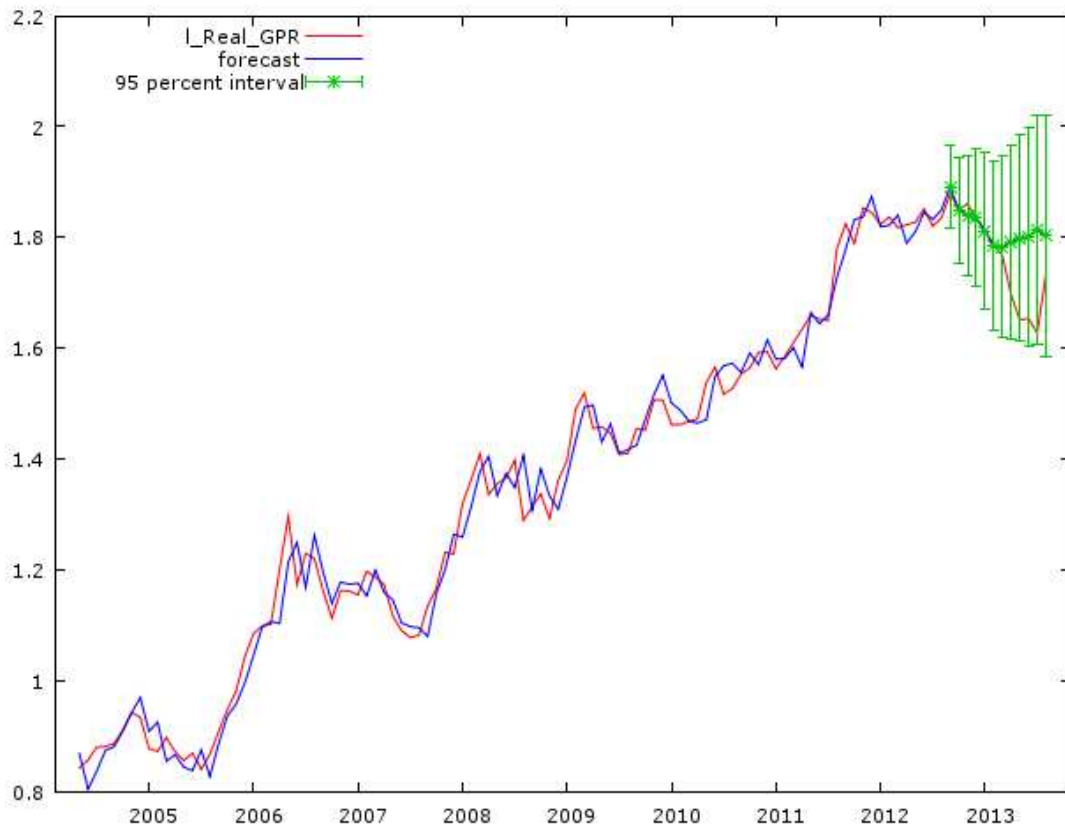
Figure 2. Forecasted values: The magnitude of error

*Table 4: Forecast evaluation statistics*

| | |
|---|---|
| Mean Error | -0.054101 |
| Mean Squared Error | 0.0077018 |
| Root Mean Squared Error | 0.08776 |
| Mean Absolute Error | 0.058156 |
| Mean Percentage Error | -3.2585 |
| Mean Absolute Percentage Error | 3.4775 |
| Theil's U | 2.0555 |
| Bias proportion, UM | 0.38003 |
| Regression proportion, UR | 0.07974 |
| Disturbance proportion, UD | 0.54023 |

## 4.4 Robustness Exercises

We carry out a series of exercises to assess the robustness of our results. For our model to be reliable, the estimated error terms should be normally distributed without any significant autocorrelation. This can be done using the Ljung-Box test which has good small sample properties as well. The null hypothesis of this test assumes the error terms

to be normally distributed and independent. Hence an acceptance of the null would imply that our estimates are robust. We carry out the test for the estimated error term from all the equations of the VECM. For all the equations the p-values of the test statistics are actually very high (> 0.90) which we interpret as non-existence of serial autocorrelation. The combined residual plots (Figure 3) bear testimony to our analysis. Furthermore we tried out the above exercise with alternative lag orders from 9 to 15 (benchmark model presented is for lag order 12) but the results thus obtained are qualitatively not different from the benchmark model.
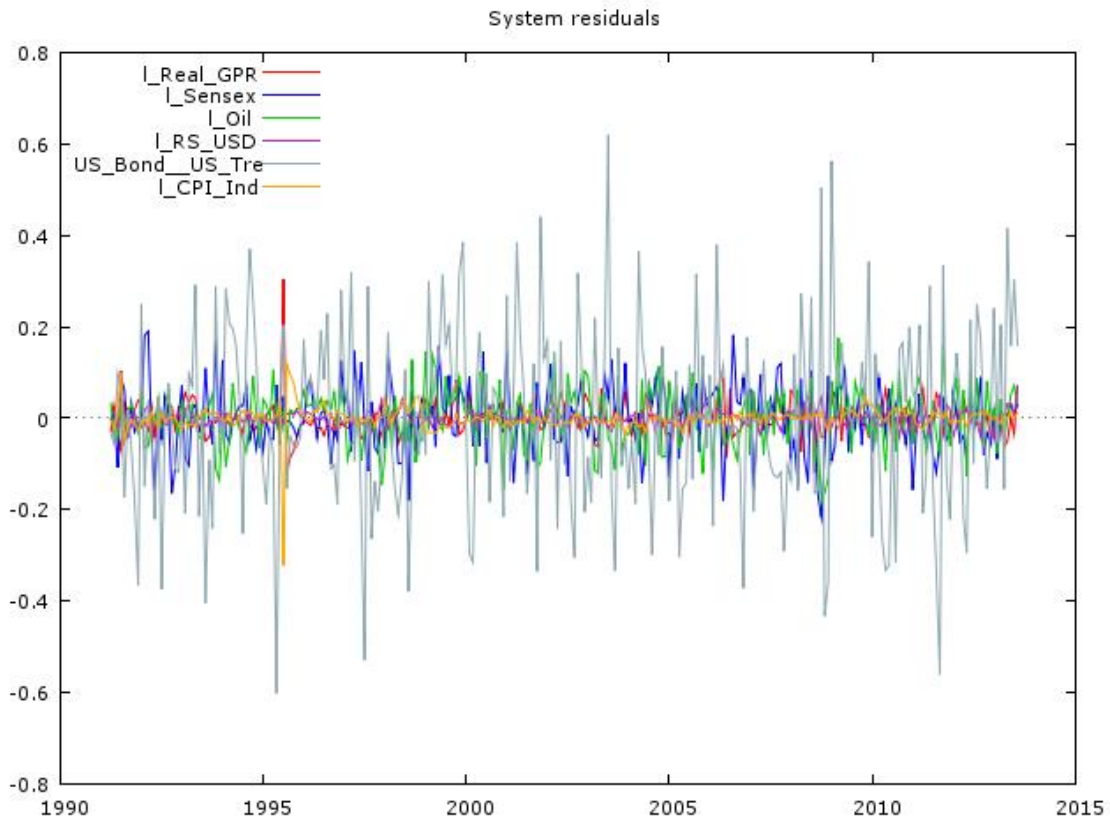


Figure 3. Forecasted values: The magnitude of error

## 5 CONCLUSION

We have modelled gold prices in India and shown it to have a long term relationship with the stock market index, exchange rate, US bond rates, oil prices and the consumer price index. We found evidence that the role of gold as a portfolio hedge dominates its use as a luxury good in India. Gold prices are negatively related with oil prices further indicating the role of gold as a hedge. Gold prices go up when the rupee is weaker implying that gold is a good hedge against the dollar. When returns from investing outside the country are high, gold price in India is low. Finally gold acts as a good inflation hedge as it moves in the same direction as CPI. We found evidence that the above variables are able to forecast gold over a 12-month horizon better than a random walk model can.

One implication of our results is that since gold seems to be a useful portfolio hedge as well as inflation hedge, government policies to curb the import of gold may be futile.

Yet the large amounts of gold imports are a cause for concern as they have kept India's current account deficit high leading to pressure on the rupee. Our research suggests that policies that directly address the causes of inflation and provide alternative investment opportunities for retail investors may better serve the objective of bringing down gold imports.

Future work in this area can proceed in several directions. In terms of methodology, alternative approaches such as Copula, Artificial Neural Networks, Fourier Transformation and Wavelet Analysis can be employed to assess the scope for improvement in the forecasting power. Other research approaches such as behavioural finance models can be tested using micro data on investors' personal choices to study its influence on gold prices. Studies can compare the gold holding decisions of households and corporate houses to evaluate the consumption vis-à-vis investment motives behind gold purchase.

## REFERENCES

| | |
|---|---|
| [1] | Baur, D.G. and Lucey, B.M. (2010), Is Gold a Hedge or a Safe Haven? An Analysis of Stocks, Bonds and Gold, *The Financial Review*, Volume 45, Issue 2, Pages 217–229. |
| [2] | Capie, F., Mills, T.C. and Wood, G. (2005), Gold as a hedge against the dollar, *Journal of International Financial Markets, Institutions and Money*, Vol. 15, Issue 4, pp. 343–352. |
| [3] | Worthington, A.C. and Pahlavani, M. (2007), Gold investment as an inflationary hedge: cointegration evidence with allowance for endogenous structural breaks, *Applied Financial Economics Letters*, Vol. 3, Issue 4, pp. 259-262. |
| [4] | Ng, S., and Perron, P. (1995), Unit root tests in ARMA models with data-dependent methods for the selection of the truncation Lag, *Journal of the American Statistical Association*, Vol. 90, pp. 268-281. |
| [5] | Schwert, W. (1989), Test for Unit Roots: A Monte Carlo Investigation, *Journal of Business and Economic Statistics*, Vol. 7, pp. 147-159. |
| [6] | Engle, R.F., and Granger, C.W.J. (1987), Co-integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, Vol. 55, Issue 2, pp. 251-76 |
| [7] | Enders, W, (2004), Applied Econometric Time Series, John Wiley & Sons, Hoboken, NJ. |
| [8] | Johansen, S. (1988), Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*,Vol. 12, Issue 2-3, pp. 231-254. |
| [9] | Johansen, S. (1991), A Bayesian Perspective on Inference from Macroeconomic Data: Comment, *Scandinavian Journal of Economics*, Vol. 93, Issue 2, pp. 249-51. |
| [10] | Johansen, S. and Juselius, K. (1990), Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics*, Vol. 52, Issue 2, pp. 169-210. |
| [11] | Reboredo, J. C. (2013). Is gold a hedge or safe haven against oil price movements?. *Resources Policy*, *38*(2), 130-137. |

| [12] | Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking & Finance*, *34*(8), 1886-1898. |
| [13] | Reboredo, J. C. (2013). Is gold a safe haven or a hedge for the US dollar? Implications for risk management. *Journal of Banking & Finance*, *37*(8), 2665-2676. |
| [14] | Ghosh, D., Levin, E. J., Macmillan, P., & Wright, R. E. (2004). Gold as an inflation hedge?. *Studies in Economics and Finance*, *22*(1), 1-25. |