12-2016

Forecasting the Spot Price of Corn: Methods and Assessment

Daniel G. Halonen
Saint Cloud State Uniteristy, hada0803@stcloudstate.edu

Forecasting the Spot Price of Corn: Methods and Assessment

by

Daniel George Halonen

A Thesis

Submitted to the Graduate Faculty of

St. Cloud State University

in Partial Fulfillment of the Requirements

for the Degree of

Master of Science

in Applied Economics

September, 2016

Thesis Committee:
Ming Chien Lo, Chairperson
Mana Komai
David Robinson

Abstract

Of the current techniques used to forecast agricultural commodity prices, none carries as high of a cost as a supply and demand analysis. Because of this expense, firms that have the ability to produce forecasts that rely on supply and demand analysis, do not update their models very frequently. In this paper we will examine if statistical methodologies can provide price forecasts at least as accurate at supply and demand analysis techniques. Both statistical as well as supply and demand models will be evaluated at one, three, six, nine, and twelve month horizons. These horizons are typical for price forecasts that can be used for buying and selling activities, contract negotiations, and production decisions. Of the models we investigated an autoregressive model was found to provide the best price forecasts over a shorter horizon. Whereas a vector autoregressive model was shown to provide the best price forecasts beyond a six-month horizon. This study reveals that a few statistical techniques have the ability to outperformed models that incorporate supply and demand analysis in forecasting the price of U.S. corn.

Table of Contents

List of Figures

I. Introduction

Forecasting the price of agricultural commodities has become a hot topic comprising an extensive amount of research. Many different econometric techniques have been adopted to forecast agricultural commodity prices. Yet, many of the techniques which accurately predict prices carry a high cost, and, therefore, they are not commonly updated frequently. Often current forecasting methods are so costly that they are only updated, at most, on a monthly basis. For example, the USDA's WASDE report comes at a very high cost, approximately $3.2 billion in 2014, because it involves gathering and summarizing a large amount of global information on every factor of supply and demand. This process requires extensive surveys from a large sample of farmers in order to ascertain an estimation of crop progress, estimated crop plantings, a survey of yields, an estimate of crop disappearance, a value of total remaining commodity inventories, and an analysis of imports, exports, and future sales contracts. This analysis is normally undertaken by governmental agencies. Better predictions of future the commodity prices can lower the risk taken by the firms that use the commodity and will result in a lower price to the consumer. As agricultural commodities are used in a wide variety of goods that consumers depend on to meet their daily needs, many firms involved in the commodities trade end up pricing their goods by taking the volatile swings, common in the agricultural markets, into account. These overpricing practices cause consumers to pay more than is efficient. If firms who use agricultural commodities had a better understanding of what their inputs would cost, then they could price their products in a more efficient manner. This paper will evaluate whether a few of the basic forecasting models currently used can provide producers with the needed information to accurately price their goods in a cost effective manner.

In the searching for a method that is both timely and cost effective we will evaluate three different forecasting techniques for both their ability and accuracy in predicting the spot price of corn in the United States. Out of sample rolling forecasts at one, three, six, nine, and twelve month horizons will be examined using a structure similar to that of Meese and Rogoff (1983). Horizons of one, three, and six months provide participants near term price forecasts for decisions surrounding near term buying and selling activities. Both nine and twelve-month horizons provide information that buyers can use for contract negotiations and producers can use for production decisions. This methodology will be used to examine whether pure statistical forecasts can outperform more elaborate models that incorporate supply and demand style estimations. Statistical forecasting processes will be compared to the price estimate forecasts provided in the USDA's World Agricultural Supply Demand Estimation (WASDE) model's price over equivalent time frames. The main results can be shown as follows: First, over a shorter time horizon, less than three months, single variable autoregressive models can outperform supply and demand analysis techniques. Second, over a longer horizon other factors, such as the price of a commodity in another country, have the ability to outperform supply and demand techniques.

II. Literature Review

Commodity price forecasting is not a new topic, and a few of the earliest works were written nearly 100 years ago. Keynes (1931) suggested that investors who have longed a futures contract are looking to receive not only their expected value but also a risk premium for purchasing the contract. Kaldor (1939) described what he later came to call "the theory of storage," where he suggests that a futures basis has two main components. The first, a forgone interest rate lost as a result of having to borrow in order to buy the commodity. The second, a convenience yield which measures the benefit gained by holding the commodity for another period. These two components are widely accepted due to their intuitive appeals and are supported by evidence from subsequent research.

Other methods to forecast commodities are motivated by the apparent link between commodity products and macroeconomic conditions. Commodity products have been, and continue to be, a key part of macroeconomic policy, planning, and formulation (Bhardwaj, et al 2014). The relationship between commodity products and economic policy changes can be shown by the aggregate index of non-oil commodities which has been treated as a macroeconomic variable whose movements are related to prevailing macroeconomic conditions (Borsztein and Reinhart 1994). A simple example of a macroeconomic policy change is a change in the currency exchange rate. This change can cause a decrease or an increase in a given country's balance sheets hereby changing the demand for agricultural commodities. These changes lead to both supply and demand shifts which cause volatile swings in commodity prices. These price fluctuations cause volatility which leads to heteroscedastic tendencies. As a result, many researchers have turned to methods known for their ability to deal with heteroscedasticity and still provide accurate price forecasts.

Among the forecasting techniques used few seem more applicable to the task than the Box-Jenkin ARMA model. This is because agricultural price data tends to be noisy and is commonly characterized by short and long term price fluctuations. These noisy tendencies have even caused some researchers to think of commodity price data and financial data separately. Figlewski, (1981) and Kamara (1993) show that methods intended for financial data may not be automatically applied to commodity markets as the characteristics for financial instruments and for agricultural commodities may be quite different. One of the reasons commodity data has been viewed as different from financial data is that financial data is less prone to rapid and volatile changes from supply and demand concerns (Bhardwaj, et al 2014). Supply and demand concerns stem from changes in weather patterns, import and export policies, and changes on the consumer demand side, as well as the producer supply side of commodity products.

Despite the volatility of commodity data, theoretical works relating commodity price data to financial data have shown that they do in fact share similar tendencies. Yang, (2005) suggests that agricultural commodities tend to be “highly developed, integrated, and permit a significant amount of risk transfer by hedging through forward contracts.” This could be caused by a number of strategies aimed at risk mitigation being adopted into commodities markets and heavier industry utilization of these strategies. One major shift in agricultural markets was the introduction of the Commodity Futures Trading Commission (CTFC) in (1974). This expanded regulation of the commodities exchanges allowed for the development of single-stock futures. Then, in 1992 the first electronic futures trades were made on the CME Globex platform. These events allow for faster and more efficient trades to take place in the futures market place while alleviating the risk of one party defaulting on a futures contract.

The CME Group outlines this point by stating, “The original intent of introducing futures contracts was to help alleviate the risk of price fluctuations associated with purchasing a commodity product needed at a future point in time (CME Group 2008).” This is done by creating a contract that allows the buyer to secure a price with a seller for the future delivery of a commodity product at a price agreed upon today. Currently, physical delivery of a commodity only occurs on a minority of contracts as many are canceled out by traders who later purchase a covering position (The National Futures Association 2006). The adoption of futures contracts allows traders the ability to purchase a covering position. This activity also tends to create volatility in agricultural markets a few days before a contract reaches expiration as traders tend to try and close open positions so as to avoid taking delivery of a commodity product.

The level of volatility found in agricultural markets is typically high enough to suggest the use of ARIMA, ARCH or GARCH model (see Lama 2015) as a recommended method to deal with the heteroscedasticity present in the data. T. Xiong et al. (2015) suggests "agricultural commodity price forecasting is considered a challenging task due to the fact that the prices are highly volatile, complex and dynamic, and are of great interest to finance researchers, market practitioners, and policymakers." Methods to stabilize the violate swings have been developed. Case in point, Lama et, al., (2015) uses a modified GARCH model to reduce the volatility in a commodity price series. Often these stabilization techniques leave forecasts only mildly improved at a one or two period horizon over that of a first order autoregressive model. This has given rise to large entities and consultancy groups undertaking the development of extensive supply and demand models to gain an accurate prediction of the market price over an extended horizon. Yet, these supply and demand models used in the commodity industries normally carry a steep cost.

The USDA's National Agricultural Statistical Services (NASS) is, at the time of this paper, forecasted to spend 6% of its 3.2-billion-dollar budget to obtain and summarize high quality information about agricultural commodities (USDA 2014). The NASS then uses this information to develop a supply/demand analysis. This analysis is released to the general public around the second week of the month and is called the World Agricultural Supply and Demand Estimation (WASDE). The WASDE is a revered report that has the tendency to cause large swings in agricultural prices. These swings are similar to the movement in stock price indices following a meeting of the Federal Reserve Bank to discuss interest rates.

There are a few other consultancy groups who provide agricultural forecasts based on supply and demand analysis commercially, yet, the WASDE is still believed to provide the most accurate information. Many consultancy groups typically adjusted their estimates to match the WASDE report. Reuters releases a survey of a conglomerate of consultancy's groups in their WASDE pre-report expectations on a monthly basis. However, even the methods employed by the NASS have been known to be inaccurate in their evaluation of agricultural prices. An excellent summary of this is given by Lama et, al., (2015), who suggest, "agricultural commodity prices respond rapidly to the actual and the presumed changes in supply and demand conditions and weather-induced fluctuations in farm production worsen the situation." This leaves many involved in the agricultural trade wondering how they can obtain an accurate expectation of a future price, especially since the reliability of the major institutions has been known to provide less than adequate information.

III. Model Specification

We will let s_t be the spot or cash price equivalent for U.S. Number 2 Yellow corn, and we define Δs_t as the log difference of the spot price. We will examine the forecast performance of an autoregressive model (AR) and a vector auto regressive model (VAR). The lag distribution for both the AR and the VAR models will be chosen based on the lag order which minimizes the value of Akaike information criterion (AIC) test statistic

$$AIC = -2\left({}^{L}/_{T}\right) + 2\left({}^{k}/_{T}\right)$$

where L is the log-likelihood value, k is the number of parameters and T the number of observations. We will then evaluate if the AR or VAR processes perform better than the USDA's WASDE forecast over the same horizon based on three different measures of the forecast error. We also include a random walk model which allows us to examine if our forecast is better than a model which is equivalent to a coin flip.

Equation (1) is the standard form of an autoregressive regressive model with p lags:

$$\Delta s_t = \alpha_0 + \alpha_1 \Delta s_{t-1} + \cdots + \Delta s_{t-p} + \varepsilon_t \qquad (1)$$

We apply the model to the first differences instead of the levels because price data are non-stationary.

The vector autoregressive model is given by simultaneously computing equation (2) and (3):

$$\Delta s_t = c_1 + a_{1,1}\Delta s_{t-1} + a_{1,2}\Delta s^*_{t-1} + \cdots + a_{p,1}\Delta s_{t-p} + a_{p,2}\Delta s^*_{t-p} + \varepsilon_t \qquad (2)$$

$$\Delta s^*_t = c_2 + b_{1,1}\Delta s_{t-1} + b_{1,2}\Delta s^*_{t-1} + \ldots + b_{p,1}\Delta s_{t-p} + b_{p,2}\Delta s^*_{t-p} + \varepsilon^*_t \qquad (3)$$

To compute the forecast of s_t *k* periods ahead, i.e. s_{t+h} from time *t*, we first forecast the cumulative changes $\sum_1^h \Delta\hat{s}_t$ and then add it to s_t

The model used for the random walk with drift is given as equation (4)

$$s_t = s_{t-1} + \varepsilon_t \qquad (4)$$

Random walk is a non-stationary process and can be appropriately applied to the data at their level. In all the equations listed above we assume that ε_t is a randomly distributed error term with a mean of zero.

The recursive method is implemented as followed: Starting with an estimation sample from the beginning to December 2013, a point forecast for k period ahead is produced. Then we increase the sample by one data point (e.g. January 2014 in the second iteration), estimate, and forecast. We repeat these steps until the data are exhausted. For example, for k = 1, our forecast period covers January 2014-December 2015 with 24 point-forecasts; for k = 6, the period becomes July 2014-December 2015 with 18 point-forecasts. Our measurement of the accuracy will follow the methodology of Meese and Rogoff (1983), and we will evaluate each model based on the Mean Error ME, Root Mean Squared Error RMSE, and Mean Absolute Error MAE statistics listed below.

$$Mean\ Error\ (ME) = \sum_{c=0}^{N_k - 1} \frac{[F(t+c+k) - A(t+c+k)]}{N_k}$$

$$Mean\ Absolute\ Error\ (ME) = \sum_{c=0}^{N_k - 1} \frac{|[F(t+c+k) - A(t+c+k)]|}{N_k}$$

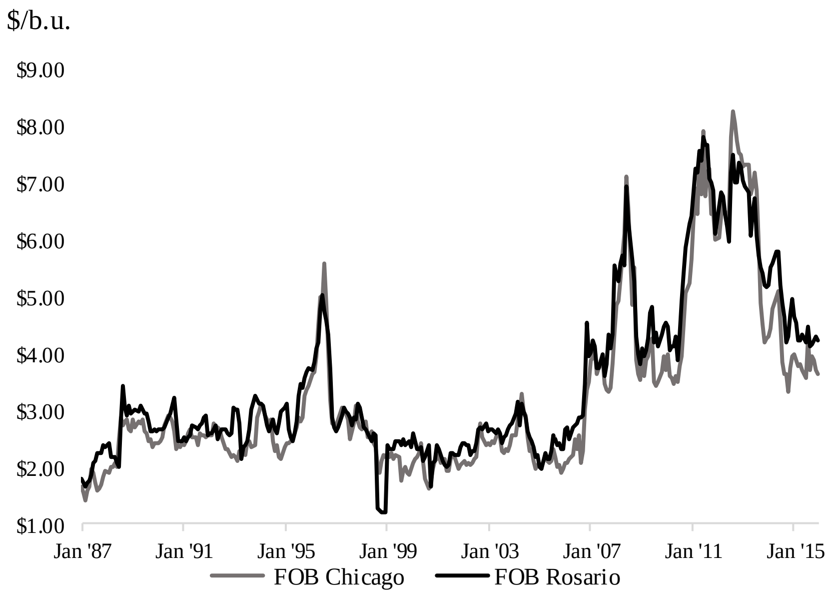
$$Root\ Mean\ Squared\ Error\ (RMSE) = \left\{ \sum_{c=0}^{N_k - 1} \frac{[F(t+c+k) - A(t+c+k)]^2}{N_k} \right\}^{1/2}$$

To start the evaluation of our forecasts we define k as the forecast horizon, where $k = 1, 3, 6, 9, 12$, (t) is given as the period in which the forecast begins. The variable N_k is the total number of forecasts estimated during the period, and we let A(t) define the actual spot price and F(t) as the forecast value. Following Meese and Rogoff, we use RMSE as our principle measure to evaluate the effectiveness of the forecast. RMSE is especially useful in this analysis as it provides a means to evaluate if there are especially large variations in our forecast. We also include MAE and ME as they provide a method to evaluate whether or not the models tend to under or over predict. Comparing RMSE and MAE we can get a sense of how large the variation in the error of our models is.

IV. Data

The price we are interested in is the spot price of corn in Chicago, United States. In our bivariate forecast model the price of corn from the port of Rosario, Argentina is chosen to help improve our forecasts of the United States corn price. Including the second price series, allows our bivariate model a method to help explain sudden spikes in price that maybe caused by rapid changes in supply. Economic theory shows that lower inventories (less supply) in one location typically lead to price spikes, while higher inventories lead to price declines. This is shown by Coleman (2009) who states, "Rational, risk-neutral arbitrageurs ship goods between centers if the expected future price in the importing center is at least as large as the price in the exporting center plus the cost of transport." The arbitrage activity that takes place in agricultural markets does not occur unless supply changes are expected to lead to price changes. The arbitrage activity that takes place suggests that commodity markets are subject to the Law of One Price which, according to Hanninen (1998), states "prices of homogenous commodities, defined in a common currency, are equal throughout the world." Provided the Law of One Price holds, our bivariate adjustment process will hold long run predictive power.

The U.S. price data are from the Chicago Mercantile Exchange (CME) and are non-seasonally adjusted prices for the U.S. Number 2 Nearby Corn Contract. This data is gathered from the CME and represents an aggregated average of the spot prices paid for U.S. Number 2 corn at the exchange. We gather a series that spans from August of 1986 and ends December of 2015. The Argentina price series comes from The Rosario Cereals Exchange and are reported on a monthly basis in Argentina Pesos per metric tonne of maize beginning August of 1986. To be able to use the Argentina maize price in our bivariate model we follow two conversions. First, exchange rates are used to convert the Argentine price in terms of U.S. Dollars per metric tonne. This conversion is done using the USD/ARS exchange rate listed on Forex. Second, the final adjustment from metric tonnes to bushel is done using the metric tonne to bushel conversion rate as listed by the U.S. Department of Commerce Weights and Measures division. We included graphs showing both the non-currency adjusted and the currency adjusted Rosario, Argentina corn price as well as the Chicago, U.S. corn prices. These graphs show that the exchange rate adjusted price is much better suited to our purposes than prices before the currency adjustment.



Source: Quandl, CME, BCR, and Forex
FOB Rosario is Calculated by the author based on Bolsa de Comercio de Rosario (BCR) and Forex data
Chicago Merchantile Exchange (CME) and Forex prices were retrived through Quandl.com

Figure 1: U.S. & Argentina Spot Price of Corn

Source: Bolsa de Comercio de Rosario (BCR)

Figure 2: Argentina Spot Price of Corn (Unadjusted exchange rate)

These two price series, Chicago U.S. and Rosario Argentina, are selected for several reasons. First, there are a large number of similarities in the nearby agricultural markets, including the types of commodity they produce, the technology available to grow these crops, and support from local firms to aid in the crop production process. Second, both Chicago and Rosario frequently trade large volumes of corn and soybeans. Third they are both in close proximity to major growing areas of a large variety of commodities. Forth, both cities have major ports near trading hubs with extensive networks allowing for large vessels to load and unload as well as similar import and export taxation policies. Finally, both areas have extensive historical data publically available allowing us to obtain a time series dating back to August of 1986. Many other countries that trade heavily in corn do not have an extensive a dataset that is as easily accessible.

V. Results

For the AR and the VAR models, we first estimate the whole sample in order to find the optimal order of lag. We set a maximum possible number of lags to 12, and the Akaike Information Criterion (AIC) selects an order of 12 for the AR and 10 for the VAR. Meese and Rogoff (1983) show that the random walk model outperformed all models especially in the short horizon. However, our univariate and bivariate models tend to outperform the random walk. The AR(12) was found to provide the smallest mean error ME, MAE, and RMSE for a one-month ahead forecast. The VAR(10) model outperformed all other methodologies on a 3, 6, 9, and 12 month horizon based on MAE and RMSE. In brief the autoregressive models dominate in all horizons. That the univariate AR(12) is the best for the very short horizon while the bivariate VAR(10) is the best for the longer horizon is a sensible result and indicates that a Law of One Price adjustment process has long run predictive power. The results of ME are shown by table 1 in the appendix. The results of MAE and RMSE are also shown in the appendix by tables 2 and 3 respectively.

To verify if our results are robust we change our sample to look at the forecast period from January 2013 to December 2014. The results prove to be robust for RMSE, our principle measure, but are inconsistent for ME and MAE. The autoregressive model held steady and outperformed all other models for a one horizon forecast. Our VAR model outperform the USDA's forecasts at three and six month horizons. The USDA's forecast outperformed all other forecasts at a nine and twelve month horizon based on ME, and at a nine month horizon based on MAE. Tables showing the results of ME, MAE, and RMSE used to verify robustness are detailed in table 4, 5, and 6 in the appendix.

VI. Conclusion

In conclusion, this study has shown that statistical techniques can outperform supply/demand analysis in forecasting the price of U.S. corn. This is a very useful finding as this methodology can add value to firms that rely on corn. As the methods described in this study come at a relatively low cost, compared to the methods employed by the USDA, many of our methods could be used to help producers realize a higher profit by pricing their products more efficiently. Further analysis could be used to extend this study. One possibility would be to examine if other models, such as generalized autoregressive conditional heteroskedastic (GARCH) models or autoregressive conditional heteroskedastic (ARCH) models, outperform the models this study investigated. Another key possibility to investigate is whether other country's corn price series have a significant impact on our VAR forecasting methodology. This examination should also include an analysis of why the countries studied increase or decrease the accuracy of the VAR methodology.

References

Bhardwaj, S. P., et al. "An Empirical Investigation of Arima and Garch Models in Agricultural Price Forecasting." Economic Affairs: A Quarterly Journal Of Economics 59, no. 3 (September 2014): 415-428. Business Source Premier, EBSCOhost (accessed July 23, 2015).

Coleman, Andrew. "Storage, slow transport, and the law of one price: theory with evidence from nineteenth-century U.S. corn markets." Review Of Economics And Statistics, no. 2 (2009): 332. Academic OneFile, EBSCOhost (accessed May 23, 2016).

Figlewski, Stephen, and Paul Wachtel. "The Formation of Inflationary Expectations." Review Of Economics & Statistics 63, no. 1 (February 1981): 1. EBSCO MegaFILE, EBSCOhost (accessed November 1, 2015).

Hänninen, R.H. "The law of one price in United Kingdom soft sawnwood imports – A cointegration approach." Forest Science 44, no. 1 (1998): 17-23. Scopus®, EBSCOhost (accessed May 24, 2016).

Kaldor, N. "Speculation and economic stability." Review Of Economic Studies 7, no. 1 (January 1, 1939): 1-27. Scopus®, EBSCOhost (accessed June 9, 2015).

Kamara, Avraham. "Production Flexibility, Stochastic Separation, Hedging, and Futures Prices." The Review of Financial Studies, 1993., 935, JSTOR Journals, EBSCOhost (accessed September 24, 2015).

Keynes, J. M. "[Mr. Keynes' Theory of Money]: A Rejoinder." The Economic Journal, 1931., 412, JSTOR Journals, EBSCOhost (accessed November 5, 2015).

Lama, Achal, et al. "Modelling and Forecasting of Price Volatility: An Application of GARCH and EGARCH Models." Agricultural Economics Research Review 28, no. 1 (January 2015): 73-82. EBSCO MegaFILE, EBSCOhost (accessed September 10, 2015).

Meese, Richard A., and Kenneth Rogoff. "Empirical Exchange Rate Models of The Seventies." Journal Of International Economics 14, no. 1/2 (February 1983): 3-24. EBSCO MegaFILE, EBSCOhost (accessed January 5, 2015).

National Futures Association, The. "An Educational Guide to Trading Futures and Options on Futures." (2006): http://www.nfa.futures.org/nfa-investor- information/publication-library/opportunity-and-risk-entire.pdf. (accessed June 4, 2015)

Xiong, Tao, et al. "A combination method for interval forecasting of agricultural commodity futures prices." Knowledge-Based Systems (2015): 92. Academic OneFile, EBSCOhost (accessed September 19, 2015)

Yang, Chao-Hsiang, Chi-Tai Lin, and Yu-Sheng Kao. "Exploring stationarity and structural breaks in commodity prices by the panel data model." Applied Economics Letters 19, no. 4, (March 2012): 353-361. EBSCO MegaFILE, EBSCOhost (accessed April 27, 2015)

Appendix A: List of Tables

Table 1
Mean Forecast Error
(Maximum Forecast Period: 2014-2015)

Model:	Random Walk	Auto Regressive	Vector Auto Regressive	USDA Forecast
Horizon				
1 Month	0.0067	0.0045	0.0044	-0.0237
3 Month	0.0191	0.0216	0.0136	-0.0275
6 Month	0.0965	0.0443	0.0289	-0.0202
9 Month	0.0872	0.0584	0.0336	-0.0435
12 Month	0.0894	0.0747	0.0461	-0.0493

Table 2
Mean Absolute Forecast Error
(Maximum Forecast Period: 2014-2015)

Model:	Random Walk	Auto Regressive	Vector Auto Regressive	USDA Forecast
Horizon				
1 Month	0.0423	0.0153	0.0301	0.0536
3 Month	0.0884	0.0262	0.0249	0.0569
6 Month	0.0595	0.0443	0.0269	0.0543
9 Month	0.1139	0.0584	0.0336	0.0450
12 Month	0.1305	0.0747	0.0461	0.0512

Table 3
Root Mean Square Forecast Error
(Maximum Forecast Period: 2014-2015)

Model:	Random Walk	Auto Regressive	Vector Auto Regressive	USDA Forecast
Horizon				
1 Month	0.0561	0.0186	0.0381	0.0719
3 Month	0.1163	0.0326	0.0300	0.0748
6 Month	0.1451	0.0485	0.0363	0.0748
9 Month	0.1476	0.0626	0.0401	0.0609
12 Month	0.1588	0.0788	0.0513	0.0664

Table 4
Mean Forecast Error
(Maximum Forecast Period: 2013-2014)

Model:	**Random Walk**	**Auto Regressive**	**Vector Auto Regressive**	**USDA Forecast**
Horizon				
1 Month	0.0262	0.0035	-0.0044	-0.0561
3 Month	0.0944	0.0228	0.0174	-0.0608
6 Month	0.2096	0.0488	0.0359	-0.0477
9 Month	0.3088	0.0711	0.0527	0.0042
12 Month	0.3828	0.0873	0.0633	-0.0021

Table 5
Mean Absolute Forecast Error
(Maximum Forecast Period: 2013-2014)

Model:	**Random Walk**	**Auto Regressive**	**Vector Auto Regressive**	**USDA Forecast**
Horizon				
1 Month	0.0578	0.0165	0.0352	0.1018
3 Month	0.1417	0.0273	0.0279	0.1100
6 Month	0.2447	0.0488	0.0372	0.1017
9 Month	0.3088	0.0711	0.0527	0.0600
12 Month	0.3828	0.0873	0.0633	0.0603

Table 6
Root Mean Square Forecast Error
(Maximum Forecast Period: 2013-2014)

Model:	**Random Walk**	**Auto Regressive**	**Vector Auto Regressive**	**USDA Forecast**
Horizon				
1 Month	0.0777	0.0195	0.0491	0.1573
3 Month	0.1847	0.0321	0.0317	0.1643
6 Month	0.2904	0.0535	0.0443	0.1493
9 Month	0.3447	0.0745	0.0562	0.0774
12 Month	0.4091	0.0909	0.0661	0.0809

Appendix B: Matlab Programs

Note: Any program error is the fault of the author.

This procedure compute the estimates for the parameters of an AR(p) model given *p*. Likelihood values and the AIC statistics are reported as well.

```
function[beta,log_lik,AIC]=ar(data,p)
T=size(data,1);
YY=data(p+1:T);
XX=zeros(T-p,p+1);
XX(:,1)=ones(T-p,1);
for i=1:p;
    XX(:,i+1)=data(p-i+1:T-i);
end;
beta=inv(XX'*XX)*(XX'*YY);
u_hat=(YY-XX*beta);
n=size(u_hat,1);
SSR=sum((u_hat).^2);
sigma=sqrt(SSR/(n-1));
sigma2=sigma.^2;
log_lik=0.0;
for i=1:n;
    f=log((1/(sqrt(2*pi*sigma2)))*exp(-u_hat(i)^2/(2*sigma2)));
```

```
    log_lik=log_lik+f;
end;
k=size(beta,1);
T=size(YY,1);
AIC=-2*(log_lik/T)+2*(k/T);
end
```

This procedure performs the recursive forecasts based on an AR(p) model discussed in the text.

```
function[f]=arf(Y,p,start,h,m)
f=zeros(m-start-h,1);
 for j=1:(m-start-h)+1;
    St=Y(1:start+j-1);
    m=size(St,1);
    [b,log_lik,AIC]=ar(St,p);
    St2=St(m-p:m);
    aYf=zeros(h,1);
    for k=1:h;
        nn=size(St2,1);
        Yf=b(1)+b(2:p+1)'*flipud(St2(nn-p+1:nn));
        aYf(k)=Yf;
        St2=[St2;Yf];
   end;
    f(j)=sum(aYf);
end;
end
```

This procedure sorts the data to prepare for the VAR estimation.

```
function[YY,XX]=sortdata(data,p)

T=size(data,1);

n=size(data,2);

YY=data(p+1:T,:);

XX=zeros(T-p,n*p+1);

XX=ones(T-p,1);

for j=1:n;

    for i=1:p;

        XX=[XX data(p-(i-1):T-i,j)];

    end;

end;

end
```

This procedure computes the estimates for parameters of a VAR model. The AIC statistics is computed as well.

```
function[beta,se,e,omega,loglik,AIC]=var(data,p2)
[YY,XX]=sortdata(data,p2);
T=size(data,1);
n=size(data,2);
beta=inv(XX'*XX)*(XX'*YY);
e=YY-XX*beta;
SSE=diag(e'*e);
sigma2s=SSE/(T-p2-n*p2-1);
sigmas=sqrt(sigma2s);
se=zeros(n*p2+1,n);
for i=1:n;
    se(:,i)=sqrt(diag(inv(XX'*XX)*sigma2s(i)));
end;
omega=det((1/(T-p2))*(e'*e));
loglik=-0.5*(T-p2)*(n*(1+log(2*pi))+log(omega));
AIC=-2*loglik/(T-p2)+2*n*(p2*n+1)/(T-p2);
end
```

This procedure performs the recursive forecasts based on an VAR(p) model discussed in the text.

```
function[VarF]=VARF(data,p2,start2,h,m)
L=size(data,2);
```

```
VarF=zeros(m-start2-h,L);
for j=1:(m-start2-h)+1;
    St=data(1:start2+j-1,:);
    m=size(St,1);
    [beta,se,e,omega,loglik,AIC,SIC]=var(St,p2);
    St2=St(m-p2+1:m,:);
    aYf=zeros(h,L);
    for k=1:h;
        nn=size(St2,1);
        Yf=zeros(1,L);
        for i=1:L;
            Yf(i)=beta(1,i)+(beta(2:p2*L+1,i)')*reshape(flipud(St2(nn-p2+1:nn,:)),[],1);
        end;
       aYf(k,:)=Yf;
        St2=[St2;Yf];
    end;
    VarF(j,:)=sum(aYf);
end;
end
```

This procedure performs the recursive forecasts based on a random walk model discussed in the text. A drift is allowed as an option.

```
function[YfR]=rw(Y,start1,n,h,c)
YfR=zeros(n-start1-h,1);
for i=start1:n-h;
   a=mean(Y(2:i)-Y(1:i-1))*(c==1);
   Yf_temp=Y(i);
   for j=1:h;
      Yf_temp=Yf_temp+a;
   end;
   YfR(i-start1+1)=Yf_temp;
end;
end
```

The final portion of our Matlab program has four main components. First, we define the criteria we are interested in. This program has several parts we define our starting point (start), that is the point in the data we want to begin estimating our forecast. We select the horizon (h) we are interested in evaluating. Then we specify the maximum number of models we want to calculate the AIC statistic for (p). Second, we call our data into Matlab and perform any data transformation we are interested in. Third, we call on the functions that we have previously created in order to run each forecast. Fourth, we call on the newly created forecast vectors and the actual data over the same time period to calculate the comparison statistics as in Meese and Rogoff (1983).

```
clear;
 data=xlsread('data.xlsx',1,'b2:c354');
 n=size(data,1);
 h=12;
start=329;
 p=12;
 c=0;
 Y=data;
 Y1=log(Y(:,1));
 Y2=log(Y(:,2));
 Yd1=log(Y1(2:n))-log(Y1(1:n-1));
 Yd2=log(Y2(2:n))-log(Y2(1:n-1));
 Yvar=[Yd1 Yd2];

 [YfRW]=rw(Y1,start,n,h,c);
 AICm=zeros(p,1);
for i=1:p;
```

```
        [beta,log_lik,AIC,SIC]=ar(Yd1,i);
        AICm(i)=AIC;
end;
[a,b]=min(AICm);
Par=b;
[f]=arf(Yd1,Par,start,h,n);
YfAR=Y1(start+h:n)+f;
AICmat=zeros(p,1);
loglikmat=zeros(p,1);
for i=1:p;
    [beta,se,e,omega,loglik,AIC,SIC]=var(Yvar,i);
    AICmat(i)=AIC;
    loglikmat(i)=loglik;
end;
[AICmin,AICmin_i]=min(AICmat);
loglik_l=loglikmat(AICmin_i);
Pvar=AICmin_i;
[VarF]=varf(Yvar,Pvar,start,h,n);
YfVAR=Y1(start+h:n)+VarF(:,1);

Actual=Y1(start+h:n);
meRW=mean(YfRW-Actual);
```

```
meYfAR=mean(YfAR-Actual);meYfVAR=mean(YfVAR-Actual);
maeRW=mean(abs(YfRW-Actual));
maeYfAR=mean(abs(YfAR-Actual));
maeYfVAR=mean(abs(YfVAR-Actual));
rmseRW=sqrt(mean((YfRW-Actual).^2));
rmseYfAR=sqrt(mean((YfAR-Actual).^2));
rmseYfVAR=sqrt(mean((YfVAR-Actual).^2));
```