William A. Orme

WORKING PAPER SERIES

encouraging creative research

ARIMA Modeling With Intervention to Forecast and Analyze Chinese Stock Prices

Jeffrey E. Jarrett, Eric Kyper

2010/2011 No. 10

This working paper series is intended to facilitate discussion and encourage the exchange of ideas. Inclusion here does not preclude publication elsewhere. It is the original work of the author(s) and subject to copyright regulations.
ARIMA Modeling With Intervention to Forecast and Analyze Chinese Stock Prices

Jeffrey E Jarrett, Ph.D.
University of Rhode Island (USA)
jejarrett@mail.uri.edu

Eric Kyper, Ph.D.
Lynchburg College (USA)
kyper@lynchburg.edu
ABSTRACT

In this study, we demonstrate the usefulness of ARIMA-Intervention time series analysis as both an analytical and forecast tool. The data base for this study is from the PACAP-CCER China Database developed by the Pacific-Basin Capital Markets (PACAP) Research Center at the University of Rhode Island (USA) and the SINOFIN Information Service Inc, affiliated with the China Center for Economic Research (CCER) of Peking University (China). This data base is recent and not fully explored in any published study. The forecasting analysis indicates the usefulness of the developed model in explaining the rapid decline in the values of the price index of Shanghai A shares during the world economic debacle stating in China in 2008. Explanation of the fit of the model is described using the latest development in statistical validation methods. We note that the use of a simpler techniques although parsimonious will not explain the variation properly in predicting daily Chinese stock prices. Furthermore, we infer that the daily stock price index contains an autoregressive component.
Introduction

In this study, we report an analysis not heretofore reported in the literature and for the data base collected. Our purpose is twofold; (1) to analyze data by implementing ARIMA-Intervention analysis (Box and Tiao, 1975); and to assess the prediction and forecasting of Chinese stock market prices over a lengthy enough period of time where stock prices fluctuated during varying temporal economic movements. We study the Chinese equity markets to understand the predictable properties in this fast growing but volatile market. Previous studies of the Chinese equity markets include Jarrett, Pan and Chen (2009) who studied the relationship between the Macroeconomy of China as it relates to the equity markets of Shanghai and Shenzhen. In addition, Thomas (2001) discussed in detail the workings of the Shanghai equity market over its illustrious history offering insights into the characteristics of the equity markets. His study included the behavior of both domestic and foreign investors in four equity markets. The history discussed is of the twentieth century and not the twenty-first century. Others (Eun and Huang, 2007; Ng and Wu, 2007; Shenoy and Ying, 2007, Weili et al., 2009, Wang et al., 2004 and Wang et al., 2007) investigated the rapid growth in the Chinese equity markets and why they became increasingly important for investors in global markets. In addition, Bailey, Cai, Cheung and Wang (2006) studied the Shanghai Equity Market and discovered some characteristics concerning order balances of individuals, and institutional investors. Our purpose, here, is to investigate the Chinese equity markets to explain and discover if certain anomalies noted in more developed equity markets are present or not in the Chinese markets.

Following the study of Zhong, Gui and Lui (1999) for the Chinese Bourses (Shanghai and Shenzhen), we selected a newer and finer data base from the PACAP-CCER China Database developed by the Pacific-
Basin Capital Markets (PACAP) Research Center at the University of Rhode Island (USA) and the SINOFIN Information Service Inc, affiliated with the China Center for Economic Research (CCER) of Peking University (China). The length of data was for approximately ten years resulting in a time period where analysis can lead to interpretable results. Smaller time periods such as two or three or even four years are usually too small to reduce the effects of disturbances in economic data and more specifically do not produce enough degrees of freedom such that one may identify significant events and factors. A small number of degrees of freedom in sample data may not lead to determining “statistically significant” events even when they exist.

The regression will include for the response variable is the daily price index for the entire data base on daily returns produce by the PACAP-CCER Greater China Database. These data are part of the website created for research on Pacific-Basin nations. [The website and its content is a copyright of PACAP-CCER. All rights reserved. If you have any question or comment, please e.mail to chinadat@etal.uri.edu.]. The data base contains information from 01/1990 to 12/2009. The length of time is long enough for serious investigation which produce results that is not highly influenced by business cycles but long enough for us to create a useful ARIMA-Intervention model for forecasting.

**ARIMA Modeling and Intervention Analysis**

We follow the methods of **stochastic time-series ARIMA** modeling to analyze and find a model for prediction of changes, variation sand interruptions (interventions) in the movement of stock prices for China over time. ARIMA modeling with intervention may help analysts understand whether economic events such as a “global financial crisis” have impact on the change in Chinese stock prices if it exists (Shanghai A-share simple arithmetic mean index). This will provides information for financial analysts, economic forecasters and decision-makers to plan for and cope with sudden movements in the prices of Chinese stock. Previously, the application of ARIMA modeling with and without intervention was implemented in such applications as system scheduling and simulation (Ip, 1997; Ip et al., 1999), tourism forecasting (Du Preez and Witt, 2003, and Cho 2001) investigation and forecast of economic factors (Chung et al., 2008) the financial crisis in China’s Manufacturing (Chung et al., 2009) demand in telemarketing centers (Bianchi, Jarrett and Hanumara, 1998) and the impact analysis on air travel demand (Lai and Lu, 2005).
As discussed before by Box and Tiao (1975) an intervention model is of the general form:

$$Y_t = V(B) I_t + N \quad (1)$$

where $I_t$ is an intervention or dummy variable that is defined as $I_t = 1$, for $t = T$ and $I_t = 0$ for $t \neq T$.

To identify the correct model for prediction and explanation, ARIMA modeling consists of three steps. First, we identify the order of the differencing, i.e., $d$. This preliminary step is essential to produce a stationary time series model for identification. There are several solutions including the Augmented Dickey-Fuller (ADF) test and those following the Ljung and Box (1978) analysis. In this analysis, we chose both methods although each method produced the same results. Either method will yield the correct order of the left hand side variable (price). Second, one diagnoses and identifies the parameters autoregressive and moving-average portions of the ARIMA model, i.e., the $p$ and $q$ terms. $p$ refers to the autoregressive component and $q$ refers to the moving-average component of the model. The autocorrelation function (ACF), partial autocorrelation function (PACF) and the cross-autocorrelation function (CACF) are important to tentatively estimate the parameters of the model. In turn, statistical measures provide the evidence to support the determination of an appropriate model for estimation and prediction. Third, one diagnoses the residuals (noise). The Portmanteau Q-statistic is in turn permits one to judge the fit of the model to the data based on the ACF of the residuals. Based on the results of this three step procedure which diagnoses the data, we are led to choosing a lag of one to produce and stationary series. In turn, the examination of the various autocorrelation and partial autocorrelation functions; we are led to concluding that an autoregressive (AR) model having one term fits the data studied. We define the intervention as the present financial crises initially triggered in July, 2007 which spread first to Europe and then to China. Hence, data was assigned a value of 1 for the variable for $I_t$, [an intervention, the variable denoted by inter], beginning July 2007 and 0 before the intervention period. For $V$ (volume), the data is the volumes of transactions corresponding to the daily price index.

**The Analysis of the ARIMA Intervention Model**

We include the financial crisis as the intervention of the time series which results in an ARIMA-Intervention model. As the global financial crisis extended to China in September 2008 and beyond, the intervention was initiated at that time. In the Appendix, we have the results in tabular form for both the model estimates and the initial $T$-statistics for the model. Note, the coefficients have $T$-statistics largely
different from zero with p-values of 0.0000. This indicates that all the estimated coefficients are nonzero. One way to write the equation of the model is as follows:

\[ Y_t = 9.937475 + 0.499991 (\Delta Y_t) + (2.30E-09) V_t -2.324061 I_t + 0.992153 Y_{t-1} \]

When we evaluate the differenced term, \((\Delta Y_t)\), the equation for the model becomes:

\[ 0.500009 Y_t = 9.937475 -0.499991 Y_{t-1} + (2.30E-09) V_t -2.324061 I_t + 0.992153 Y_{t-1} \]

Last, we evaluate the AR (1) term 0.992153 \(Y_{t-1}\) to form the final estimation model:

\[ 0.500009 Y_t = 9.937475 + 0.492162 Y_{t-1} + (2.30E-09) V_t -2.324061 I_t \]

What we learn from the estimated model equation is that the intervention is significant and that prices were inversely related to financial crisis. This is what we would expect the results to be and is likely to be corroborated by studies using other methods for analyzing time series data having been related to the economics of the environment. Furthermore, the volume effect, although statistically significant, leads us to conclude that there is only a tiny effect on the volume of transaction on prices. The differencing which was necessary to build a time series that was stationary for modeling purposes was also significant and positive. Last, the significant autoregressive coefficient of the first order, i.e., AR (1), led to the estimated coefficient of 0.492162 for lagged value of \(Y_t\), i.e., \(Y_{t-1}\).

**Additional Analytical Results**

In addition the analytical table in the Appendix further gives diagnostic information about the estimated model and its equation. The coefficient of determination and the adjusted coefficient of determination are both extremely large and are the same because of the large number of degrees of freedom, i.e., 0.987864. The standard error of regression (3.065267) is relatively small. The log likelihood statistic is -5756615 and is the ingredient in likelihood ratio tests. A **likelihood ratio test** is used to compare the fit of two models one of which is nested within the other. Both models are fitted to the data and their log-likelihood recorded. A null hypothesis is often stated by saying the parameter \(\theta\) is in a specified subset \(\Theta_0\) of the parameter space \(\Theta\). Higher values of the statistic mean that the observed outcome was more than or equally likely or nearly as likely to occur under the null hypothesis as compared to the alternate, and the null hypothesis cannot be rejected. Since our results indicate a very large value for the log likelihood, we cannot reject the the null hypothesis.
The F-statistic of 46137147 with a probability of 0.0000 indicates that the regression coefficients (except for the constant or intercept term) are nonzero. The probability (p-value) indicates the marginal significance of the F-test.

The AIC (Akaike information criterion, 1974) indicates the usefulness of model selection for non-nested alternative. Smaller values for AIC is preferred. As in our illustration, one best chooses the length of the lag distribution by choosing the lowest value of the AIC, 5.078149 is for our estimated model. The Schwarz criterion(1978) is 5.078156. It is an alternative to the AIC and penalizes more greatly for additional coefficients. The two criterions are so similar that either will lead to the conclusion that this model is best. Last, the Hannan-Quinn Criterion (1979) yields an additional test for model selection similar to the other two noted before. The result is again very similar to the AIC and Schwarz criterions already considered.

The well-known Durbin-Watson Statistic is the most universally accepted testing criterion for measuring the magnitude of autocorrelation in the residuals from a model. A statistic of 2.245173 would lead one to conclude that one should accept the hypothesis of the autocorrelation coefficient equal zero. With the absence of autocorrelation in the residuals from forecast, one concludes the previous tests of significance are useful and the model is available for predictive purposes. The Portmanteau Q-statistic calculated along with the autocorrelation function of the residual resulted in the same conclusion. Hence, the estimated model was a good fit for the time series. If we were to fit the model without the intervention the AIC, Schwarz Criterion and Hanna-Quinn Criterion would all have larger values indicating also that the model with the intervention was a better fit to the data. The coefficient of determination adjusted for degrees of freedom was also small for the model without intervention.

**Summary and Conclusions**

The stock market price index in China was modeled by the methods of ARIMA-Intervention analysis and produced a fit for one to analyze and draw conclusion concerning how the index behave over time. The results indicate that the World Financial Crisis included China and affected its stock as well as its manufacturing industry. The intervention effect as measured in this study was alarming and clearly felt by the financial industry in China as well as its other component. The negative effect of the stock market decline is likely to correlate with the decline in Chinese manufacturing during the period and to have negative impacts on the Chinese Macroeconomy. In addition, the results corroborate that daily prices of
Chinese equity securities have an autoregressive component. Whether this is a permanent or temporary component of the time series requires a more exhaustive study involving long term modeling of financial time series as exemplified by Ray, Jarrett and Chen (1997) in a study of the Japanese stock market index. One very basic conclusion is that we indicated that the use of intervention analysis is very useful in explaining the dynamics of the impact of serious interruptions in an economy and the changes in the time series of a price index in a precise and detailed manner.

References


 Founded in 1892, the University of Rhode Island is one of eight land, urban, and sea grant universities in the United States. The 1,200-acre rural campus is less than ten miles from Narragansett Bay and highlights its traditions of natural resource, marine and urban related research. There are over 14,000 undergraduate and graduate students enrolled in seven degree-granting colleges representing 48 states and the District of Columbia. More than 500 international students represent 59 different countries. Eighteen percent of the freshman class graduated in the top ten percent of their high school classes. The teaching and research faculty numbers over 600 and the University offers 101 undergraduate programs and 86 advanced degree programs. URI students have received Rhodes, Fulbright, Truman, Goldwater, and Udall scholarships. There are over 80,000 active alumnae.

The University of Rhode Island started to offer undergraduate business administration courses in 1923. In 1962, the MBA program was introduced and the PhD program began in the mid 1980s. The College of Business Administration is accredited by The AACSB International - The Association to Advance Collegiate Schools of Business in 1969. The College of Business enrolls over 1400 undergraduate students and more than 300 graduate students.

**Mission**

Our responsibility is to provide strong academic programs that instill excellence, confidence and strong leadership skills in our graduates. Our aim is to (1) promote critical and independent thinking, (2) foster personal responsibility and (3) develop students whose performance and commitment mark them as leaders contributing to the business community and society. The College will serve as a center for business scholarship, creative research and outreach activities to the citizens and institutions of the State of Rhode Island as well as the regional, national and international communities.

The creation of this working paper series has been funded by an endowment established by William A. Orme, URI College of Business Administration, Class of 1949 and former head of the General Electric Foundation. This working paper series is intended to permit faculty members to obtain feedback on research activities before the research is submitted to academic and professional journals and professional associations for presentations.

An award is presented annually for the most outstanding paper submitted.