IMPACTS OF PERMANENT AND TEMPORARY SHOCKS ON OPTIMAL LENGTH OF MOVING AVERAGE TO FORECAST A TIME SERIES

By

YOONSUK LEE

Bachelor of Science in Agricultural Economics
Kangwon National University
Chuncheon, South Korea
2003

Master of Science in Agricultural Economics
Oklahoma State University
Stillwater, Oklahoma
2009

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2013

IMPACTS OF PERMANENT AND TEMPORARY

SHOCKS ON OPTIMAL LENGTH OF MOVING

AVERAGE TO FORECAST A TIME SERIES

Dissertation Approved:

Dr. B. Wade Brorsen

Dissertation Adviser

Dr. Brian Adam

Dr. Eric Devuyst

Dr. Jaebeom Kim

Outside Committee Member

.

ACKNOWLEDGEMENTS

Over the past four years, I have received support and encouragement from a great number of people. It is my great pleasure to thank all the people who made this dissertation possible. It is difficult to overstate my sincere gratitude to my advisor, Dr. B. Wade Brorsen. Without his patience, guidance and his great direction, I would never have been able to finish my dissertation.

I would like to thank my mother who is always in my heart and my brother who got me started in studying.

I would also like to thank my uncle and aunt for supporting and encouraging me. Finally, thanks to Maeda Kenta for sharing mathematical and statistical knowledge with me.

Name: Yoonsuk Lee

Date of Degree: MAY, 2013

Title of Study: IMPACTS OF PERMANENT AND TEMPORARY SHOCKS ON OPTIMAL LENGTH OF MOVING AVERAGE TO FORECAST A TIME SERIES

Major Field: Agricultural Economics

Abstract: Moving averages are often used for forecasting and the optimal length of the moving average depends on the size and frequency of structural breaks. A new time series model is proposed to describe permanent shocks related to structural breaks and temporary shocks with probability distributions. In the proposed model, permanent shocks are captured by a Poisson-jump or a Bernoulli-jump process, and temporary shocks are independent and identically normally distributed. This model requires a time series to have negative autocorrelation created by overdifferencing the temporary shocks. The proposed model is adapted to allow for positive autocorrelation by permitting autocorrelation of the jump process. The models are estimated with Oklahoma hard red winter wheat basis, Illinois corn basis and soybean basis, money stock, stock prices, total employment and total unemployment rate macroeconomic series. The parameters of the models are the probability of occurrence of jumps, the variance and the mean of the jump process, a time trend, and the variance of temporary shocks. The parameters are estimated with generalized method of moments estimation. In order to deal with autocorrelation in each series, we add an additional moment condition about autocorrelation to the generalized method of moments estimation. Most shocks are permanent shocks. The findings imply that shorter moving averages are the best for forecasting these series. The developed models are used to estimate the relative impacts of permanent and temporary shocks on the optimal length of moving average to use for forecasts. One year is the optimal length due to the large proportion of permanent shocks occur. The autoregressive integrated moving average (ARIMA) model with outliers is selected as a competing model. The proposed models for both a Poisson-jump model and a Bernoulli-jump model fit actual series better than the competing ARIMA models with outliers.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

Introduction

Forecasting is an important aim of time-series analysis. The occurrence of structural breaks increases uncertainty and decreases the accuracy of forecasts in time series analyses. A moving average is a common and simple method of forecasting a time series. A moving average forecast using historical data works best if the mean is constant or changes slowly or infrequently. In order to better understand and predict, several researchers have compared the forecasting performance of the simple moving average method and various regression models. Jiang and Hayenga (1997) applied several methods of forecasting corn and soybean basis and they found three-year moving averages worked relatively well. However, the three-year-average method modified with current market information and seasonal autoregressive integrated moving average (ARIMA) models outperformed the simple three-year-moving average method in out of sample forecasting tests. Sanders and Manfredo (2006) considered a variety of time series models and concluded that even when the time series model produced better forecasts than a five-year moving average, the accuracy gained from advanced time series models was relatively small.

While there is agreement that moving averages compete well with time series approaches, there is much less agreement about the length of moving average to use. The optimal length of moving average is sensitive to the effects of structural changes (or breaks). Hatchett, Brorsen, and Anderson (2010) argued that the longer moving averages are optimal when little or no structural change occurs, but shorter moving averages are optimal following structural changes. They suggested that when a structural change has occurred, the previous year's basis or an alternative approach should be used to forecast it. The existence of structural breaks in time-series analysis is a long-standing problem and a variety of approaches have been attempted to identify and to estimate the effects of structural breaks (Chen and Hong, 2012).

We take as a starting point the existence of permanent shocks associated with structural breaks in a time series and we assume permanent shocks produce leptokurtic features of series. The permanent shock should have a different distribution than temporary shocks. The study proposes a new approach of time-series analysis to estimate the probability and relative size of structural changes. The proposed model allows two types of shocks: transitory shocks and permanent shocks. Even though separating shocks into permanent and temporary effects has a long history, there is a key difference between the existing models to decompose temporary and permanent effects of shocks and the new time series process proposed in the study. The proposed stochastic processes are designed by choosing a probability distribution process to model permanent shocks. In traditional time series analysis, macroeconomic variables were separated into a trend component assumed deterministic and a cyclical component assumed to be transitory (Banerjee and Urga, 2005). Nelson and Plosser (1982) challenged such a traditional view that the trend is deterministic and augured that possible stochastic features of the trend caused by permanent shocks should be considered. In their paper, the timing of shocks that have permanent effects on the long-run level of most macroeconomic aggregates is known and the shocks are related to the Great Depression and the first oil-price crisis. A set of literature related to the structural breaks of Nelson and Plosser (1982) typically only considers one or two structural breaks and captures structural breaks using indicator variables. Perron (1989) similarly argued for the need to isolate some unexpected and discrete economic events and considered them as permanent effects of time series. These time series processes with structural breaks often have a completely different set of parameters before and after the break and the way they are identified will miss small breaks. While most post literature has considered one or two are exogenous structural breaks, several later literature has treated structural breaks as endogenous ones. Lumsdaine and Papell (1997) conducted unit-root tests with two endogenous breaks under the alternative hypothesis and found that the case against the random walk is strengthened. Lee and Strazicih (2003) allowed for two breaks under the null hypothesis, using a minimum Lagrange multiplier test. Campbell and Mankiw (1987) and Zivot and Andrews (1992) suggested that current shocks are a combination of temporary and permanent shocks, and that the long-run response of a series to a current shock depends on the relative importance or size of the two types of shocks. Pesaran, Pettenuzzo, and Timmerman (2006) develop a Bayesian forecasting approach that considers the possibility of structural breaks when the number of structural breaks is known. In most studies associated with structural breaks, these structural breaks are treated as dummy variables to indicate the absence or presence of their effects that may be expected to shift the level of series. In a number of studies dealing with structural breaks in time series analysis, inference may be organized in several different ways; stationary or non-stationary processes, known or unknown break points, multiple breaks or single break, estimation in single-equation or systems, or any of these in combination (Banerjee and Urga, 2005).

Structural breaks can be defined in terms of any parameter. Here we define a structural break as a change in the long-run mean. We develop a single stochastic process that can estimate the probability of break occurrences as well as a distribution for the size of structural breaks and therefore the idea that we impose a probability distribution for permanent shocks is necessary. The effects of permanent and temporary shocks are obvious in many time series but they are unobserved components in the series. In order to build a new stochastic process including permanent and temporary shocks, we use state space modeling that provides an explicit structural framework for the unobserved components. The proposed models are estimated, starting with Oklahoma hard red winter wheat basis, and Illinois corn basis and soybean basis for harvest. In addition, for the validation of the developed model with other data series, we use three more series (money stock, stock prices and total employment) out of fourteen macroeconomic series that Nelson and Plosser (1982) used to estimate impacts of structural breaks on unit root processes.

In the paper, the concept of structural changes is associated with abrupt breaks (or shocks) occurring at discrete time points within a given period and having permanent effects on a market. The potential source of shocks in time series can be a change in production and consumption, a change of government policies, technological advances, a change of transportation cost and weather, and so on. In the paper, we assume these permanent shocks cause the leptokurtic features that have higher peaks around the mean compared to a normal distribution, which leads to a distribution with fat tails. Several financial models incorporate the leptokurtic features to better match a financial time-series and to provide accurate forecasts. Merton (1976) introduced a jump-diffusion model that combines a Poisson-jump process for discrete shocks and a standard geometric Brownian motion diffusion process. The proposed model uses a Poisson jump process to model the frequency and size of permanent shocks associated with structural breaks, called jumps, which occur independently of one another. Transitory shocks are exploited by an independent and identically normal distribution. The resulting model differs from the jump-diffusion model of Merton (1976) since the normally distributed errors are all temporary shocks that cause negative autocorrelation in first differences. In addition to adapting a Poisson-jump process to describe permanent shocks, we estimate a Bernoulli-jump process that only allows for one permanent shock per observation period. The Bernoulli-jump process has the advantaged of nesting classic time-series models such as a random walk model with drift and a linear time trend model.

After we develop a new stochastic time series model, we consider how well the developed model fits the data. We select an ARIMA model with outliers as a competing model for the model calibration tests. In the ARIMA model with outliers, outliers can be a level shift outlier or a temporary shock that is referred to as a transient change outlier. Fox (1972) classified outliers as additive outliers or innovative outliers. Tsay (1988) suggested three classes of outliers: level shift, transient change, and variance change. An additive outlier (AO) affects only a single observation and after this disturbance, the series returns to its normal path as if nothing has happened. An innovative outlier (IO) is an unusual innovation affecting all later observations. A level shift (LS) outlier changes the level or mean of the series after a shift. A transient change outlier (TC) causes an initial impact like an additive outlier but the effect takes a few periods to disappear after a change occurred.

Variance change outliers affect the variance of the observed data by a certain magnitude. The combination of outliers could be used to estimate the effects of different shocks. We combine level shift and transient change outliers to match the concept of permanent and temporary shocks considered in the developed model.

One of our interests is to determine how well the developed models are calibrated to a data series. It is achieved by Kolmogorov-Smirnov test, testing a maximum distance between empirical cumulative distribution curves from historical and simulated time series. The test statistic of the KS test is the maximum difference of empirical cumulative density functions of the data and the estimated model. The model calibration test is not to assess whether a particular model is true, but rather understand which features of data it can explain because models are rejected by the data (Hnatkovska et al. 2012).

Objective

The general objective is to develop a new time series model to better describe permanent shocks and temporary shocks with appropriate probability distributions and based on the developed model, we determine the optimal length of moving average to use forecasts.

1. The first specific objective is to develop a new time series stochastic process to better describe the behavior of permanent and temporary shocks; a permanent shock is reflected by a Poisson jump process and a Bernoulli jump process, respectively and a temporary shock is explained by an independent and identically distributed normal distribution.

2. The second objective is to determine the effects of the relative importance of permanent shocks on the optimal length of moving averages to use in forecasting.
3. The last objective is to determine whether the proposed model is well calibrated through the indirect inference of comparing with ARIMA models with outliers.

The new stochastic model developed in this study is estimated, using a generalized method of moments (GMM) estimation. For GMM estimation, each moment equation is derived from first order condition of the assumed log-likelihood function of the developed model. Even in the presence of autocorrelation, maximum likelihood that does not consider the autocorrelation is still consistent. The existence of autocorrelation created by over-differencing a temporary shock is dealt with by adding an additional moment equation. Based on the developed model, we determine the impacts of permanent shocks on the optimal length of moving average to use for forecasting time series. In order to evaluate accuracy of forecasts, we use root mean squared error (RMSE). Lastly, we consider whether the developed model is well calibrated to the actual data. We select an ARIMA model with outliers as a competing model. The Kolmogorov-Smirnov test for two samples is used to test whether the estimated models are well calibrated.

Literature Review

There is a need to look for a new approach for time-series analysis. It is important to understand movement of a time series. Understanding structural breaks is essential to improve the accuracy of forecasting. We propose describing the structural breaks by an appropriate probability distribution. There are permanent shocks and temporary shocks.

Traditional time-series analyses have tended to distinguish between processes where shocks have a permanent effect and those where they do not. The distinction between stationary first-order autoregressive (AR (1)) processes, where all shocks are temporary, and the random walk process is a common way to determine permanent effects of shocks in the series (Engle and Smith, 1999). However, AR (1) process could not explain that a series responds not only to permanent shocks but also to temporary by modeling a process. Several studies have challenged this view by finding empirical evidence of not only permanent shocks but also temporary shocks in a series. Campbell and Mankiw (1989), Lee (1995), and Engle and Smith (1999) suggested that a current shock are a combination of temporary and permanent shocks, and that the long-run response of a series to a current shock depends on the relative importance or size of the two types of shocks. Beveridge and Nelson (1981) introduced a decomposition process of an autoregressive integrated moving average (ARIMA) for a univariate time series to separate effects of the permanent components representing long term change and temporary components reflecting short term changes. Fountis and Dickey (1986) extended this decomposition to vector autoregressive (VAR) models with a single unit root (stochastic trend). Stock and Watson (1988) generalized the process of Fountis and Dickey (1986) to cases having several unit roots. Clarida and Taylor (2003) proposed a decomposition method for univariate as well as multivariate nonlinear processes to analyze the permanent and temporary components, using real US GNP from 1947 to 1998. Although these decomposition processes of univariate time series found separate effects of the trend cycle and seasonality from permanent shocks, their estimation procedure depends on Gaussian white noise processes. In the study, the developed time-series models do not decompose temporary and permanent shocks, but instead we impose different probability distributions processes to reflect permanent and temporary shocks.

The permanent shocks related to structural breaks is defined as the long-horizon level forecast of the series, or the part that remains after all transitory dynamics have disappeared (Clarida and Taylor, 2003). All possible breaks occurring at the discrete time points in a series could be attributed to some random external variables to the series. Modeling impacts of permanent shocks in time-series analyses has become a key point. Several researchers have treated permanent shocks as indicator variables and these variables are removed from the noise function of the Nelson and Plosser (1982) data. With Nelson and Plosser (1982) data, Perron (1989) showed that the ability to reject a unit root decreases when the stationary alternative is true and an existing structural break is ignored in the process. He allowed for the presence of a one-time change in the level or in the slope of the trend function under both the null and alternative hypotheses. The one-time change could occur due to either the 1929 crash or the 1973 oil price shock. He used a modified Dickey-Fuller unit root test that includes a dummy variable to control one fixed structural break. Zivot and Andrews (1992) treated breakpoints as endogenous variables that are estimated rather than fixed, and determined the asymptotic distribution of the estimated breakpoint test statistic. They tested a unit-root that allows for the estimated breaks in the trend function under the alternative hypothesis. They found that there is less evidence against the unit-root hypothesis than Perron's (1989) finds for many of the data series, but for several of the series such as industrial production, nominal GNP, and real GNP there is stronger evidence against the unit-root hypothesis. Nunes, Newbold and Kuan (1997) allowed a structural break under both the null and alternative hypotheses, using Nelson and Plosser's (1982) data. In addition to permitting one structural break for both null and alternative hypotheses, they emphasized the importance of the size of Zivot and Andrews's (1992) breakpoint test statistic. Their testing procedure showed that the unit root hypothesis is failed to reject for any series at the 5% level, and for real GNP the hypothesis is rejected at the 10% level. These studies mainly consider as the cause of permanent changes either the Great Crash or the oil-price shock and both ones. The matter of their studies is whether the points of the occurrence of structural breaks are given or are estimated. In the study, we less focus on the time points occurring at structural breaks. We are interested in the probability of permanent shocks related to structural breaks in the given period and the distribution of a size of permanent shocks. Since Perron (1989) who suggested that it might need to isolate some unexpected and discrete economic events and considered them as permanent effects of time series, the implication of structural breaks when testing for unit root processes has been emphasized. However, the assumption of only one structural break given a time-period has implied the possibility of multiple structural breaks. Lumsdaine and Papell (1997) discussed that the inference about the break points themselves is less sensitive than inference about the assumption about the number of breaks, and therefore, the results about tests of the unit root hypothesis are sensitive to the number of breaks in the alternative specification. They tested the unit root hypothesis against an alternative of two breaks and rejected the unit root hypothesis at the 5% level for seven of the 13 series and for two more series at the 10% level. Lee and Strazicih (2003) argued the computation of critical values based on the assumption of no breaks under the unit root hypothesis (Zivot and Andrews ,1992; Lumsdain and Papell, 1997) might lead to the erroneous reject on of the unit root hypothesis. They allowed for two breaks under both the null and alternative hypotheses, and they concluded the rejection of the null hypothesis (a unit root with two breaks) clearly implies trend stationarity, using a minimum Lagrange multiplier test. Although they showed the processes imposing two endogenous breaks are more reasonable to analyze data series than those imposing a break, their models could not provide the possibility that a process containing breaks more than two is appropriate. Pesaran, Pettenuzzo, and Timmerman (2006) proposed a Bayesian forecasting approach that considers the possibility of structural breaks when the number of structural breaks is known. They considered the number of breaks as well as the size of breaks for the nominal three-month U.S. T-bill rate from July 1947 to December 2002 in their model and modeled the break process, using a hierarchical hidden Markov chain (HMC) approach under the assumption that the parameters within each break segment are drawn from some common meta-distribution. They found the HMC approach worked well in forecasting the series, however, when forecasting many periods ahead or when breaks occur relatively frequently, this approach is unlikely to show satisfactory forecasts. When we consider the case of more than one structural break, the distinction between a series with a unit root and a stationary series with nonconstant deterministic components is less clear (Hansen, 2001). In the paper, we assume that multiple structural breaks are captured by Poisson-jump process and also consider Bernoulli-jump process that takes value 1 with a success probability of one permanent shock and value 0 with a failure probability of no permanent shock.

With the importance of dealing with structural breaks in time series estimations, various approaches for detecting and handling structural breaks have been provided. In the paper, we have a particular interest in the existing method that detects and handles structural breaks in an autoregressive integrated moving average (ARIMA) process. An ARIMA model with outliers is selected as a competing model with the proposed model. There exists some controversy and not everyone agrees with the use of outlier methods, but, it is common to treat structural breaks the indicator variables in a time-series analysis (Proietti, 2008). The presence of structural breaks influences the autocorrelation structure of a time series, and therefore they easily mislead the conventional Box-Jenkins procedure (Tsay, 1988). The ARIMA model with outliers is to identify outliers and remove the impacts of outliers from a series to better understand the structure of a series (Chang, Tiao and Chen, 1988). Fox (1972) derived the likelihood ratio criteria for testing the presence of outliers and for identifying additive outlier and innovation outliers. Tsay (1988) suggested the structure breaks allow level shift and variance change, and level shift is classified as a permanent level change and a temporary change. Although the classification of outliers provides with advantages to determine the impacts of different outliers, there is an ambiguous identification between level shift and innovation outliers since a level shift and an innovative outlier are identical on a random walk. Sanchez and Pena (2003) argued the ARIMA model with outliers may misidentify level shifts as innovative outliers and this procedure may fail to identify patches of outliers due to the masking effects. Balke (1993) reported a level shift in a stationary time series is identified as an innovation outlier. In our study, however, the alternative procedures proposed by several studies to avoid this confusion is not an interesting subject. In order to match the concept of permanent and temporary shocks considered in our paper, we use a combination of level change and transient change. Level shifts and transient changes can have more serious effects on point forecasts even when outliers are not close to the forecast region (Trivez 1993).

These time-series models with structural break specifications often have a completely different set of parameters before and after the break and the way they are identified will miss small breaks. In addition, inference about unit roots depends on the number of break dates exogenously or endogenously permitted (Nunes, Newbold and Kuan, 1996). However, these time-series models with structural breaks could not address distributional features from variations in a price series. In order for the improved understanding of irregular events associated with structural breaks that produce fat-tailed distributions, we need to select an appropriate distribution for the data. Price variations within a various type of data series can present common empirical properties. These properties could be seen as various assumptions or constraints that a probabilistic model may provide better understandings of time-series behavior. Almost all financial asset prices such as US and worldwide stock indices, individual stocks, foreign exchange rates, interest rates, etc. display a high peak and asymmetric heavy tails (Kou 2008). This is, they have a leptokurtic distribution. In order to incorporate the leptokurtic features of financial prices, various models have been proposed in finance. Merton (1976) introduced the jump-diffusion model for the stock price. He categorized the total change in the stock price into normal and abnormal types. The normal variations in price are due to temporary disturbances between supply and demand, changes in capitalization rates, changes in the economic outlook, or other new information that causes incremental changes in the stock's value. In Merton (1976)'s jump-diffusion model, the temporary disturbance is modeled by a standard geometric Brownian motion with a constant variance per unit time and it has a continuous sample path. The abnormal variations in the stock price are due to the arrival of important new information about the stock that has more than a marginal effect on price. Such information arrives at discrete points in time by its nature. He treated this component as a "jump" process reflecting the non-marginal impact of the information. The assumption of discontinuities due to the discrete breaks is consistent with the observed leptokurtosis in the distributions of many financial variables (Hall et al. 1989).

In several studies associate with jump-diffusion models, jump-diffusion models have been adjusted by different methods. Kou (2002) proposed a double exponential jump-diffusion model for option prices. He differentiated the double exponential jump-diffusion model from Merton's jump-diffusion model, by assuming that jump sizes from a Poisson process are double exponentially distributed. The advantage of the double exponential jump-diffusion model over other financial models is that it can capture the asymmetric leptokurtic features as well as the volatility smile and lead to analytical solutions for many option prices. However, there is a weak point in jump-diffusion models. They cannot capture the volatility clustering effects that can be captured by other models such as a stochastic volatility model. An affine jump-diffusion model is one that combines jump-diffusion and stochastic volatility processes (Duffie and Singleton, 2000).

Most data in macroeconomics and finance come in the form of time series. Many finance models require often identical and independent random variables and assume constant variance, normal distribution and the absence of autocorrelation (Hull, 2005). To deal with an assumption of a constant variance in classical financial models, Bollerslev (1986) proposed a generalized autoregressive conditional heteroskedasticity (GARCH) process that can successfully account for the volatility of price being time dependent. However, the jump-diffusion-based models assume there exists small or no autocorrelation. Kou (2008) found the magnitude of autocorrelations in a variable is quite small, only about -0.05 to 0.05 in the daily closing prices of S&P 500 index from January 2, 1980 to December 31, 2005 and it is even smaller for weekly and monthly returns. Applying an assumption of zero autocorrelation to a jump-diffusion process is a common way to deal with autocorrelation. However, with time-series data, where the observations follow a natural ordering through time, there is a high possibility that successive errors will be correlated with each other. Ignoring the presence of autocorrelation in time-series models could lead to poor predictions in a short-run. Tomek and Myers (1993) discussed time-series properties of commodity prices. Features of commodity price behaviors over time could be categorized in immediate response by the high degree of positive autocorrelation in price levels and occasional breaks that can appear in a price distribution. Besides the effects of permanent shocks in a series, it is an important step to analyze the existence of autocorrelation in time-series analyses to show the effect of temporary shocks. In our study, the temporary shocks created by overdifferencing imply autocorrelation in a series. Therefore, a first-order moving average process models the autocorrelation of the change in a series in our study.

Appropriately modeling impacts of permanent shocks related to structural breaks is essential for time-series forecasts. Structural breaks could be the main source of forecast failure (Pesaran et al. 2006). For example, recently the Oklahoma red wheat and Illinois corn and soybean basis series have become more difficult to forecast. Figures 1 to 3 display Oklahoma red wheat and Illinois corn and soybean basis series for harvest. Several forecasting methods have been applied to improve the accuracy of basis forecasts. Jiang and Hayenga (1997) applied reported that the simple three –years-average forecast method outperformed several forecasting techniques with corn and soybean basis behavior in different locations. In out of sample tests, however, the three-years-average model supplemented with current market information and seasonal ARIMA model provided the best forecasts. Sanders and Manfredo (2006) compared the forecast abilities within a variety of time series models for soybeans, soybean meal, and soybean oil. They argued neither a 5-year average basis forecast commonly used may be the most accurate forecasting method nor the complicated time models do outperform the simple moving average methods. These studies have found favor in the simple moving average method than in other forecasting methods, and when current market information play an important role of improving accuracy of forecasting. In this study, the focus is on the proper length of moving average to use for forecasts when structural changes have been observed in a series.

The length of moving average could depend on the size or frequency of structural breaks. Hatchett, Brorsen, and Anderson (2010) discussed that the longer moving averages are optimal when little or no structural change occurs, but shorter moving averages are optimal following structural changes. They suggested that when a structural change has occurred, the previous year's basis or an alternative approach should be used to forecast it. For forecasting purposes a model of the stochastic process underlying the structure breaks address questions such as how often breaks are likely to occur over the forecasting sample, how large such breaks will be and at which dates they occur (Pesaran et al, 2006).

The proposed model is built to estimate a probability of occurrence of and size of permanent shocks related to structural shocks. If there exist permanent shocks captured by a Poisson process then a change of series responds according to the jump size distribution, while if there do not exist permanent shocks then the series follows a Gaussian white noise process.

CHAPTER II

Model Development

We consider the impacts of permanent and temporary shocks simultaneously on a series. Permanent and temporary shocks in our study are treated by different distributions; permanent shocks are reflected by a Poisson-jump process and a Bernoulli-jump process, and temporary shocks are represented by an independent identical normal distribution. In order to explain a stochastic process of permanent and temporary shocks, we adapt the framework of state space model. In state space analysis, the unobserved dynamic process at time t is referred to as the state of the time series (Commandeur and Koopman, 2007). The state is modeled in the state equation that is a key component in state space modeling. In the state equation, time dependencies in a time series are dealt with by letting the state at time $t + 1$ be a function of the state at time t (Commandeur and Koopman, 2007). An original permanent-jump and temporary shocks model is required to a series having negative autocorrelation because of overdifferencing temporary shocks. Thus, we adjust the original model for a series having positive autocorrelation. Since all series used in the study are dependent, it needs to prove the weak consistence of maximum likelihood estimation with dependent observations.

Jump-Diffusion Model

Jump diffusion processes are widely used in finance to model asset prices because asset return distributions tend to be leptokurtic having heavier tails than those of a normal distribution. The classical financial model, Black-Scholes assumes geometric Brownian motion and thus cannot capture leptokurtic feature of a distribution (Kou, 2008). In order to incorporate the leptokurtic feature, many alternative models have been proposed. The jump-diffusion model is one of them and this model has been applied successfully with stock and foreign currency prices displaying large price changes over a small time interval (Hilliard and Reis, 1998). Merton (1976) discussed the total change in the stock price is categorized into two types of changes: normal and abnormal. The normal variations in price are modeled by a standard geometric Brownian motion with a constant variance per unit of time and it has a continuous sample path. The abnormal variations in price are represented by a jump process reflecting discrete breaks in time. Jump-diffusion models proposed by Merton (1976) could be expressed as (Kou, 2002):

(1) $$X_t = \sigma W_t + \mu t + \sum_{i=1}^{N_t} Y_i,$$

where W_t is a standard Brownian motion with $W_0 = 0$, N_t is a Poisson process with average probability κ , constants μ and σ are the drift and volatility of the diffusion part respectively, jump size Y_i are independent and identically distributed random variables, and the random process W_t, N_t, and random variable Y_i are assumed to be independent. Many finance models often require identical and independent random variables and assume constant variance, absence of autocorrelation and a normal distribution (Hull, 2005). Especially, the restriction of zero autocorrelation is commonly applied to finance models. However, most data in macroeconomics and finance come in the form of time series. Ignoring the presence of autocorrelation in time-series models could lead to erroneous results.

Permanent-Jump and Transitory Shocks Model

With the characteristic structure of state space models, describe a series of unobserved values, $a_1, \ldots, a_t$ called the states, with a set of observations, $x_1, \ldots, x_t$, and the states can be specified in a state equation (Commandeur and Koopman, 2007, Durbin and Koopman, 2012). The level component model of the state space model can be formulated as:

(2) $$x_t = a_t + v_t,$$

(3) $$a_t = a_{t-1} + \tau_t,$$

where a_t is the unobserved level at time t, $t = 1, \ldots, n$, if a_t does not change from time to time then it becomes an intercept in a regression model, v_t is the observation disturbance and τ_t is the level disturbance, these two error terms are assumed to be independently, identically and normally distributed with zero mean and variances $\sigma_v^2$and σ_τ^2, respectively, The equation (2) is called the observation or measurement equation, and the equation (3) is called the state equation. This local level model could be a random walk model if the variance of v_t is equal to zero.

The similarity between a local level model and a random walk model could be shown as:

(4) $$x_t = a_{t-1} + \tau_t,$$

(5) $$\Delta x_t = \tau_t,$$

where a_{t-1} is equal to x_{t-1}, Δx_t is the first differences of x_t. The equation (5) implies that the local level model is equivalent to an ARIMA (1,1,0) model. However, Commandeur and Koopman (2007) discussed differences between ARIMA model and state space approaches to time series analysis. While ARIMA models are concerned with the short-term dynamics only and thus are primarily concerned with forecasting only, the state space models provide an explicit structural (state equation) for the decomposition of time series in order to diagnose all the dynamics in the time series data simultaneously.

The study adapts the framework of the local level model. We assume that a Poisson distribution provides a more proper model for the features of permanent shocks related with structural breaks that yield a fat-tailed distribution. Based on the local level model, we build the permanent-jump and temporary–diffusion model. As an alternative model of commodity prices, the paper assume that each of permanent shock and temporary shock produces different impacts on a market and thus we impose different distributional form on permanent and temporary shocks, respectively. A permanent shock occurs at discrete time points and the impact of the shocks remains on the market. This permanent shock is modeled by a Poisson-jump process and a Bernoulli-jump process, respectively. Due to overdifferencing temporary shocks, the temporary error term implies a first-order moving average. The temporary shock follows an independent and identical normal distribution.

For the model mixed with the Poisson-jump process, a combination of the two different shocks is expressed by the state space approach:

(6) $$Series_t = \mu_t + \varepsilon_t,$$

(7) $$\mu_t = \mu_{t-1} + \gamma + \sum_{q=1}^{Q_t} Jump_{qt},$$

where $Series_t$ is data series at time t, μ_t is an unobserved variable which assumed to include unexpected and discrete structural changes, ε_t is the transitory shocks and follows independent and identical normal distribution with a zero mean and a variance, σ_ε^2, and $E(\varepsilon_t \varepsilon_{t-1}) = 0$, γ is a drift, $Jump_{qt}$ follows independent and identical normal distribution with a mean, μ_J, and a variance, σ_J^2, Q_t represents the number of permanent shocks occurring in a given time period, t, and follows a Poisson process with parameter λ which is a probability of structural breaks, and ε_t and Q_t are independent. The first difference is applied to convert the stochastic processes to achieve stationarity. One risk of this data transformation is the possibility of overdifferencing. First differencing creates a first order moving average (MA) process from the temporary shocks. The resulting data generating process can be expressed as:

(8) $$\Delta Series_t = Series_t - Series_{t-1} = \gamma + \sum_{q=1}^{Q_t} Jump_{qt} + \nu_t,$$

$$\nu_t = \varepsilon_t - \varepsilon_{t-1},$$

where the autocorrelated error ν_t is replaced by the stationary and non-autocorrelated error ε_t. The relationship between ν_t and ε_t is given by a non-invertible first-order moving average process. In the paper, all series are transformed to the first difference: $\Delta series_t = series_t - series_{t-1}$. The difference transformation in the proposed model induces a non-invertible moving average process ($\nu_t = \varepsilon_t - \varepsilon_{t-1}$) in the transformed model. We face the problems associated with estimating the proposed model containing the non-invertible moving average process based on the likelihood function. A number of researchers have studied the properties of different estimators of a non-invertible moving average process. In practice, the most common approach to the non-invertible moving average process is to set the initial condition, $\varepsilon_0 = 0$, because the conditional maximum likelihood estimator for a parameter of first-order autocorrelation merely requires finding the value of the parameter which minimizes the sum of squares function (Pierce, 1971). Since the initial condition has diminishing influence on series as time gets large for invertible processes, all of the estimators share the large sample distributional properties of the likelihood estimator. Plosser and Schwert (1997) argued that the efficiency properties of the error term in the regression equation are affected by the difference transformation, but the values of the regression coefficients are not substantially affected by overdifferencing if an MA parameter is estimated from the difference.

Ball and Torous (1983) introduce a model mixed with a Bernoulli jump process model for jump features of stock prices instead of a model mixed with a Poisson jump process. A Poisson process counts the number of events occurring in a given time period and a Poisson process in a jump-diffusion model might capture all possible type of discrete events. But, such discontinuous events ought not to be very often. A Bernoulli process is a finite or infinite sequence of binary random variables that take only two values, canonically 0 (no event) and 1 (one event). The Bernoulli process in a jump-diffusion model is that over a given time period either no event occurs in a price series or one event occurs with probability λ. A merit of this Bernoulli-jump and temporary-diffusion model is that it could nest traditional time-series models. From the equations (9) and (10), if the probability of one permanent shock (P) is equal to zero and the variance of temporary shocks (σ_e^2) is equal to zero and the mean (μ_B) is equal to zero then the Bernoulli-jump model nests a linear time trend. If $P = 1$ then the Bernoulli-jump model nests a random walk model.

The data generating process mixed with a Bernoulli process is expressed as :

(9) $$Series_t = a_t + e_t,$$

(10) $$a_t = a_{t-1} + \beta + B_t J_t,$$

where $Series_t$is a series at time t, a_t is an unobserved variable which assumed to explain a permanent shock occurring in a given period of time, t, e_t are the transitory shocks and follows an independent and identical normal distribution with a zero mean and a variance, σ_e^2, and $E(e_t e_{t-1}) = 0$, β is a drift, J_t follows an independent and identical normal distribution with a mean, μ_B and a variance, σ_B^2, B_t represents one permanent shock in a fixed time period and follows a Bernoulli (P) process, P is the probability of one permanent shock in a given time interval, t , and B_t and J_t are independent. From the model mixed with the Bernoulli process, the data generating process for a change in a series is:

(11) $$\Delta series_t = series_t - series_{t-1} = \beta + B_t J_t + u_t,$$

$$u_t = e_t - e_{t-1}.$$

The relationship between u_t and e_t is given by a first-order moving average process. A merit of this Bernoulli-jump and temporary-diffusion model is that it nests traditional time-series models. From the equation (9), if the probability of one permanent shock (P) is equal to zero and the variance of temporary shocks (σ_e^2) is equal to zero and the mean

(μ_B) is equal to zero then the Bernoulli-jump model becomes a linear time trend model. If $P = 0$ and $\sigma_e^2 = 0$ then the Bernoulli-jump model becomes a random walk model.

In Poisson-jump and Bernoulli-jump models derived above, the overdifferencing of the shock term creates negative autocorrelation in the first differenced series. Since the requirement of negative autocorrelation is too restrictive for other time series, we consider the case where there exist positive autocorrelation in a series and thus we assume autocorrelation in permanent shocks as well. The state equations (7) and (10) are adjusted to consider both cases of autocorrelation.

From equation (7) for the Poisson-jump process, the data generating process is rewritten as:

(12)
$$\mu_t = \mu_{t-1} + \gamma^* + \sum_{q=1}^{Q_t} Jump_{q,t} + \rho \cdot \sum_{q=1}^{Q_{t-1}} Jump_{q,t-1},$$

where ρ is autocorrelation at lag one of $\sum_{q=0}^{Q_t} Jump_{q,t}$. The adjusted data generating process for positive autocorrelation is:

(13)
$$\Delta Series_t = \gamma^* + \sum_{q=1}^{Q_t} Jump_{qt} + \rho \cdot \sum_{q=1}^{Q_{t-1}} Jump_{q,t-1} + \varepsilon_t^* - \varepsilon_{t-1}^*.$$

From equation (10) for the Bernoulli-jump process, the adjusted data generating process is expressed as:

(14)
$$a_t = a_{t-1} + \beta + B_t J_t + \eta \cdot B_{t-1} J_{t-1}.$$

where η is autocorrelation at lag one. The adjusted data generating process is:

(15) $$\Delta series_t = \beta^* + B_t J_t + \eta \cdot B_{t-1} J_{t-1} + e_t^* - e_t^*.$$

Autoregressive Integrated Moving Average Model with Outliers

The method for detecting outliers in autoregressive integrated moving average processes is based on Tsay (1988)'s five classifications of outliers. We especially focus on outliers for level shift and transient change because of the in similarity to the proposed model. The autoregressive integrated moving average (ARIMA) process for univariate time series model could be written as:

(16) $$\Delta^d Y_t \Phi(L) = \Theta(L) u_t,$$

where Y_t is a data series, Δ is the difference operator of degree d, L is the lag operator such that $LY_t = Y_{t-1}$, $\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ and $\Theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$ are polynomials in L of degrees p and q, respectively.

Based on the outlier specification of Tsay (1988), after allowing for the existence of outliers, we consider a series, Z_t, contaminated by outliers instead of the series, Y_t. The contaminated series, Z_t, could be expressed as:

(17) $$Z_t = g(t) + Y_t,$$

where $g(t)$ is a function representing the exogenous distribution of Y_t such as different types of outliers. Tsay (1988) suggested that the function $g(t)$ could be deterministic or stochastic depending on the types of disturbances. He considered the stochastic case only for the variance change of outliers. In the paper, we estimate the deterministic model of

$g(t)$ as a competing model since we focus on the level shift and transient change of outliers. The deterministic function of $g(t)$ can be written as:

(18) $$g(t) = w_0 \cdot \frac{\Phi(L)}{\Theta(L)} \cdot \zeta_t^{(D)},$$

where $\frac{\Phi(L)}{\Theta(L)}$ defines the characteristic of outliers, $\zeta_t^{(D)}$ is a dummy variable for outliers occurring at time point D, $\zeta_t^{(D)} = 1$ if $t = D$ and $\zeta_t^{(D)} = 0$ if $t \neq D$.

The combination of level shift and transitory changes in an ARIMA process with outliers is used due to the similarity to the concept of permanent and temporary shocks in the developed model. According to Tsay's (1988) classification of outliers, for a level shift (LS) outlier, when $w_0 = w_{LS}$ and $\frac{\Phi(L)}{\Theta(L)} = \frac{1}{(1-L)}$, this is a level change model because $Z_t = Y_t$ for $t < D$ but $Z_t = w_{LS} + Y_t$ for $t \geq D$. The model says that a level shift of magnitude w_L occurs at time $t = d$ and the change is permanent. For a transient change (TC), when $w_0 = w_{TC}$ and $\frac{\Phi(L)}{\Theta(L)} = \frac{1}{(1-\theta L)}$ where $0 < \theta < 1$. This model describes a disturbance that affects Y_t for $t \geq D$. However, the effect decays exponentially with rate θ and initial impact w_{TC} and thus the change is temporary.

Weak Consistency for Maximum Likelihood Estimation with Dependent Observations

Observations are independent if the sampling of one observation does not affect another observation, however, when we are dealing with time series, it is likely that there is some relationships between a given time series and a lagged time series over consecutive periods. In this paper, we face a case in which the observations are not independent. In order to show the consistency of the known joint density of observations in the proposed models, we adopt Heijmans and Magnus (1986)'s Theorem 2 shows that consistent maximum likelihood estimation with dependent observations. Heijmans and Magnus (1986) proposed two theorems on the weak consistency of the maximum likelihood (ML) estimator obtained from generally dependent observations under the assumption that the joint density of the observations is known.

According to their arguments, the Theorem1 contains conditions that are necessary as well as sufficient, while the conditions for Theorem 2 are somewhat stronger but more readily applicable than Theorem 1. In theorem 2, they added a condition of the normalizing function for the behavior of log-likelihood ratio.

First, some notations are defined based on Heijmans and Magnus (1986). $\mathbb{N} = \{1,2,\dots\}$ and $\mathbb{R}^h$ is the Euclidean space of dimension $h > 1$. $N(\theta)$ is a neighborhood of a point $\theta \in \Gamma \subset \mathbb{R}^h$ is an open subset of Γ which contains θ. Ɓ means *Borel* measurable. E and var are mathematical expectation and variance. Let $\{\Delta series_1, \Delta series_2, \dots\}$ be a sequence of random time series variables, not necessarily independent or identically distributed. For each $t \in \mathbb{N}$, let $\Delta series_t = (\Delta series_1, \Delta series_2, \dots, \Delta series_t)$ be defined on the probability space $(\mathbb{R}^n, \mathcal{B}_t, P_{t,\theta})$ with values in $(\mathbb{R}^t, \mathcal{B}_t)$, where $\mathcal{B}_t$ denotes the minimal *Borel* field on $\mathbb{R}^t$ and P is the probability of a sequence of $\Delta series_t$.

Before discussing the consistent estimation with the known likelihood function of dependent observations, we need to discuss that the given joint density functions are measurable. We assume that a likelihood function $f_t(\Delta series_t\,;\,\theta)$ of the permanent-jump and temporary-diffusion model is known. For every $\Delta series_t \in \mathbb{R}^t$, the real-valued likelihood function,

$$\Lambda_t(\theta) = \Lambda_t(\theta; \Delta series_t) = \log L_t\,(\Delta series_t\,;\,\theta), \quad \theta \in \Gamma,$$

and $\Lambda_t(\theta) = \log L_t(\theta)$ is given in the paper. The true (but unknown) value of $\theta \in \Gamma$ is denoted by θ_0. An MLE estimate of θ_0 is a value $\hat{\theta}_t(\Delta series_t) \in \Gamma$ for every $\Delta series_t \in \mathbb{R}^t$ with

$$L_t\big(\hat{\theta}_t(z); z\big) = \sup_{\theta \in \Gamma} L_t(\theta; z).$$

Heijmans and Magnus (1986) discussed that since the supremum is not always attained, and thus the values of $\sup_{\theta\in\Gamma} L_n(\theta; z)$ does not necessarily exist everywhere on $\mathbb{R}^t$. However, if Γ is a compact subset of $\mathbb{R}^h$, then $\sup_{\theta\in\Gamma} L_n(\theta; z)$ always permits solution and the function, $L_t\big(\hat{\theta}_t(z); z\big)$, can be chosen as a measurable function. If there exists a measurable function $\hat{\theta}_t$ from $\mathbb{R}^t$ into Γ such that $\sup_{\theta\in\Gamma} L_t(\theta; z)$ holds for every $\Delta series_t \in \mathbb{R}^t$, an MLE estimator of $\theta_0 \in \Gamma$ exists surely.

The basic assumptions of the measurability of the assumed likelihood function for the proposed models are 1) $L_t(\Delta series_t\,;\,\theta)$ is a measurable function of $\Delta series_t$ for $\theta \in \Gamma$,; 2) $L_t(\theta; \Delta series_t)$ is continuous function of θ for every $\Delta series_t \in \mathbb{R}^t$; 3) Γ is a compact subset. In addition to the assumptions of the measurability, several estimating sequences $\{\hat{\theta}_n\}$ are allowed.

In order to obtain the weak consistency of an MLE with serially correlated observations under the known likelihood function, we adopt other two conditions from Heigmans and Magnus's Theorem 2 with basis measurability conditions:

A. for every $\theta \in \Gamma$, $\theta \neq \theta_0$, there exists a sequence of non-random non-negative quantities $k_t(\theta, \theta_0)$, which may depend on θ and θ_0, such that
   a) $\lim_{t\to\infty} \inf k_t(\theta, \theta_0) > 0$,
   b) $\operatorname{p}\lim_{t\to\infty} \left(\frac{1}{k_t(\theta,\theta_0)}\right)\left(\Lambda_t(\theta) - \Lambda_t(\theta_0)\right) = -1$;

B. for every $\theta \neq \theta_0 \in \Gamma$ there exists a neighborhood $N(\theta)$ of θ such that
   $\lim_{n\to\infty} \mathrm{P}\left[\left(\frac{1}{k_t(\theta,\theta_0)}\right) \sup_{\phi\in N(\theta)}\left(\Lambda_t(\theta) - \Lambda_t(\theta_0)\right) < 1\right] = 1$.

The condition A is supportive one for the proposed models. Heijmans and Magnus (1986) discussed the use of Kullback-Leibler information as normalizing function, k_t. The Kullback-Leibler information is criterion of evaluating model's similarity. The normalizing function k_n is not required to be continuous in either θ and θ_0, and is expressed as:

(19) $$k_t(\theta, \theta_0) = -\mathrm{E}\left(\Lambda_t(\theta) - \Lambda_t(\theta_0)\right).$$

For the developed model in the study, the conditions A and B are applied to prove the weak consistency of MLE with a serially correlated observation. Each series used in the study is not independent of each other. With a dependent time series, a density function can be expressed:

(20)
$$l_t^*(\Delta series_1, \dots, \Delta series_t) = \prod_{i=1}^{t} l_i^*(\Delta series_i, \boldsymbol{\theta}^*) \neq \prod_{i=1}^{t} l_i(\Delta series_i, \boldsymbol{\theta}).$$

where a $l_i(\Delta series_i, \boldsymbol{\theta})$ is a log-likelihood function with an independent time series and it is given, $l_i^*(\Delta series_i, \boldsymbol{\theta})$ is a log- likelihood function with a dependent time series, $l_i(\Delta series_i, \boldsymbol{\theta})$ defines a distribution on $\mathbb{R}^n$ corresponding $\boldsymbol{\theta}$ in Ω, where Ω is a matrix space and a subset of $\mathbb{R}^s$, $n \ni s$, $\boldsymbol{\theta}$ is a vector of parameters from MLE with an independent series, and define $\boldsymbol{\theta}^*$ is a vector of parameters from MLE with a dependent variable. The normalized $k_t(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ function can be rewritten as:

(21)
$$-\mathrm{E}\big(l_t(\boldsymbol{\theta}) - l_t^*(\boldsymbol{\theta}^*)\big) = -\mathrm{E}\big(l_t(\boldsymbol{\theta})\big) + \mathrm{E}\big(l_t^*(\boldsymbol{\theta}^*)\big),$$
$$= k_t(\boldsymbol{\theta}, \boldsymbol{\theta}^*),$$

where $k_t(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \to \infty$ for every $\boldsymbol{\theta} \neq \boldsymbol{\theta}^*$ based on the given assumption, $\lim_{n\to\infty} \inf k_t(\boldsymbol{\theta}, \boldsymbol{\theta}^*) > 0$, in the condition A. Thus, if $t \to \infty$ for every $\boldsymbol{\theta}^* \neq \boldsymbol{\theta}$ then

(22)
$$\frac{var\big(l_t(\boldsymbol{\theta}) - l_t^*(\boldsymbol{\theta}^*)\big)}{k_t(\boldsymbol{\theta}, \boldsymbol{\theta}^*)} \to 0,$$

and the condition B holds for the developed models. We do not fully complete the proof of weak consistency of MLE with a dependent observation and we leave a more complete proof to later works.

CHAPTER III

Procedure

We estimate the permanent-jump and temporary-diffusion model proposed in the paper by a generalized method of moments (GMM) procedure. The generalized method of moments (GMM) has become an essential estimation procedure in various areas of applied economics and finance since Hansen (1982) introduced the power of the GMM estimators with statistical theory (Jagannathan, et al. 2002). Hansen and Hodrick (1980) and Hansen and Singleton (1982) showed important applications of the GMM approach in the case where time-series data are used through their empirical analyses of foreign exchange markets and asset pricing, respectively. For the new stochastic time-series process proposed in the study, we apply the GMM framework for estimation. For GMM estimation of parameters of the developed models, we adapt an alternative approach for generating moment conditions proposed by Gallant and Tauchen (1996). Their idea of generating moment conditions is to use the derivative of the log density of a given model with respect to the parameters of the model.

Since we assume permanent shocks follow both a Poisson-jump process and a Bernoulli-jump process, the proposed models rely on a particular probability density function and distributional assumption. Having restricted the distribution of error terms, we first apply the maximum likelihood estimation (MLE) to estimate parameters of the proposed model. If the proposed model closely approximates the distribution of observed data then the hypothetical parameter vectors, $\tilde{\boldsymbol{\theta}}$ for a model mixed with Poisson jump process and $\tilde{\boldsymbol{\xi}}$ for a model mixed with a Bernoulli jump process, equal the true parameter vector, $\boldsymbol{\theta}$ (or $\boldsymbol{\xi}$) , and the hypothetical density function of first difference series, $f\left(\Delta series_t; \tilde{\boldsymbol{\theta}} \text{ (or } \tilde{\boldsymbol{\xi}})\right)$, becomes the true one. The MLE estimate of the unknown true parameter vector, $\boldsymbol{\theta}$ (or $\boldsymbol{\xi}$), is the $\tilde{\boldsymbol{\theta}}$ (or $\tilde{\boldsymbol{\xi}}$) that maximizes the likelihood function. The maximization is equivalent to maximizing the log likelihood function because the log transformation is a monotone transformation. With the MLE method, however, we ignore autocorrelation in a series. Harris (1999) argued that moving average (MA) terms in a time series model complicates the estimation problem since the least squares are no longer linear in the parameters and thus an MLE estimation of the time-series model with MA components face computational difficulty to obtain numerical optimization. Since the proposed time-series models involve a MA term created by overdifference of temporary shocks, the GMM estimation is applied to handle autocorrelation cause by the MA term.

The data generating processes for Poison-jump (equation (8)) and Bernoulli-jump (equations (11)) induce negative autocorrelation in first differenced series. Therefore, the models only require a time series having negative autocorrelation and it is too restrictive for a vast number of time series. Therefore, we adjust the developed models for positive autocorrelation; equation (12) for the Poisson-jump model and equation (13) for the Bernoulli-jump model. Each model has its log likelihood function but the log likelihood function is based on the case where autocorrelation is ignored.

The log likelihood function of the permanent-Poisson jump and temporary-diffusion model for a series having negative autocorrelation can be expressed as:

(23)
$$l(\tilde{\boldsymbol{\theta}}, \Delta series_t) = \sum_{i=1}^{T} log\left[\sum_{q=0}^{\infty} \frac{e^{-\tilde{\lambda}}\tilde{\lambda}^q}{q!} \frac{1}{\sqrt{2\pi(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2)}} exp\left(\frac{-(\Delta series_t - \tilde{\gamma} - q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2)}\right)\right],$$

where $l(\tilde{\boldsymbol{\theta}}, \Delta series_t)$ is the log likelihood function mixed with Poisson and normal distributions, $\tilde{\boldsymbol{\theta}}$ is a vector of five parameters $(\tilde{\gamma}, \tilde{\sigma}_\varepsilon^2, \tilde{\mu}_J, \tilde{\sigma}_J^2, \tilde{\lambda})$ estimated from the permanent-jump and transitory-diffusion model, $\tilde{\lambda}$ is an average of jump probability measuring the occurrence rate of discrete structural breaks by Poisson distribution, $\tilde{\sigma}_J^2$ is variance and $\tilde{\mu}_J$ is mean of jump process, respectively, $\tilde{\gamma}$ is a drift and $\tilde{\sigma}_\varepsilon^2$ is a variance of transitory shocks, and the variance of transitory shocks has to be doubled $(2\tilde{\sigma}_\varepsilon^2)$ due to the moving average term, $\varepsilon_t - \varepsilon_{t-1}$. In order for a series having positive autocorrelation, the log-likelihood function of the adjusted model is:

(24)
$$l^*(\tilde{\theta}^*, \Delta series_t) = \sum_{i=1}^{T} \log\left[\sum_{q=0}^{\infty}\sum_{i=0}^{\infty} \frac{e^{-\tilde{\lambda}^*} \cdot (\tilde{\lambda}^*)^q}{q!} \cdot \frac{e^{-\tilde{\lambda}^*} \cdot (\tilde{\lambda}^*)^i}{i!} \cdot \frac{exp\left(\frac{-(\Delta series_t - \tilde{\gamma}^* - \tilde{\mu}_j^* - i \cdot \tilde{\rho} \cdot \tilde{\mu}_j^*)^2}{2(2 \cdot \tilde{\sigma}_\varepsilon^{2*} + q \cdot \tilde{\sigma}_J^{2*} + i \cdot \tilde{\rho} \cdot \tilde{\sigma}_J^{2*})}\right)}{\sqrt{2\pi(2 \cdot \tilde{\sigma}_\varepsilon^{2*} + q \cdot \tilde{\sigma}_J^{2*} + i \cdot \tilde{\rho} \cdot \tilde{\sigma}_J^{2*})}}\right],$$

where $l^*(\tilde{\theta}^*, \Delta series_t)$ is the likelihood function, $\tilde{\theta}^*$ is a vector of six parameters, $(\tilde{\gamma}^*, \tilde{\mu}_j^*, \tilde{\sigma}_\varepsilon^{2*}, \tilde{\sigma}_J^{2*}, \tilde{\lambda}^*, \tilde{\rho})$, $\tilde{\rho}$ is a parameter of autocorrelation between permanent shocks.

For the model combined with a Bernoulli-jump process, the log-likelihood function without considering the autocorrelation is:

$$ll\left(\tilde{\boldsymbol{\xi}}, \Delta series_t\right) = \sum_{t=1}^{T} log\left[\sum_{j=0}^{1} \tilde{P}^j (1-\tilde{P})^{1-j} \frac{exp\left(\frac{-\left(\Delta series_t - \tilde{\beta} - j \cdot \tilde{\mu}_B\right)^2}{2(2\tilde{\sigma}_e^2 + j \cdot \tilde{\sigma}_B^2)}\right)}{\sqrt{2\pi(2\tilde{\sigma}_e^2 + j \cdot \tilde{\sigma}_B^2)}}\right], \tag{25}$$

where $\tilde{\boldsymbol{\xi}}$ is a vector of five parameters $(\tilde{\beta}, \tilde{\sigma}_e^2, \tilde{\mu}_B, \tilde{\sigma}_B^2, \tilde{P})$ estimated from the permanent-jump and transitory-diffusion model that a jump is represented by a Bernoulli distribution. $\tilde{P}$ is jump probability of one discrete event captured by a Bernoulli distribution, $\tilde{\sigma}_B^2$ and $\tilde{\mu}_B$ are the variance and mean of one permanent shock, respectively, and $\tilde{\beta}$ is a drift and $\tilde{\sigma}_e^2$ is a variance of temporary shocks. For the adjusted model with Bernoulli-jump process, the log-likelihood function can be expressed as:

$$ll^*\left(\tilde{\boldsymbol{\xi}}^*, \Delta series_t\right) = \sum_{t=1}^{T} \log\left[\sum_{j=0}^{1}\sum_{r=0}^{1} \left(\tilde{P}^*\right)^j (1-\tilde{P}^*)^{1-j} \cdot \left(\tilde{P}^*\right)^r (1-\tilde{P}^*)^{1-r} \cdot \frac{exp\left(\frac{-\left(\Delta series_t - \tilde{\beta}^* - j \cdot \tilde{\mu}_B^* - r \cdot \tilde{\eta} \cdot \tilde{\mu}_B^*\right)^2}{2(2 \cdot (\tilde{\sigma}_e^*)^2 + j \cdot (\tilde{\sigma}_B^*)^2 + r \cdot \tilde{\eta} \cdot (\tilde{\sigma}_B^*)^2)}\right)}{\sqrt{2\pi(2 \cdot (\tilde{\sigma}_e^*)^2 + j \cdot (\tilde{\sigma}_B^*)^2 + r \cdot \tilde{\eta} \cdot (\tilde{\sigma}_B^*)^2)}}\right] \tag{26}$$

where $ll^*\left(\tilde{\boldsymbol{\xi}}^*, \Delta series_t\right)$ is the likelihood function, $\tilde{\boldsymbol{\xi}}^*$ is a vector of six parameters, $\left(\tilde{\beta}^*, \tilde{\mu}_B^*, \tilde{\sigma}_e^{2*}, \tilde{\sigma}_B^{2*}, \tilde{P}^*, \tilde{\eta}\right)$, $\tilde{\eta}$ is a parameter of autocorrelation at lag one of permanent shock.

Based on the log likelihood function of the models, we compute first-order conditions and these first order conditions become moment equations for GMM estimation. Each moment equation is derived from the known log-likelihood function based on the Gallant and Tauchen (1996)'s approach. Gallant and Tauchen (1996) proposed an alternative way to generate a moment equation. Gallant and Tauchen (1996) defined the score generator that is a model assumed to be close to true model to compute moment conditions for a GMM estimator. According to their definition of the score generator, the observed data $\{\tilde{y}_t, \tilde{x}_t\}_{t=1}^{\infty}$ are assumed to have been generated from the sequence of densities, $\{P_1(x_1|\varsigma^o), \{P_t(y_t|x_t, \varsigma^o)\}_{t=1}^{\infty}\}_{\varsigma\in R}$ where ς^o is true value of the parameter ς and R is the parameter space. The given model is said to be smoothly embedded within the score generator, $\{f_1(x_1|\theta), \{f_t(y_t|x_t, \theta)\}_{t=1}^{\infty}\}_{\theta\in R}$ for every $\varsigma \in R$ and $P_1(x_1|\varsigma) = f_1(x_1|g(\varsigma))$ for every $\varsigma \in R$. Under the condition of a score generator, they derived the first order condition of the log density of a score generator function with respect to the parameters of the score generator function. In the paper, we assume that a series is generated from the mixture densities of Poisson (or Bernoulli) and normal. Based on the log-likelihood function assumed true, we generate moment conditions through applying Gallant and Tauchen's (1996) approach. There is an advantage to used Gallant and Tauchen's (1996) approach is that the estimator is nearly fully efficient even though a given parametric model does not require perfectly nesting the true model for a time-series.

We address the choices of the number of moments. For series having negative autocorrelation, we obtain five parameter estimates, using MLE estimation; a drift (γ), a variance (σ_ε^2) of a diffusion process, and a probability (λ) of permanent shocks, mean (μ_J) and variance (σ_J^2) of a jump process. For series having positive autocorrelation, we obtain six parameters from the adjusted model; a drift (γ^*), a variance (σ_ε^{2*}) of a diffusion process, and a probability (λ^*) of permanent shocks, mean (μ_J^*) and variance (σ_J^{2*}) of a jump process, autocorrelation (ρ) between jump processes. We first choose moments matching parameters. Each moment equation is an expectation of a first order condition of the given log-likelihood function in the paper:

(27)
$$\bar{m}_l(\widetilde{\boldsymbol{\theta}}) = \frac{1}{N}\sum_{t=1}^{N}\frac{d}{d\theta}l(\widetilde{\boldsymbol{\theta}}, \Delta series_t),$$

where $\bar{m}_l(\theta)$ is an expectation of ith moment equation, $l = \{1, \dots, 5\}$, each moment equation is evaluated at $\widetilde{\boldsymbol{\theta}} = (\tilde{\gamma}, \tilde{\sigma}_\varepsilon^2, \tilde{\mu}_J, \tilde{\sigma}_J^2, \tilde{\lambda})$ or $\tilde{\boldsymbol{\xi}} = (\tilde{\beta}, \tilde{\sigma}_e^2, \tilde{\mu}_B, \tilde{\sigma}_B^2, \tilde{P})$ and should be close to zero for large value of N. We apply the same computation procedure for the adjusted model. The five moment equations for the Poisson-jump process are:

(28)
$$\bar{m}_1 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial l(\widetilde{\boldsymbol{\theta}}, \Delta series_t)}{\partial\hat{\gamma}}\right) = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\sum_{q=0}^{Q}\left(\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot(\Delta series_t-\hat{\gamma}-q\cdot\hat{\mu}_J)\cdot e^{\frac{-(\Delta series_t-\tilde{\gamma}-q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2+q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(2\cdot\hat{\sigma}_\varepsilon^2+2\cdot q\cdot\hat{\sigma}_J^2)\cdot(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2+q\cdot\hat{\sigma}_J^2)}}\right)}{\sum_{q=0}^{Q}\left(\frac{1}{2}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot e^{\frac{-(\Delta series_t-\tilde{\gamma}-q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2+q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2+q\cdot\hat{\sigma}_J^2)}}\right)}\right),$$

(29)
$$\bar{m}_2 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial l(\widetilde{\boldsymbol{\theta}}, \Delta series_t)}{\partial\hat{\mu}_J}\right) = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\sum_{q=0}^{Q}\left(\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot(\Delta series_t-\hat{\gamma}-q\cdot\hat{\mu}_J)\cdot q\cdot e^{\frac{-(\Delta series_t-\tilde{\gamma}-q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2+q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(2\cdot\hat{\sigma}_\varepsilon^2+2\cdot q\cdot\hat{\sigma}_J^2)\cdot(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2+q\cdot\hat{\sigma}_J^2)}}\right)}{\sum_{q=0}^{Q}\left(\frac{1}{2}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot e^{\frac{-(\Delta series_t-\tilde{\gamma}-q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2+q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2+q\cdot\hat{\sigma}_J^2)}}\right)}\right),$$

(30)
$$\bar{m}_3 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial l(\widetilde{\boldsymbol{\theta}}, \Delta series_t)}{\partial\hat{\sigma}_\varepsilon^2}\right)$$
$$= \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\sum_{q=0}^{Q}\left(\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot(\Delta series_t-\hat{\gamma}-q\cdot\hat{\mu}_J)^2\cdot e^{\frac{-(\Delta series_t-\tilde{\gamma}-q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2+q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(2\cdot\hat{\sigma}_\varepsilon^2+2\cdot q\cdot\hat{\sigma}_J^2)^2\cdot(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2+q\cdot\hat{\sigma}_J^2)}} - \frac{1}{4}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot e^{\frac{-(\Delta series_t-\tilde{\gamma}-q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2+q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}\cdot\pi}{(q!)(\pi(2\cdot\hat{\sigma}_\varepsilon^2+q\cdot\hat{\sigma}_J^2))^{3/2}}\right)}{\sum_{q=0}^{Q}\left(\frac{1}{2}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot e^{\frac{-(\Delta series_t-\tilde{\gamma}-q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2+q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2+q\cdot\hat{\sigma}_J^2)}}\right)}\right),$$

(31) $$\bar{m}_4 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial l(\widetilde{\boldsymbol{\theta}}, \Delta series_t)}{\partial \hat{\sigma}_J^2}\right)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\sum_{q=0}^{Q}\left(\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot(\Delta series_t - \hat{\gamma} - q\cdot\hat{\mu}_J)^2\cdot q\cdot e^{\frac{-(\Delta series_t - \tilde{\gamma} - q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(2\cdot\hat{\sigma}_\varepsilon^2 + 2\cdot q\cdot\hat{\sigma}_J^2)^2\cdot(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2 + q\cdot\hat{\sigma}_J^2)}} - \frac{1}{4}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot e^{\frac{-(\Delta series_t - \tilde{\gamma} - q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}\cdot\pi\cdot q}{(q!)(\pi(2\cdot\hat{\sigma}_\varepsilon^2 + q\cdot\hat{\sigma}_J^2))^{3/2}}\right)}{\sum_{q=0}^{Q}\left(\frac{1}{2}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot e^{\frac{-(\Delta series_t - \tilde{\gamma} - q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2 + q\cdot\hat{\sigma}_J^2)}}\right)}\right),$$

(32) $$\bar{m}_5 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial l(\widetilde{\boldsymbol{\theta}}, \Delta series_t)}{\partial \hat{\lambda}}\right)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\sum_{q=0}^{Q}\left(-\frac{1}{2}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot e^{\frac{-(\Delta series_t - \tilde{\gamma} - q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2 + 2\cdot q\cdot\hat{\sigma}_J^2)}} + \frac{1}{2}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot q\cdot e^{\frac{-(\Delta series_t - \tilde{\gamma} - q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{\lambda\cdot(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2 + q\cdot\hat{\sigma}_J^2)}}\right)}{\sum_{q=0}^{Q}\left(\frac{1}{2}\cdot\frac{e^{-\hat{\lambda}}\cdot\hat{\lambda}^q\cdot e^{\frac{-(\Delta series_t - \tilde{\gamma} - q\tilde{\mu}_J)^2}{2(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2)}}\cdot\sqrt{2}}{(q!)\sqrt{\pi(2\cdot\hat{\sigma}_\varepsilon^2 + q\cdot\hat{\sigma}_J^2)}}\right)}\right)$$

The same computations for moment equations are applied for the Bernoulli permanent-jump and temporary-diffusion model and are provided in an Appendix (A1).

The GMM estimation based on parameters and moment conditions of the given likelihood function ignores the existence of autocorrelation at lag one that is characteristic of a first-order moving average process. The presence of autocorrelation could cause underestimation of the standard errors of the parameter estimates. In order to consider autocorrelation, we provide additional information about autocorrelation for a GMM procedure. The additional moment equation reflects the autocorrelation caused by overdifferencing the transitory shocks. This moment equation for autocorrelation is derived by equating an empirical autocorrelation and the theoretical autocorrelation and can be expressed as:

(33) $$m_6 = \text{Empirical Autocorrelation} - \text{Theoretical Autocorrelation}$$

$$= \frac{Cov(\Delta series_t, \Delta series_{t-1})}{\sqrt{var(\Delta series_t) var(\Delta series_{t-1})}} - \frac{-2 \cdot \hat{\sigma}_\varepsilon^2}{\hat{\lambda} \cdot \hat{\sigma}_J^2 - \hat{\lambda}^2 \cdot \hat{\mu}_J^2 + 2 \cdot (2 \cdot \hat{\sigma}_\varepsilon^2)}$$

where the empirical autocorrelation is computed form the actual data set and theoretical autocorrelation is computed from the data generating process assumed to be true. Based on the data generating processes, we compute theoretical autocorrelation between $\Delta series_t$ and $\Delta series_{t-1}$. From Poisson-jump data generating process of equation (4), theoretical autocorrelation is computed as:

(34) $$corr(\Delta series_t, \Delta series_{t-1})$$

$$= \frac{cov(\Delta series_t, \Delta series_{t-1})}{\sqrt{var(\Delta series_t)} \cdot \sqrt{var(\Delta series_{t-1})}}$$

$$= \frac{E[(\Delta series_t - \overline{\Delta series})(\Delta series_{t-1} - \overline{\Delta series})]}{var(\Delta series_t)}$$

$$= \frac{E[(\varepsilon_t - \varepsilon_{t-1})(\varepsilon_{t-1} - \varepsilon_{t-2})]}{var(\Delta series_t)}$$

$$= \frac{E(-\varepsilon_{t-1} \cdot \varepsilon_{t-1})}{var(\gamma + \sum_{q=1}^{Q_t} Jump_{q,t} + \varepsilon_t - \varepsilon_{t-1})}$$

$$= \frac{E(-\varepsilon_{t-1} \cdot \varepsilon_{t-1})}{var(\sum_{q=1}^{Q_t} Jump_{q,t}) + var(\varepsilon_t) + var(\varepsilon_{t-1})}$$

$$= \frac{-\sigma_\varepsilon^2}{\lambda \cdot \sigma_J^2 - (\lambda^2 \cdot \mu_J^2) + 2 \cdot \sigma_\varepsilon^2},$$

where for $var(\sum_{q=1}^{Q_t} Jump_{qt})$

$$var\left(\sum_{q=1}^{Q_t} Jump_{q,t}\right) = E\left(\sum_{q=1}^{Q_t} {Jump_{q,t}}^2\right) - E\left(\sum_{q=1}^{Q_t} Jump_{q,t}\right)^2$$

$$= E\left({Jump_{0,t}}^2 + {Jump_{1,t}}^2 + \cdots + {Jump_{Q_t,t}}^2\right)$$

$$- E\left(Jump_{0,t} + Jump_{1,t} + \cdots + Jump_{Q_t,t}\right)^2$$

$$= \lambda \cdot \sigma_J^2 - \left(\lambda \cdot \mu_J\right)^2.$$

The same computation for theoretical autocorrelation is applied for the Bernoulli-jump model and is in the Appendix (A2).

For positive autocorrelation, the adjusted data generating process (equation (12)) is used to compute a theoritical correlation between $\Delta series_t$ and $\Delta series_{t-1}$. The procedure of computation for the adjusted Poisson-jump process is derived as:

(35)
$$corr(\Delta series_t, \Delta series_{t-1}) = \frac{cov(\Delta series_t, \Delta series_{t-1})}{\sqrt{var(\Delta series_t)} \cdot \sqrt{var(\Delta series_{t-1})}}$$

$$= \frac{E[(\Delta series_t - \overline{\Delta series})(\Delta series_{t-1} - \overline{\Delta series})]}{var(\Delta series_t)}$$

$$= \frac{E\left(\sum_{q=0}^{Q_t} Jump_{q,t} + \rho \cdot \sum_{q=0}^{Q_{t-1}} Jump_{q,t-1} + \varepsilon_t - \varepsilon_{t-1}\right)}{\sqrt{var(\Delta series_t)}}$$

$$\cdot \frac{\left(\sum_{q=0}^{Q_{t-1}} Jump_{q,t-1} + \rho \cdot \sum_{q=0}^{Q_{t-2}} Jump_{q,t-2} + \varepsilon_{t-1} - \varepsilon_{t-2}\right)}{\sqrt{var(\Delta series_{t-1})}}$$

$$= \frac{E\left[\rho \cdot \left(\sum_{q=0}^{Q_{t-1}} Jump_{q,t-1}\right)^2 - (\varepsilon_{t-1})^2\right]}{var\left(\gamma + \sum_{q=0}^{Q_t} Jump_{q,t} + \rho \cdot \sum_{q=0}^{Q_{t-1}} Jump_{q,t-1} + \varepsilon_t - \varepsilon_{t-1}\right)}$$

$$= \frac{\rho \cdot E\left(\sum_{q=0}^{Q_{t-1}} Jump_{q,t-1}\right)^2 - E(\varepsilon_{t-1})^2}{var\left(\sum_{q=1}^{Q_t} Jump_{q,t}\right) + var\left(\rho \cdot \sum_{q=0}^{Q_{t-1}} Jump_{q,t-1}\right) + var(\varepsilon_t) + var(\varepsilon_{t-1})}$$

$$= \frac{\rho \cdot var\left(\sum_{q=0}^{Q_{t-1}} Jump_{q,t-1}\right) - var(\varepsilon_{t-1})}{var\left(\sum_{q=1}^{Q_t} Jump_{q,t}\right) + \rho^2 \cdot var\left(\sum_{q=0}^{Q_{t-1}} Jump_{q,t-1}\right) + var(\varepsilon_t) + var(\varepsilon_{t-1})},$$

$$= \frac{\rho \cdot \left(\lambda \cdot \sigma_J^2 - \left(\lambda^2 \cdot \mu_J^2\right)\right) - \sigma_\varepsilon^2}{(1+\rho^2) \cdot \left(\lambda \cdot \sigma_J^2 - \left(\lambda^2 \cdot \mu_J^2\right)\right) + \sigma_\varepsilon^2},$$

where for $var\left(\sum_{q=1}^{Q_t} Jump_{qt}\right)$

$$var\left(\sum_{q=1}^{Q_t} Jump_{q,t}\right) = E\left(\sum_{q=1}^{Q_t} {Jump_{q,t}}^2\right) - E\left(\sum_{q=1}^{Q_t} Jump_{q,t}\right)^2$$

$$= E\left({Jump_{0,t}}^2 + {Jump_{1,t}}^2 + \cdots + {Jump_{Q_t,t}}^2\right)$$

$$- E\left(Jump_{0,t} + Jump_{1,t} + \cdots + Jump_{Q_t,t}\right)^2$$

$$= \lambda \cdot \sigma_J^2 - \left(\lambda \cdot \mu_J\right)^2,$$

for $\rho^2 \cdot var\left(\sum_{q=0}^{Q_{t-1}} Jump_{q,t-1}\right)$,

$$\rho^2 \cdot var\left(\sum_{q=1}^{Q_t} Jump_{q,t}\right) = \rho^2 \cdot \left(E\left(\sum_{q=1}^{Q_t} {Jump_{q,t}}^2\right) - E\left(\sum_{q=1}^{Q_t} Jump_{q,t}\right)^2\right)$$

$$= \rho^2 \cdot \left(\lambda \cdot \sigma_J^2 - \left(\lambda \cdot \mu_J\right)^2\right).$$

The same computation for theoretical autocorrelation of series having positive autocorrelation is applied for the Bernoulli-jump model and is in the Appendix (A3).

With additional moment condition for autocorrelation, there are more moment equations, $k = \{1, \ldots, 6\}$, than parameters, $l = \{1, \ldots, 5\}$, $(k \geq l)$, and we expect to see an improvement of estimation performance. Andersen and Sorensne (1996) argued including more information in the form of additional moment restriction improves estimation performance for a given degree of precision in the estimate of the weighting matrix, but in small samples, this must be balanced against the deterioration in the estimate of the weighting matrix as the number of moments expands.

With the moment equations, GMM estimation is used to estimate the parameters of the proposed model. The GMM estimator is defined by choosing $\widetilde{\boldsymbol{\theta}}$ to minimize q:

(36) $$q = \bar{m}_k{}' W_k(\widetilde{\boldsymbol{\theta}}) \bar{m}_k,$$

where $\widetilde{\boldsymbol{\theta}}$ is a vector of five parameters, $\widetilde{\boldsymbol{\theta}} = (\tilde{\gamma}, \tilde{\sigma}_\varepsilon^2, \tilde{\mu}_J, \tilde{\sigma}_J^2, \tilde{\lambda})$, $\bar{m}_k$ represents the *k*th moment equations, $k = \{1, \ldots, 6\}$, including additional moment equation of autocorrelation, five moment equations are an expectation of a first order conditions of the log-likelihood function and the additional moment equation in order for autocorrelation is a difference between empirical autocorrelation and theoretical, $W_k(\widetilde{\boldsymbol{\theta}})$ is a positive-definite, symmetric weighting matrix that can depend on sample information, and $W_k(\widetilde{\boldsymbol{\theta}}) = \left(\frac{1}{N}\sum_{t=1}^{N} m_k{}' m_k\right)^{-1}$. A weighting matrix $\left(W_k(\widetilde{\boldsymbol{\theta}})\right)$ is essential to obtain an optimal GMM estimator. Newey and West (1987) found an optimal GMM estimator is obtained when W_k is a consistent estimator of $(S_k)^{-1}$. Based on Newey and West (1987), the asymptotic covariance matrix of $\widetilde{\boldsymbol{\theta}}$ or $(\tilde{\boldsymbol{\xi}})$ for a GMM estimation of the proposed model could be expressed as:

(37) $$V_{\widetilde{\boldsymbol{\theta}},k} = \left(H'_{\widetilde{\boldsymbol{\theta}},k} S_k H_{\widetilde{\boldsymbol{\theta}},k}\right)^{-1},$$

where $H_{\widetilde{\boldsymbol{\theta}},k} = \frac{\sum_{t=1}^{N} E\left[m_{t,\widetilde{\boldsymbol{\theta}},k}(\widetilde{\boldsymbol{\theta}})\right]}{N}$, $m_{t,\widetilde{\boldsymbol{\theta}},k}(\widetilde{\boldsymbol{\theta}})$ is the $(l \times k)$ matrix of partial derivative of $m_{t,k}(\widetilde{\boldsymbol{\theta}})$ with respect to $\widetilde{\boldsymbol{\theta}}$ where $m_{t,k}(\widetilde{\boldsymbol{\theta}})$ is a partial derivative of log-likelihood functions, and $S_k = \frac{\sum_{t=1}^{N} E\left[m_{t,k}(\widetilde{\boldsymbol{\theta}})m_{t,k}(\widetilde{\boldsymbol{\theta}})'\right]}{N}$ if $m_{t,k}$ is serially uncorrelated. The efficient GMM estimator is constructed with a weighting matrix. Greene (2011) discussed that the asymptotic covariance matrix is a function of a weighting matrix. He also discussed that different choices of computing a weighting matrix for the efficient GMM estimator produce different estimates, but the estimator is consistent for any weighting matrix.

In the study, GMM estimation is used with first differenced series. The differenced series created autocorrelation. In order to deal with autocorrelation, we adapt the advantage of GMM estimation that can have more moment equations than parameters. We concern the estimates of GMM with first difference in a series can be biased. A Monte Carlo method is used to estimate the variance of the parameters of all the change in observations that are not independent, since we assume that Poisson (or Bernoulli) and normal distributions of error terms are met by the data. In using a Monte Carlo method, we investigate the finite-sample properties of GMM procedures for conducting inference about standard deviation of first differenced series. The simulation method has been used primarily to obtain information on the small sample properties of asymptotically valid estimators and test statistics, or to calibrate the distribution of test statistics (Pesaran and Pesaran, 1993). The empirical standard deviation of a series of Monte Carlo replications of estimators can be used to approximate the standard error of an estimator. First we draw *M*=10,000 independent samples with same size of an actual data set, where based on the parameters estimated from the permanent-jump and temporary-diffusion model. Second, we estimate the parameters, $\hat{\theta} = (\hat{\gamma}, \hat{\sigma}_{\varepsilon}^2, \hat{\mu}_J, \hat{\sigma}_J^2, \hat{\lambda})$, of the developed model for each generated sample. We obtain *M* numbers of parameters,

$\hat{\theta}_m = (\hat{\gamma}_m, \hat{\sigma}^2_{\varepsilon,m}, \hat{\mu}_{J,m}, \hat{\sigma}^2_{J,m}, \hat{\lambda}_m)$ for m=1,…,M. Then we compute standard error, $se(\hat{\theta})$, by $se(\hat{\theta}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}_b - \tilde{\theta})^2}$, where $\tilde{\theta} = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b$. The standard error of the sample mean is the estimate of the standard deviation of samples means for some sample drawn from the population. The standard error of running Monte Carlo simulation is therefore the estimate of the standard deviation of values returned from running many Monte Carlo simulations.

Test Statistics and Bernoulli-Jump Model under the Alternative Hypothesis

In the study, a benefit of the Bernoulli-jump model is that it nests the classic time-series models such as a random walk model with drift and a linear trend model comparing with the Poisson-jump model. In order to determine whether the Bernoulli-jump model encompass a random walk model with drift and a linear trend model, we conduct the hypotheses tests for the nested models, using Monte Carlo methods. Since the small-sample size of the Wald tests exceeds its asymptotic size and increases sharply with the number of hypotheses being jointly tested (Burnside and Eichenbaum, 1996), Wald test statistics based on the asymptotic approximations used in the study becomes unreliable. In addition, we conduct hypotheses test even there exist nuisance parameters are not identified under the null hypotheses. The Monte Carlo method can be used to estimate the distribution of asymptotically pivotal statistics. Monte Carlo estimate of the variance to construct the Wald statistic provide more reliable small sample inference than the usual asymptotic Wald test in the optimally weighted GMM estimators (Bond and Windmeijer, 2003). Hansen (1996) argued showed the conditional *p*-value transformation provided an asymptotic distribution free of nuisance parameters through Monte Carlo methods. In our approach for an asymptotic distribution, Monte Carlo method is appropriate. Under the null hypotheses of a random walk model with drift and a linear trend model, we generate 10,000 sets of time series, each of length T=actual series. For the test of the nested random walk model with drift, the data generating process is:

(38) $$series_{t,i} = seriesi_{t-1,i} + \beta_i + J_{t,i}$$

where is the simulated series at ith sample, $i = 1, \ldots, N$ and the same size as an actual data, $t = 1, \ldots, T$, β_i is a drift at ith sample, and $J_{t,i}$ is an error term. For the test of the nested linear trend model, the data generating process is:

(39) $$series_{i,t} = \beta_i \cdot t_i + J_{i,t}$$

$$\Delta series_{i,t} = \beta_i + J_{i,t} - J_{i,t-1}$$

where β_i is a time trend coefficient at ith set, and t_i is time at ith set.

The hypotheses under the random walk model with drift in equation (38) are:

(40) $$H_0: \sigma_e^2 = 0, P = 1$$

$$H_A: \sigma_e^2 > 0 \text{ and /or } P < 1,$$

where σ_e^2 and P are parameters from the unrestricted Bernoulli-jump model, and the null hypothesis that a probability of one jump (P) from is equal to one and a variance of temporary shocks (σ_e^2) is equal to zero is tested.

The hypotheses under the linear trend model in the equation (39) are:

(41)
$$H_0: P = 0$$
$$H_A: P > 0,$$

where the null hypothesis that a probability of one permanent shock (P) is equal to zero is tested. In the case, when the null hypothesis is $P = 0$, parameters, μ_B and σ_B^2 are not identified and they are nuisance parameters. According to Hansen (1996), if a conditional transformation is analogous to an asymptotic *p*-value and it has an asymptotic uniform distribution under the null hypothesis then the asymptotic null distribution is free of nuisance parameters.

The test statistic for Wald test is derived as:

(42)
$$W_F = \frac{a(\tilde{\theta})' \left[A(\tilde{\theta}) \cdot Avar(\tilde{\theta}) \cdot A(\tilde{\theta})'\right]^{-1} a(\tilde{\theta})}{R} \xrightarrow{d} F(R, T-K),$$
$$Avar(\tilde{\theta}) = W_k(\tilde{\boldsymbol{\theta}}) = \left(\frac{1}{N}\sum_{t=1}^{N} m_k{}' m_k\right)^{-1},$$

where $a(\cdot)$is a vector-valued function, $A(\cdot)$ is the Jacobian of $a(\cdot)$ and is of full row rank h , $A(\cdot)_{h\times g} = \frac{\partial a(\cdot)}{\partial}$, $h \leq g$. Under the null hypotheses, the test statistics is that $\left({W_F}/{R}\right)/R \xrightarrow{d} F(R, T-K)$.

Optimal Length of Moving Average for Actual and Stochastic Time Series

A moving average method is one of forecasting methods. In a simple moving average method, the forecast for next period will be equal to the average of a specified number of the most recent observation, with each observation receiving the same emphasis. In order to determine the effects of permanent shocks and temporary shocks on optimal length of moving the stochastic series are simulated from the developed models (equations (8), (11), (13) and (15)). Moving average methods use the simple average of the previous *N* years:

(43) $$\Delta se\widehat{ri}es_i(N) = \frac{1}{N}\sum_{s=1}^{N} \Delta series_{i-s},$$

where $\Delta series_i$ is a series or simulated series, $\Delta se\widehat{ri}es_i$ is series forecast, *N* is the moving average interval.

In order to evaluate an optimal length of moving average in actual data and the data simulated from the proposed models, root mean squared errors (RMSE) is applied. The root mean squared error is the square of the difference between the values actually observed and values predicted by a model such as equation (26) and is expressed as:

(44) $$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Delta series_i - \Delta se\widehat{ri}es_i(N))^2},$$

where N is number of years, $N = 1,2, \dots ,5$. The lowest value of RMSE is selected as the optimal length of moving average.

Autoregressive Integrated Moving Average with Outliers

We determine whether a basis price series has a unit root using the augmented Dickey-Fuller (ADF) test according to a general procedure for a time-series model. Time series in levels are tested first, and then in first differences if necessary. Based on the ARIMA specification to use how many autoregressive and moving average parameters to include, we estimate ARIMA models for basis in a first difference until we found the properties of white noise, using Akaike information criterion (AIC) to identify the best structure. Now, we identify the types of outliers in a given data series, but at unknown time points. After detecting outliers, we treat these times as known, and estimate the outlier parameters with parameters of the specified ARIMA model. We could use several methods to deal with the existence of outliers (Franses and Haldrup, 1994). One approach is to consider robust estimation of the model by attaching less weight to extreme observations. Another approach is to remove the outliers' effects with dummy variables. We include dummy variables in the auxiliary augmented Dickey-Fuller regression. This paper first detects the presence of a level shift (permanent shocks) and a transient change in first differences. After outliers for level shifts and transient changes, we again detect outlier without restricting any types of outliers. We add those outliers to an ARIMA model as dummy variables and estimate the ARIMA model with outliers. The structure of the model may be slightly different from the ARIMA without outliers. Thus, we conduct the model identification again and make sure the new residuals exhibit white noise.

Comparison of Empirical CDF of a Series and Theoretical CDF of the Specified Distribution

For the indirect inference of a better fit to data, we use an empirical distribution function statistic. The empirical distribution function statistics are based on the comparison of distribution functions of samples (Stephens, 1974). We begin by simulating series based on the developed models in the study and the competing models, and we compare how well those models capture the dynamic characteristics of an actual series. For the comparison of empirical distribution functions of series and theoretical distribution functions of the specified distributions, Kolmogorov-Smirnov test is used. We consider four models for comparison. There are two cases for the permanent-jump and temporary diffusion model; one is that permanent-jump process follows a Poisson-jump process and the other one is that permanent-jump process follows a Bernoulli-jump process. A conventional ARIMA model and an ARIMA model with outliers are selected as the competing models.

Let $\Delta series_t,\ t = 1, \dots, N$ denote the actual time series in first difference and estimate the cumulative density function, $F(\Delta series) = P[\Delta series_t \leq \Delta series]$, using the proportion of a series that are less than $\Delta series$. The empirical cumulative density function is;

(45)
$$\hat{F}(\Delta series) = \frac{1}{N}\{\#\Delta series_t \leq \Delta series\} = \frac{1}{N}\sum_{t=1}^{N} I(\Delta series_t \leq \Delta series).$$

In the study, we test the assumption that the true distribution is the mixed distribution of Poisson (or Bernoulli) and normal. We simulate $\Delta series_{t,r}$ with $r =$

$1, \dots, 100$ and $t = 1, \dots, N$. Then we obtain the corresponding specified cumulative density function, $F(\Delta series)$. Therefore, the maximum vertical distance of Kolmogorov-Smirnov distance between the empirical CDF and theoretical CDF is:

(46) $$D_{ks} = \max_{s}\left(\left|\hat{F}(\Delta series) - F(\Delta series)\right|\right).$$

CHAPTER IV

Data

As a prime example of where both permanent and transitory shocks in time series are expected, harvest basis prices for Oklahoma hard red winter wheat, Illinois #2 corn and #1soybean are selected. Although basis will vary throughout the marketing year, the variation tends to be more predictable and less extreme than changes in the price of cash price since the carrying charge, arbitrage between the futures and cash markets, and transportation costs (Baldwin and Smith, 2011). According to visual inspection from figures 1 to 3, however, certain shocks in recent years of series make behavior of basis series less predictable. Permanent effects of shocks in grain markets change the relationship between cash and futures prices. Figures from 1 to 3 displays Oklahoma hard red wheat, Illinois corn and soybean basis series for harvest. For the validation of the developed models, we also use four series out of fourteen macroeconomic time series used by Nelson and Plosser (1982).

The basis is the difference between the local cash price and the nearby futures price. A harvest basis is considered in the paper. Harvest basis is useful for grain producers' decision. The information from harvest basis helps grain producers predict harvest prices, evaluate forward contract bids for harvest delivery, and decide whether to sell their grain at harvest or store. For Oklahoma hard red winter wheat, harvest basis is calculated as the cash price in June minus the price of the July futures contract in June. For corn, harvest basis is that the cash price in October minus the futures prices of the December contract. For soybean, harvest basis is that the futures prices of the November contract in October are subtracted from the cash price in October. Monthly average prices are used for cash and futures prices for Oklahoma red wheat, Illinois corn and soybean. The harvest basis use in the paper is annual data taking average for the specific months. Using annual data is more appropriate for the estimation of the developed model than monthly and daily data to estimate the effects of permanent shocks in the series. The characteristic of permanent shocks in the paper is that the effects of them remain in a market.

Cash prices of hard red winter wheat for Oklahoma locations were taken from the Oklahoma Department of Agriculture, Food and Forestry's weekly "Oklahoma Market Report" from 1942 through 2012. The five production areas in Oklahoma for cash prices: Frederick, Medford, Weatherford, and Kingfisher and Okarche included since May 2003, are considered and then the prices from these areas are averaged. Daily spot prices for corn and soybean for seven regions of Illinois Agricultural Marketing service and reflect the mid-range of elevator bids for each region on Thursdays of each month from 1975-2012 (Hatchett, Brorsen and Anderson, 2010 and FarmDoc, 2013).

The cash prices from the seven production areas in Illinois: northern, western, north central, south central, Wabash, west and southwest, Little Egypt are averaged. Futures prices reflect daily closing prices at the Kansas City Board of Trade (KCBT) for hard red winter wheat and at the Chicago Board of Trade (CBOT) for corn and soybean. This template is best used for directly typing in your content. However, you can paste text into the document, but use caution as pasting can produce varying results.

Figure 1. Oklahoma Hard Red Winter Wheat for Harvest Basis from 1942 to 2012

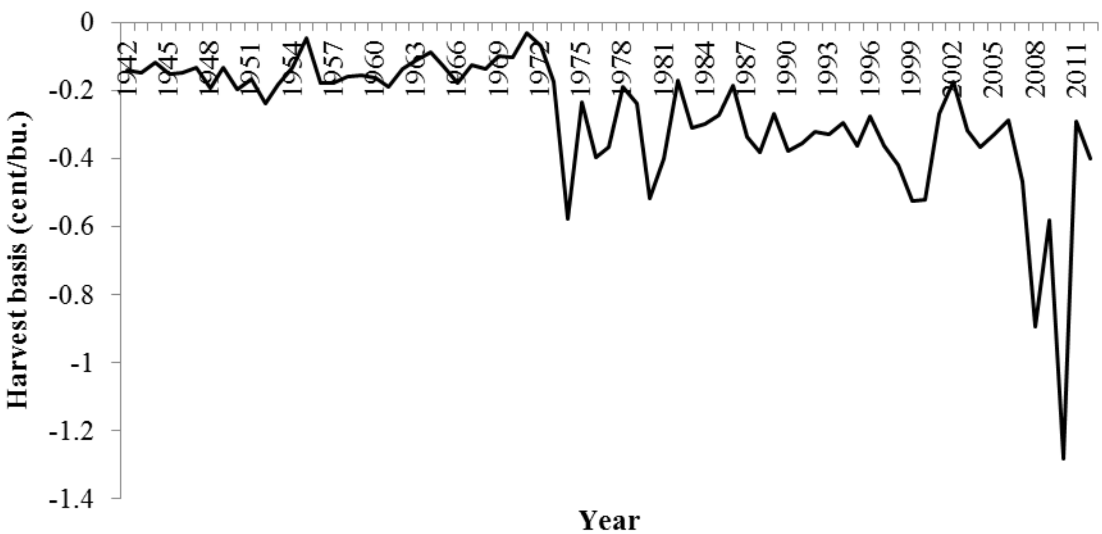


Figure 2. Illinois Corn for Harvest Basis from 1975 to 2012

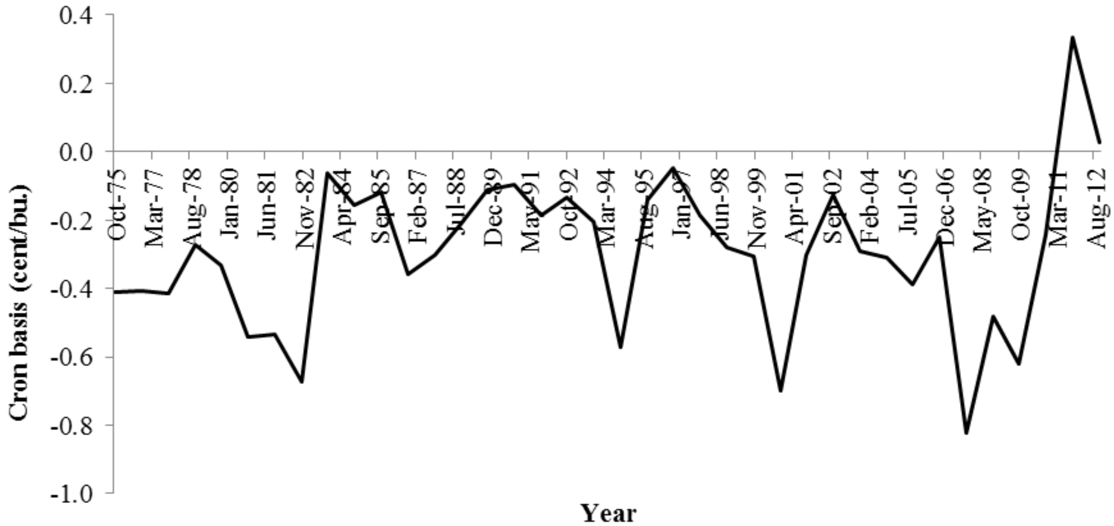
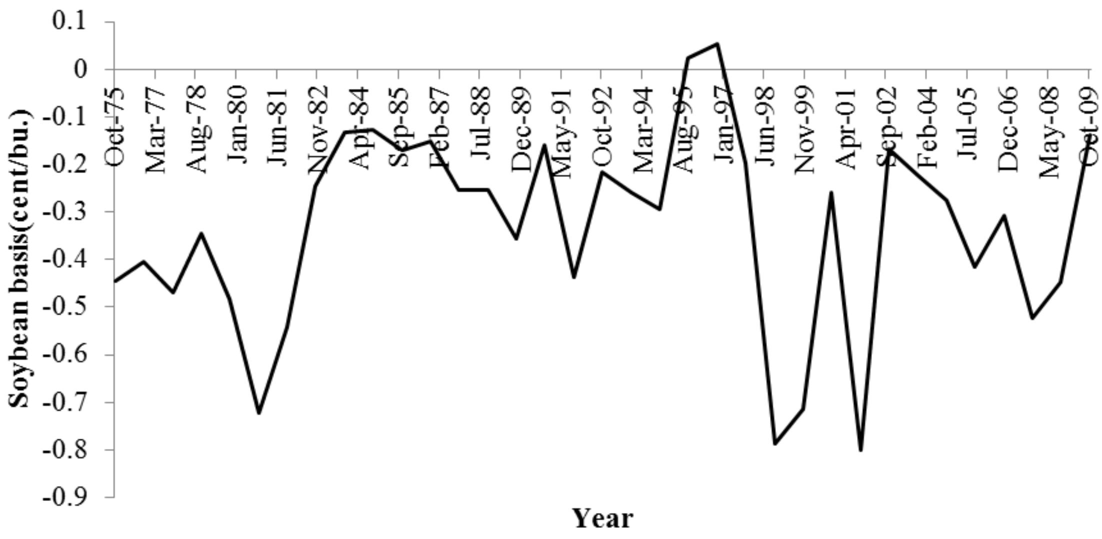


Figure 3. Illinois Soybean for Harvest Basis from 1975 to 2009



In addition to the grain basis series, we use three macroeconomic series out of fourteen series that Nelson and Plosser (1982) used to estimate impacts of structural breaks on stationary processes. Several researchers have applied Nelson and Plosser data sets to their studies to show how different the various estimation methods work with these data series (Perron, 1988, Zivot and Andrews, 1992 Nunes et al., 1997). In the study, we use Nelson and Plosser data sets to show how different the developed model can imply impact of permanent and temporary shocks. All their data can be accessed at http://korora.econ.yale.edu/phillips/data/np&enp.dat. The selected three data series for the new stochastic time series model are the total employment from 1890 to 1988, total unemployment rate from 1890 to 1988, the money stock from 1889 to 1988, and the stock prices from 1871 to 1988 and these three data sets are extended from original data set ending in 1970. The total employment, money stock and stock price follow a random walk while the series of total unemployment rate follows a stationary series. With total employment rate, we show whether the proposed model can estimate stationary series or not. Figures 4-7 display four series. These data are annual data and thus many smaller variations would average out. Since the feature of permanent shocks remaining a market forever, using high-frequency time-series data is not useful to apply the proposed model. We follow their data transformation that the four series are transformed to natural log.

For the original models and the models adjusted for positive autocorrelation, we compute first order autocorrelation in first differenced series. The harvest basis series for Oklahoma red winter wheat, Illinois corn and soybean have negative autocorrelation at lag one. The total employment, money stock and stock price display positive autocorrelation at lag one. Table 1 reports the first order autocorrelation of the series in first difference.

Figure 4. Total Employment from 1890 to 1988

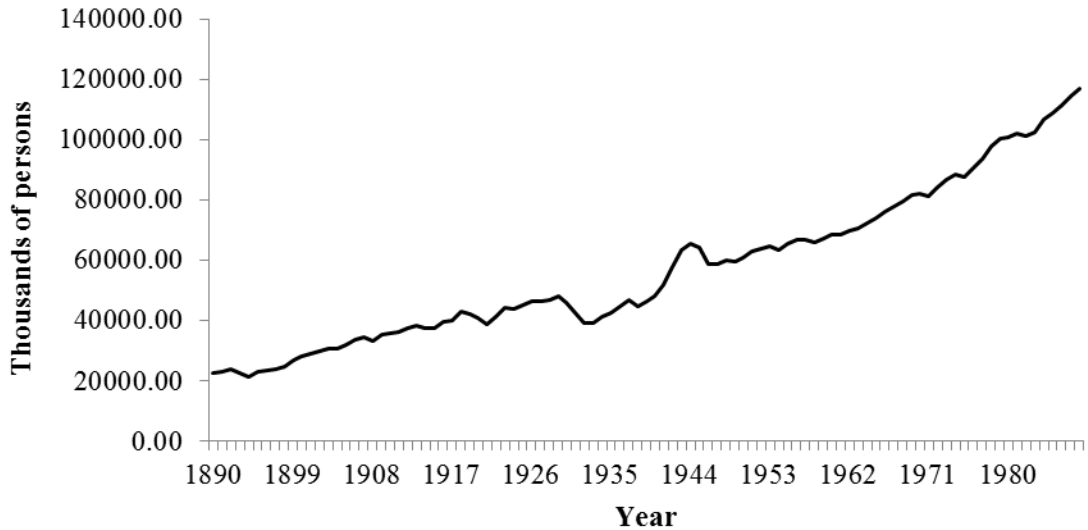


Figure 5. Money Stock from 1889 to 1988

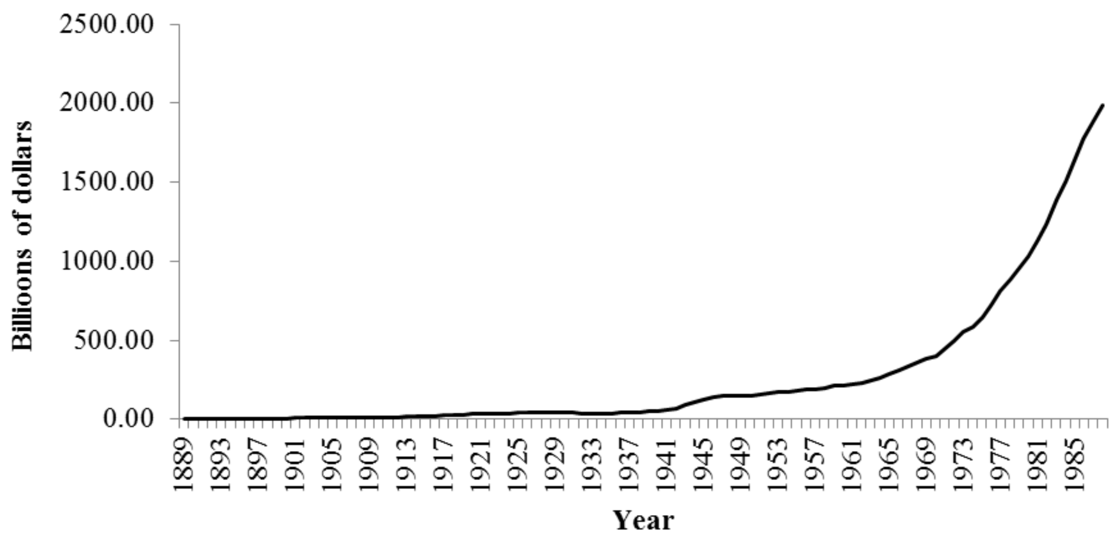


Figure 6. Stock Prices from 1871 to 1988

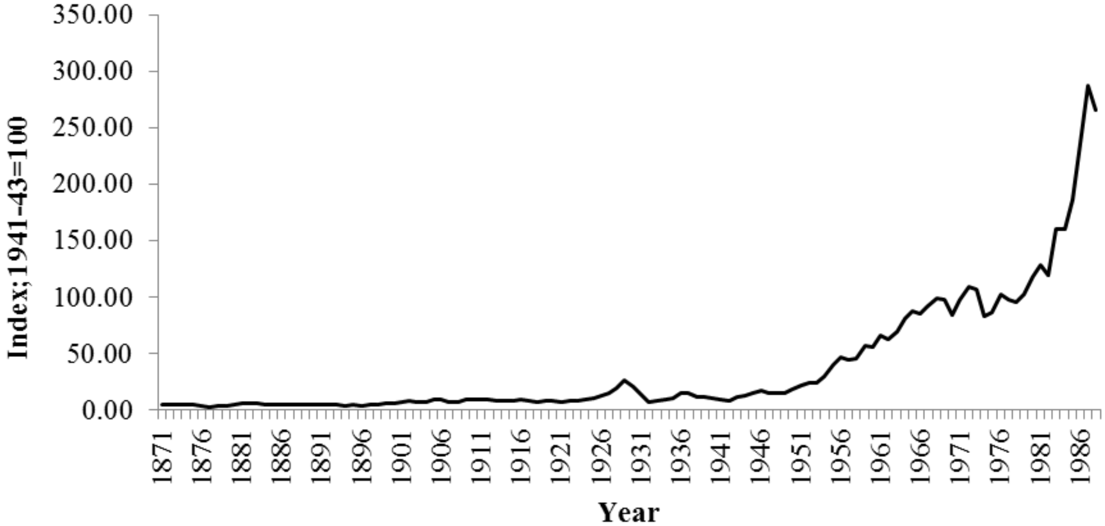


Figure 7. Total Unemployment Rate from 1890 to 1988

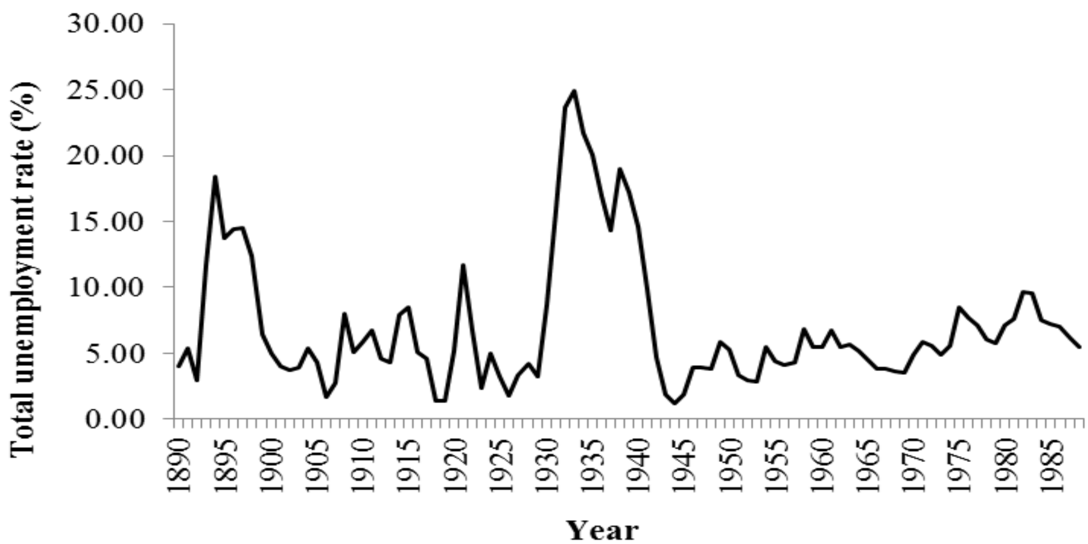


Table 1 Autocorrelation of Series in First Difference

	Autocorrelation	*P*-values for Durbin-Watson test
Oklahoma red wehat harvest basis (1964-2012)	-0.507	<.0001*
Illinois corn harvest basis (1975-2012)	-0.305	0.0566
Illinois soybean harvest basis (1975-2009)	-0.302	0.0674
Money Stock (1889-1988)	0.622	<.0001*
Stock price (1871-1988)	0.174	0.0201*
Total employment (1890-1988)	0.311	0.0005*
Total unemployment rate (1890-1988)	0.755	<.0001*

Note: One (*) indicates the rejection of the hypothesis of no first order autocorrelation at 5% significance level.

CHAPTER V

Results

Poisson-jump and Bernoulli-jump processes are applied to capture permanent shocks. With the Poisson-jump model, multiple jumps are possible to make the Poisson-jump likelihood function tractable, its infinite sum must be approximated by a finite sum. Table 2 shows that once the finite sum reaches four there is no change in the parameter estimates. In the permanent-Bernoulli jump model, the number of permanent shocks is either zero or one. The original model assumes no autocorrelation in jumps and results in negative autocorrelation in the first differenced series. For the requirement of the original model, Oklahoma hard red winter wheat and Illinois corn and soybean basis series are used since they display negative autocorrelation in first differenced series are used. For series having positive autocorrelation, the assumptions of the original model are relaxed. The money stock, stock prices and total employment have positive autocorrelation and are used for the adjusted models. Table 1 reports the autocorrelation in first differenced series. For corn and soybean bases, series show negative autocorrelation but they are failed to reject the null hypothesis of no first order autocorrelation at 5% significance level. We estimate grain basis with adjusted models as well. For macroeconomic variables, all series are positive and significant at 5% level.

Table 2. Parameter Estimates of Poisson-Jump Model with Different Number of Jumps for Oklahoma Wheat Basis

Parameters	1 Jump	2 Jumps	3 Jumps	4 Jumps	5 Jumps	6 >Jumps
Drift (γ)	-0.0082	-0.0081	-0.0080	-0.0080	-0.0080	-0.0080
	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0107)
Jump Mean $\left(\mu_J\right)$	0.0252	0.0193	0.0181	0.0180	0.0180	0.0180
	(0.1245)	(0.0983)	(0.0939)	(0.0934)	(0.0934)	(0.0934)
Variance (σ_ε^2)	0.0027	0.0025	0.0025	0.0025	0.0025	0.0025
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Jump Variance$\left(\sigma_J^2\right)$	0.1767	0.1388	0.1321	0.1315	0.1315	0.1315
	(0.0814)	(0.0693)	(0.0690)	(0.0692)	(0.0692)	(0.0692)
Probability of Jump	0.1765	0.2261	0.2373	0.2384	0.2384	0.2384
(λ)	(0.0635)	(0.0874)	(0.0967)	(0.0980)	(0.0981)	(0.0981)
AIC	-72.1	-74.7	-74.9	-74.9	-74.9	-74.9

Note: numbers in parenthesis are standard errors.

Permanent-Jumps and Temporary-Shocks Model

In table 3-12, the jump mean, the jump variance, and the probability (average of jump probability for Poisson-jump process) are from Poisson-jump and Bernoulli-jump processes, respectively, the variance is from diffusion process, and a drift is considered. These parameters are estimated with different estimation methods, the maximum likelihood estimation (MLE) ignoring the presence of autocorrelation, the generalized method of moments (GMM) with and without imposing an additional moment for autocorrelation, and GMM with Monte Carlo method to compute standard errors for the parameter estimates. From tables 3-12, we compare the values between the third and the fifth columns. The second column shows the estimated parameters with quasi-maximum likelihood estimation but in MLE estimation, we do not consider autocorrelation. For the second and fourth columns, GMM estimation is used. The third column represents the estimated parameters of GMM without considering autocorrelation and thus the coefficients of parameters are similar to the coefficients estimated from MLE. Since the small sample property of GMM, the standard errors can be different. The fourth column reports the estimated parameters of GMM with considering autocorrelation. We add an additional moment equation for autocorrelation. For possibly biased standard errors of GMM estimation, Monte Carlo method is used. The fourth and sixth columns present the estimated parameters from Monte Carlo Method. The coefficients of parameters in the fifth column are used to estimate other objects of this study. Tables 3-8 present the estimates of Poisson-jump and Bernoulli-jump models for the three basis series. Since there obviously exists negative autocorrelation in wheat basis series, we should not ignore autocorrelation causing temporary shocks in the series. The GMM process with autocorrelation for wheat series produces slightly different values in a jump mean and jump variance from GMM without autocorrelation. In the study, we take the values from the GMM estimation with autocorrelation.

Table 3 reports the parameter estimates for Oklahoma hard red winter wheat basis from the permanent-jump and temporary-diffusion model both for Poisson-jump and Bernoulli-jump cases. Most time-series models that consider structural breaks treat structural breaks as indicator variables and estimate the impact of them on a series. However, we impose a Poisson distribution and a Bernoulli distribution to show a probability of occurrence of permanent shocks related to structural breaks and size of permanent shocks. In addition, we consider a distribution for the size of temporary shocks. In the Poisson-jump model from the fifth column in table 6, the jump probability is 0.2281 and in the Bernoulli-jump model, the probability of one jump is 0.1949. The estimated jump variance is relatively bigger than the estimated variance of temporary shocks for both the Poisson-jump and Bernoulli-jump processes. From the results, most shocks are permanent and it implies a shorter moving average is best for forecasting.

Table 4, 6, and 8 present the estimated parameters from the adjusted models with wheat, corn and soybean bases. These grain bases have negative autocorrelation at lag one, but only wheat basis has the statistical significance of autocorrelation. Therefore, we apply the adjusted models for these series. Due to the jump autocorrelation parameter, for Oklahoma red winter wheat, the coefficient values of jump mean and jump variance slightly are changed but the jump probability is decreased in table 4. For corn and soybean, the jump probability is obviously changed while the coefficients values of other parameters are almost close to those of original models. The existence of jump autocorrelation decreases the frequency of jumps.

Table 3. Parameter Estimates for Oklahoma Wheat Harvest Basis from Permanent-Poisson Jumps and Temporary-Shocks Model

Parameters	Without Autocorrelation			With Autocorrelation	
	MLE	GMM	Monte Carlo (1000 samples and size 71)	GMM	Monte Carlo (1000 samples and size 71)
Poisson-jump					
Drift (γ)	-0.0081	-0.0080	-0.0082	-0.0085	-0.0073
	(0.0107)	(0.0076)	(0.0128)	(0.0157)	(0.0290)
Jump mean (μ_J)	0.0193	0.0180	0.0169	0.0096	0.0094
	(0.0934)	(0.1090)	(0.1048)	(0.0792)	(0.0068)
Variance (σ_ε^2)	0.0025	0.0025	0.0026	0.0026	0.0022
	(0.0007)	(0.0006)	(0.0012)	(0.0006)	(0.0011)
Jump variance (σ_J^2)	0.1388	0.1315	0.1288	0.1088	0.1255
	(0.0692)	(0.1156)	(0.0905)	(0.0120)	(0.3767)
Probability of Jump (λ)	0.2261	0.2384	0.2595	0.2281	0.2448
	(0.0981)	(0.1678)	(0.1598)	(0.0810)	(0.3708)
Bernoulli-jump					
Drift (β)	-0.0080	-0.0080	-0.0084	-0.0085	-0.0075
	(0.0107)	(0.0108)	(0.0092)	(0.0108)	(0.0488)
Jump mean (μ_B)	0.0209	0.0209	0.0307	0.0091	0.0067
	(0.1083)	(0.1104)	(0.1274)	(0.1053)	(0.0867)
Variance (σ_e^2)	0.0025	0.0025	0.0025	0.0026	0.0022
	(0.0006)	(0.0007)	(0.0011)	(0.0007)	(0.0082)
Jump variance (σ_B^2)	0.1536	0.1536	0.1463	0.1304	0.1417
	(0.0693)	(0.0816)	(0.1019)	(0.0358)	(0.1677)
Probability of Jump (P)	0.2059	0.2059	0.2157	0.1949	0.2041
	(0.0714)	(0.0743)	(0.1014)	(0.0572)	(0.1317)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Table 4. Parameter Estimates for Oklahoma Wheat Harvest Basis from Adjusted Permanent-Poisson Jumps and Temporary-Shocks Model

	Without Autocorrelation			With Autocorrelation	
Parameters	MLE	GMM	Monte Carlo (1000 samples and size 71)	GMM	Monte Carlo (1000 samples and size 71)
Poisson-jump					
Drift (γ^*)	-0.0072 (0.0111)	-0.0072 (0.0022)	-0.0072 (0.0138)	-0.0072 (0.0082)	-0.0087 (0.0321)
Jump mean (μ_J^*)	0.0208 (0.1366)	0.0208 (0.0004)	0.0478 (0.0611)	0.0083 (0.1643)	0.0207 (0.0338)
Variance (σ_ε^{2*})	0.0024 (0.0007)	0.0024 (0.00003)	0.0026 (0.0012)	0.0023 (0.0004)	0.0096 (0.0218)
Jump variance (σ_J^{2*})	0.1929 (0.1187)	0.1929 (0.0012)	0.1748 (0.0792)	0.1783 (0.1361)	0.1856 (0.0282)
Probability of Jump (λ^*)	0.1416 (0.0663)	0.1416 (0.0001)	0.1516 (0.0893)	0.1478 (0.0600)	0.1349 (0.0325)
Autocorrelation of jump (ρ)	0.1961 (0.2315)	0.1961 (0.0013)	0.2006 (0.0492)	0.2262 (0.0479)	0.1905 (0.0263)
Bernoulli-jump					
Drift (β^*)	-0.0072 (0.0111)	-0.0072 (0.0116)	-0.0077 (0.0153)	-0.0063 (0.0315)	-0.0076 (0.0164)
Jump mean (μ_B^*)	0.0223 (0.1468)	0.0223 (0.1504)	0.0274 (0.0091)	0.0223 (0.0721)	0.0154 (0.0257)
Variance (σ_e^{2*})	0.0024 (0.0007)	0.0024 (0.0007)	0.0028 (0.0041)	0.0012 (0.0004)	0.0092 (0.0172)
Jump variance (σ_B^{2*})	0.2065 (0.1202)	0.2065 (0.1175)	0.1353 (0.0097)	0.0447 (0.0082)	0.1972 (0.0186)
Probability of Jump (P^*)	0.1321 (0.0565)	0.1321 (0.0487)	0.1331 (0.0119)	0.1363 (0.0161)	0.1245 (0.0198)
Autocorrelation of jump (θ)	0.1902 (0.2141)	0.1902 (0.1234)	0.1919 (0.0110)	0.1919 (0.2073)	0.1854 (0.0150)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Table 5. Parameter Estimates for Illinois Corn Harvest Basis from Permanent-Poisson Jumps and Temporary-Shocks Model

	Without Autocorrelation			With Autocorrelation	
Parameters	MLE	GMM	Monte Carlo (1000 samples and size 37)	GMM	Monte Carlo (1000 samples and size 37)
Poisson-jump					
Drift (γ)	-0.0299	-0.0299	-0.0308	-0.0334	-0.0321
	(0.0353)	(0.0871)	(0.0336)	(0.0164)	(0.0608)
Jump mean (μ_J)	0.0500	0.0500	0.0498	0.0582	0.0610
	(0.0652)	(0.1035)	(0.0927)	(0.0345)	(0.0593)
Variance (σ_ε^2)	0.0045	0.0045	0.0044	0.0050	0.0046
	(0.0028)	(0.0093)	(0.0036)	(0.0016)	(0.0080)
Jump variance (σ_J^2)	0.0681	0.0681	0.0582	0.0661	0.0959
	(0.0400)	(0.0930)	(0.0298)	(0.0007)	(0.4636)
Probability of Jump (λ)	0.8336	0.8336	0.8891	0.7747	0.7910
	(0.4464)	(0.4877)	(0.2673)	(0.1363)	(0.4293)
Bernoulli-jump					
Drift (β)	-0.0304	-0.0304	-0.0348	-0.0392	-0.0341
	(0.0361)	(0.0363)	(0.0397)	(0.0441)	(0.0288)
Jump mean (μ_B)	0.0734	0.0734	0.0875	0.0927	0.0927
	(0.0893)	(0.0896)	(0.1223)	(0.1120)	(0.0235)
Variance (σ_e^2)	0.0044	0.0044	0.0043	0.0057	0.0059
	(0.0027)	(0.0019)	(0.0041)	(0.0041)	(0.0364)
Jump variance (σ_B^2)	0.0926	0.0926	0.0872	0.0854	0.0901
	(0.0377)	(0.0235)	(0.0628)	(0.0147)	(0.1785)
Probability of Jump (P)	0.5761	0.5761	0.5836	0.5247	0.5271
	(0.1323)	(0.1555)	(0.2604)	(0.0991)	(0.1442)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Table 6. Parameter Estimates for Illinois Corn Harvest Basis from Adjusted Permanent-Poisson Jumps and Temporary-Shocks Model

	Without Autocorrelation			With Autocorrelation	
Parameters	MLE	GMM	Monte Carlo (1000 samples and size 100)	GMM	Monte Carlo (1000 samples and size 100)
Poisson-jump					
Drift (γ^*)	-0.0299	-0.0291	-0.0280	-0.0581	-0.0308
	(0.0355)	(0.5818)	(0.0242)	(0.0073)	(0.0564)
Jump mean $\left(\mu_J^*\right)$	0.0495	0.0478	0.0476	0.0453	0.0456
	(0.0731)	(1.0540)	(0.0242)	(0.0055)	(0.0450)
Variance (σ_ε^{2*})	0.0045	0.0045	0.0033	0.0136	0.0133
	(0.0028)	(0.2505)	(0.0058)	(0.0005)	(0.0217)
Jump variance $\left(\sigma_J^{2*}\right)$	0.0675	0.0674	0.0628	0.0317	0.0619
	(0.0613)	(1.2557)	(0.0241)	(0.0014)	(0.0473)
Probability of Jump (λ^*)	0.4212	0.4221	0.4282	0.4303	0.4179
	(0.2331)	(1.7120)	(0.0491)	(0.0124)	(0.0506)
Autocorrelation of jump (ρ)	1.0000	1.0005	0.9890	1.0212	0.9936
	(1.3633)	(0.2999)	(0.0580)	(0.0253)	(0.0456)
Bernoulli-jump					
Drift (β^*)	-0.0301	-0.0301	-0.0309	-0.0597	-0.0348
	(0.0357)	(0.0356)	(0.0259)	(0.1408)	(0.0607)
Jump mean (μ_B^*)	0.0605	0.0605	0.0602	0.0555	0.0500
	(0.0853)	(0.0736)	(0.0297)	(0.1828)	(0.0856)
Variance (σ_e^{2*})	0.0044	0.0044	0.0037	0.0143	0.0331
	(0.0027)	(0.0019)	(0.0047)	(0.0066)	(0.0705)
Jump variance (σ_B^{2*})	0.0794	0.0794	0.0708	0.0373	0.0691
	(0.0639)	(0.0234)	(0.0249)	(0.0663)	(0.0694)
Probability of Jump (P^*)	0.3466	0.3466	0.3575	0.3553	0.3405
	(0.1456)	(0.1212)	(0.0467)	(0.0897)	(0.0743)
Autocorrelation of jump (θ)	1.0000	1.0003	0.9910	1.0236	0.9766
	(1.2821)	(0.0026)	(0.0340)	(1.6108)	(0.0563)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Table 7. Parameter Estimates for Illinois Soybean Harvest Basis from Permanent-Poisson Jumps and Temporary-Shocks Model

	Without Autocorrelation			With Autocorrelation	
Parameters	MLE	GMM	Monte Carlo (1000 samples and size 34)	GMM	Monte Carlo (1000 samples and size 34)
Poisson-jump					
Drift (γ)	-0.0139 (0.0456)	-0.0139 (0.0266)	-0.0148 (0.0487)	-0.0072 (0.0172)	-0.0089 (0.0399)
Jump mean (μ_J)	0.0176 (0.0443)	0.0176 (0.0336)	0.0192 (0.0579)	0.0139 (0.0237)	0.0160 (0.0463)
Variance (σ_ε^2)	0.0026 (0.0039)	0.0026 (0.0011)	0.0029 (0.0043)	0.0034 (0.0020)	0.0026 (0.0032)
Jump variance (σ_J^2)	0.0425 (0.0259)	0.0425 (0.0066)	0.0403 (0.0250)	0.0430 (0.0004)	0.0780 (0.4929)
Probability of Jump (λ)	1.3039 (0.7461)	1.3039 (0.2142)	1.3342 (0.3582)	1.2139 (0.1687)	1.2329 (0.4149)
Bernoulli-jump					
Drift (β)	-0.0170 (0.0465)	-0.0170 (0.0593)	-0.0259 (0.3008)	-0.0103 (0.0471)	-0.0155 (0.1718)
Jump mean (μ_B)	0.0361 (0.0805)	0.0361 (0.0911)	0.0367 (0.0987)	0.0287 (0.0847)	0.3281 (0.0955)
Variance (σ_e^2)	0.0025 (0.0037)	0.0025 (0.0048)	0.0034 (0.0045)	0.0032 (0.0052)	0.0030 (0.0050)
Jump variance (σ_B^2)	0.0750 (0.0275)	0.0750 (0.0274)	0.0690 (0.0295)	0.0742 (0.0258)	0.0764 (0.0269)
Probability of jump (P)	0.7205 (0.2215)	0.7205 (0.2591)	0.7551 (0.2636)	0.6857 (0.2085)	0.7073 (0.1292)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Table 8. Parameter Estimates for Illinois Soybean Harvest Basis from Adjusted Permanent-Poisson Jumps and Temporary-Shocks Model

	Without Autocorrelation			With Autocorrelation	
Parameters	MLE	GMM	Monte Carlo (1000 samples and size 34)	GMM	Monte Carlo (1000 samples and size 34)
Poisson-jump					
Drift (γ^*)	-0.0209	-0.0209	-0.0297	-0.0313	-0.0327
	(0.0672)	(0.0065)	(0.0289)	(0.1044)	(0.0084)
Jump mean (μ_J^*)	0.0249	0.0249	0.0209	0.0520	0.0508
	(0.0581)	(0.0073)	(0.0259)	(0.1100)	(0.0070)
Variance (σ_ε^{2*})	0.0016	0.0016	0.0015	0.0007	0.0020
	(0.0044)	(0.0029)	(0.0013)	(0.0047)	(0.0088)
Jump variance (σ_J^{2*})	0.0474	0.0474	0.0405	0.0623	0.0603
	(0.0522)	(0.0244)	(0.0165)	(0.0283)	(0.0144)
Probability of Jump (λ^*)	0.8486	0.8486	0.8566	0.6449	0.6447
	(0.7527)	(1.3111)	(0.0277)	(0.4457)	(0.0007)
Autocorrelation of jump (ρ)	0.4179	0.4179	0.4523	0.3312	0.3306
	(1.4263)	(0.5420)	(0.0631)	(0.2248)	(0.0033)
Bernoulli-jump					
Drift (β^*)	-0.0315	-0.0315	-0.0302	-0.0313	-0.0427
	(0.0604)	(0.1323)	(0.0125)	(0.1044)	(0.0639)
Jump mean (μ_B^*)	0.0529	0.0529	0.0527	0.0520	0.0395
	(0.0849)	(0.1439)	(0.0154)	(0.1100)	(0.0457)
Variance (σ_e^{2*})	0.0009	0.0009	0.0011	0.0007	0.0058
	(0.0038)	(0.0085)	(0.0005)	(0.0047)	(0.0080)
Jump variance (σ_B^{2*})	0.0745	0.0744	0.0696	0.0623	0.0498
	(0.0354)	(0.0489)	(0.0109)	(0.0283)	(0.0301)
Probability of Jump (P^*)	0.6389	0.6489	0.6348	0.6449	0.6314
	(0.3181)	(0.5687)	(0.0412)	(0.4457)	(0.0235)
Autocorrelation of jump (θ)	0.2002	0.2002	0.1972	0.3312	0.3192
	(0.3345)	(0.1879)	(0.0148)	(0.2248)	(0.0199)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Tables 9-11 report the results of the adjusted model with money stock, stock prices and total employment series that have positive autocorrelation. Money stock, stock prices and total employment series are selected out of the fourteen series. Money stock and stock price series are have relatively bigger jump variance than variance from temporary shocks, and total employment series has relatively smaller values of jump variance than variance from temporary shocks. In tables 6-8, the parameters of the jump mean, the jump variance, the probability and the autocorrelation of jumps are from Poisson-jump and Bernoulli-jump processes, respectively, the variance is from temporary shocks, and a drift. The positive autocorrelation of permanent shocks neutralize the negative autocorrelation of temporary shocks. In the study of Perron (1982) with Nelson and Plosser's (1982) fourteen microeconomic series, he imposed one structural break either at the Great Crash of 1929 or at the 1973 oil-price shock. Later studies based on Perron's(1982) paper have argued whether the time point occurring structural breaks is known and how many jumps should be imposed. With the proposed models, we do not concentrate on how structural breaks affect unit root tests.

Table 9 reports the parameter estimates for the money stock series from the adjusted permanent-jump and temporary-diffusion models both for Poisson-jump and Bernoulli-jump cases. In table 9, we compare the values between the second and the fourth columns. The second column shows the parameters estimated from GMM without an additional moment condition about autocorrelation, and the fourth column shows the parameters estimated from GMM with an additional moment condition about autocorrelation. For money stock series, the GMM process with autocorrelation produces slightly different estimates and standard errors from GMM without autocorrelation. The value of jump variance is also bigger than that of variance of temporary shocks for both a Poisson-jump process and a Bernoulli-jump process with a money stock series. The positive autocorrelation from the fourth column, 0.1814 for Poisson-jump model and 0.1714 for Bernoulli-jump model, imply that a large value of $\sum_{q=0}^{Q_t} Jump_{q,t}$ is likely to be followed by an additional effect in the next time period.

Table 10 presents the estimates using stock prices. The values of jump variance, 0.0006 for Poisson and 0.0011 for Bernoulli are smaller than the values of variance from temporary shocks, 0.0092 for Poisson and 0.0092 for Bernoulli and the probability is relatively smaller than that of other series. From table 11, the variance of temporary shocks in total employment series is close to zero (0.00003), but the variances of permanent shocks is relatively bigger for both the adjusted Poisson-jump model (0.0014) and an adjusted Bernoulli-jump model (0.0019).

Table 9. Parameter Estimates for Money Stock from Adjusted Permanent-Poisson Jumps and Temporary-Shocks Model

	Without Autocorrelation			With Autocorrelation	
Parameters	MLE	GMM	Monte Carlo (1000 samples and size 100)	GMM	Monte Carlo (1000 samples and size 100)
Poisson-jump					
Drift (γ^*)	0.0699	0.0699	0.0070	0.0699	0.0699
	(0.0060)	(0.0033)	(0.0029)	(0.0186)	(0.0030)
Jump mean (μ_J^*)	-0.0201	-0.0201	-0.0204	-0.0250	-0.0251
	(0.0322)	(0.0113)	(0.0069)	(0.0616)	(0.0037)
Variance (σ_ε^{2*})	0.0006	0.0006	0.0005	0.0006	0.0006
	(0.0004)	(0.0001)	(0.0003)	(0.0003)	(0.0031)
Jump variance (σ_J^{2*})	0.0067	0.0067	0.0059	0.0078	0.0075
	(0.0060)	(0.0014)	(0.0028)	(0.0080)	(0.0034)
Probability (λ^*)	0.2506	0.2503	0.2519	0.2066	0.2065
	(0.3326)	(0.0409)	(0.0094)	(0.1906)	(0.0029)
Autocorrelation of jump (ρ)	0.2112	0.2112	0.2113	0.1814	0.1811
	(0.5971)	(0.0024)	(0.0095)	(0.1478)	(0.0047)
Bernoulli-jump					
Drift (β^*)	0.0699	0.0699	0.0718	0.0700	0.0728
	(0.0060)	(0.0061)	(0.0084)	(0.0056)	(0.0100)
Jump mean (μ_B^*)	-0.0241	-0.0241	-0.0250	-0.0275	-0.0298
	(0.0336)	(0.0432)	(0.0223)	(0.0296)	(0.0114)
Variance (σ_e^{2*})	0.0006	0.0006	0.0007	0.0007	0.0008
	(0.0004)	(0.0005)	(0.0005)	(0.0001)	(0.0007)
Jump variance (σ_B^{2*})	0.0078	0.0078	0.0073	0.0083	0.0077
	(0.0049)	(0.0045)	(0.0031)	(0.0031)	(0.0016)
Probability of jump (P^*)	0.2119	0.2118	0.2244	0.1902	0.2257
	(0.1984)	(0.1914)	(0.0258)	(0.0673)	(0.0383)
Autocorrelation of jump (θ)	0.1939	0.1939	0.1992	0.1714	0.2030
	(0.5077)	(0.2801)	(0.0206)	(0.1361)	(0.0456)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Table 10. Parameter Estimates for Stock Prices from Adjusted Permanent-Poisson Jumps and Temporary-Shocks Model

Parameters	Without Autocorrelation			With Autocorrelation	
	MLE	GMM	Monte Carlo (1000 samples and size 118)	GMM	Monte Carlo (1000 samples and size 118)
Poisson-jump					
Drift (γ^*)	0.0467 (0.0169)	0.0467 (0.0727)	0.0475 (0.0027)	0.0469 (0.0125)	0.0471 (0.0160)
Jump mean $\left(\mu_J^*\right)$	-0.5992 (0.2858)	-0.5992 (0.2853)	-0.5997 (0.0180)	-0.6053 (0.1384)	-0.6059 (0.0171)
Variance (σ_ε^{2*})	0.0092 (0.0022)	0.0092 (0.0491)	0.0072 (0.0029)	0.0092 (0.0021)	0.0069 (0.0187)
Jump variance $\left(\sigma_J^{2*}\right)$	0.0006 (0.0334)	0.0006 (0.0162)	0.0014 (0.0032)	0.0002 (0.0144)	0.0049 (0.0102)
Probability (λ^*)	0.0160 (0.0271)	0.0160 (0.3411)	0.0181 (0.0191)	0.0155 (0.0356)	0.0145 (0.0079)
Autocorrelation of jump (ρ)	0.2766 (2.3941)	0.2767 (0.6191)	0.2803 (0.0189)	0.3198 (1.5796)	0.3127 (0.0408)
Bernoulli-jump					
Drift (β^*)	0.0465 (0.0165)	0.0465 (0.0152)	0.0488 (0.0041)	0.0464 (0.0132)	0.0472 (0.0079)
Jump mean (μ_B^*)	-0.5936 (0.1771)	-0.5936 (0.1902)	-0.6236 (0.1103)	-0.5888 (0.1779)	-0.5766 (0.0466)
Variance (σ_e^{2*})	0.0092 (0.0018)	0.0092 (0.0013)	0.0074 (0.0031)	0.0093 (0.0011)	0.0105 (0.0037)
Jump variance (σ_B^{2*})	0.0011 (0.0324)	0.0011 (0.0190)	0.0047 (0.0156)	0.0015 (0.0200)	0.0051 (0.0096)
Probability of Jump (P^*)	0.0165 (0.0166)	0.0165 (0.0159)	0.0141 (0.0055)	0.0168 (0.0152)	0.0145 (0.0066)
Autocorrelation of jump (θ)	0.2295 (1.4524)	0.2296 (0.9830)	0.1876 (0.2084)	0.2038 (0.3592)	0.1958 (0.1275)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Table 11. Parameter Estimates for Total Employment from Adjusted Permanent-Poisson Jumps and Temporary-Shocks Model

Parameters	Without Autocorrelation			With Autocorrelation	
	MLE	GMM	Monte Carlo (1000 samples and size 99)	GMM	Monte Carlo (1000 samples and size 99)
Poisson-jump					
Drift (γ^*)	0.0243	0.0244	0.0242	0.0264	0.0251
	(0.0026)	(0.0007)	(0.0016)	(0.0009)	(0.0020)
Jump mean (μ_J^*)	-0.0091	-0.0092	-0.0090	-0.0088	-0.0073
	(0.0061)	(0.0013)	(0.0021)	(0.0013)	(0.0062)
Variance (σ_ε^{2*})	0.00003	0.00003	0.00002	0.00003	0.0007
	(0.00003)	(0.00001)	(0.00003)	(5.12E-6)	(0.0001)
Jump variance (σ_J^{2*})	0.0014	0.0014	0.0014	0.0013	0.0012
	(0.0006)	(0.0001)	(0.0003)	(0.0352)	(0.0002)
Probability (λ^*)	0.6863	0.6804	0.6223	0.6837	0.6570
	(0.3487)	(0.0219)	(0.0231)	(0.0242)	(0.0511)
Autocorrelation of jump (ρ)	0.2147	0.2142	0.2148	0.2550	0.2339
	(0.3376)	(0.0178)	(0.0123)	(0.0096)	(0.0155)
Bernoulli-jump					
Drift (β^*)	0.0243	0.0244	0.0242	0.0243	0.0242
	(0.0026)	(0.0027)	(0.0023)	(0.0028)	(0.0111)
Jump mean (μ_B^*)	-0.0130	-0.0132	-0.0136	-0.0129	-0.0093
	(0.0078)	(0.0082)	(0.0029)	(0.0082)	(0.0035)
Variance (σ_e^{2*})	0.00003	0.00004	0.00003	0.00003	0.0008
	(0.00003)	(0.00003)	(0.00002)	(0.00002)	(0.0024)
Jump variance (σ_B^{2*})	0.0019	0.0019	0.0018	0.0019	0.0016
	(0.0005)	(0.0004)	(0.0004)	(0.0003)	(0.0036)
Probability of Jump (P^*)	0.4965	0.4938	0.4861	0.4982	0.5482
	(0.1351)	(0.0934)	(0.0129)	(0.0938)	(0.0157)
Autocorrelation of jump (θ)	0.1699	0.1691	0.1682	0.1824	0.1835
	(0.1870)	(0.1019)	(0.0301)	(0.0317)	(0.0113)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Table 12 reports the results of the adjusted models with total unemployment rate. Total employment, money stock and stock price follow non-stationary process while total unemployment rate follows stationary process. We estimate permanent and temporary shocks in the total unemployment rate series. The total unemployment rate has positive autocorrelation and the adjusted model is applied. From the estimated parameters of GMM estimation with considering autocorrelation, we found that jump variance (0.3527) is relatively bigger than variance (0.0298) from temporary shocks. The frequency of jump probability is about 0.5088. The test of unit roots is a first and primary step of a time-series analysis. Although we do not focus on the existence of unit root in a series in the study, we apply a stationary series to estimate a probability of permanent shocks and a distribution of the size of permanent shocks and a distribution of the size of temporary shocks.

Table 12. Parameter Estimates for Total Unemployment Rate from Adjusted Permanent-Poisson Jumps and Temporary-Shocks Model

	Without Autocorrelation			With Autocorrelation	
Parameters	MLE	GMM	Monte Carlo (1000 samples and size 99)	GMM	Monte Carlo (1000 samples and size 99)
Poisson-jump					
Drift (γ^*)	1.6119 (0.0656)	1.6190 (0.0918)	1.6140 (0.0276)	1.6658 (0.0079)	1.6588 (0.0765)
Jump mean (μ_J^*)	0.1302 (0.2086)	0.1302 (0.0185)	0.1260 (0.0471)	0.0541 (0.0068)	0.0532 (0.0632)
Variance (σ_ε^{2*})	0.0298 (0.0140)	0.0298 (0.0123)	0.0279 (0.0067)	0.0086 (0.0007)	0.0140 (0.0365)
Jump variance (σ_J^{2*})	0.3527 (0.2016)	0.3527 (0.0441)	0.3427 (0.0474)	0.2552 (0.0026)	0.2484 (0.1057)
Probability (λ^*)	0.5088 (0.1980)	0.5089 (0.3230)	0.5251 (0.0606)	0.7539 (0.0183)	0.7524 (0.0289)
Autocorrelation of jump (ρ)	1.0000 (0.8151)	1.0002 (1.5003)	1.0146 (0.0511)	0.9732 (0.0188)	0.9723 (0.0148)
Bernoulli-jump					
Drift (β^*)	1.6137 (0.0668)	1.6190 (0.2426)	1.6076 (0.0391)	1.6188 (0.6127)	1.6196 (0.0187)
Jump mean (μ_B^*)	0.1705 (0.1297)	0.1301 (0.3872)	0.1655 (0.0691)	0.1301 (0.2360)	0.1261 (0.0351)
Variance (σ_e^{2*})	0.0287 (0.0134)	0.0298 (0.0911)	0.0269 (0.0157)	0.0291 (0.0287)	0.0416 (0.0374)
Jump variance (σ_B^{2*})	0.4300 (0.2065)	0.3526 (0.3509)	0.4165 (0.0812)	0.3525 (0.2472)	0.3443 (0.0256)
Probability (P^*)	0.4040 (0.1103)	0.5088 (0.1024)	0.4184 (0.0872)	0.5088 (0.0169)	0.5004 (0.0309)
Autocorrelation of jump (θ)	1.0000 (0.7528)	1.0000 (0.1705)	1.0187 (0.0689)	1.0003 (0.1165)	0.9997 (0.0162)

Note: Numbers in parenthesis are standard errors. In GMM with and without autocorrelation, for a nonlinear minimization problem, Newton-Raphson method with line search and Newton-Raphson method with ridging method are applied. The optimization techniques stop the iteration process when the absolute function convergence criterion meets 1.e-8 or 1.e-12.

Test Statistics and Bernoulli-Jump Model under the Alternative Hypothesis

In the test for a random walk model with a drift, using the Bernoulli-jump process $H_0: \sigma_e^2 = 0, P = 1$ is tested with the simulated series with 5000 replications and the same sample size as the corresponding actual series. The Wald test statistic is asymptotically pivotal, which means that its distribution does not depend on unknown parameters. Table 13 reports the value of Wald test statistics and asymptotic critical values from the Monte Carlo simulation with 10% and 5% significance levels, respectively. For a linear trend model, we test $H_0: P = 0$.

We fail to reject the null hypothesis of $H_0: \sigma_e^2 = 0, P = 1$ for harvest wheat, corn and soybean basis series at both 10% and 5% significance levels, respectively, since the asymptotic critical values at the 10% and 5% significance levels are greater than the calculated Wald test statistics. However, we reject the null hypothesis of $H_0: P = 0$ at the 5 % significance levels. We found that we fail to reject a random walk model with drift but we could not fail to reject a linear trend model.

For money stock, stock price and total employment series, since they have a positive autocorrelation at lag one, we simulated the adjusted Bernoulli-jump model. From table 13, for the test of nested random walk model with drift, we fail to reject the null hypothesis of $H_0: \sigma_e^2 = 0, P = 1$ at the 10% significance level, respectively. For the test of the nested linear trend model, the null hypothesis of $H_0: P = 0$ can be rejected at the 5% significance levels. In macroeconomic variables, we also found that a Bernoulli-jump model fail to reject to a random walk model with drift but a Bernoulli-jump model reject a linear trend model.

Table 13. Monte Carlo Results for Nested Models Test Statistics

Nested Models	Wald Test Statistics	Asymptotic Critical Values	
		5%	10%
	Wheat basis (T=71)		
Random walk model with drift ($H_0: \sigma_e^2 = 0, P = 1$)	2.657	3.332	2.814
Linear trend model ($H_0: P = 0$)	7.191	1.799	1.641
	Corn basis (T=38)		
Random walk model with drift ($H_0: \sigma_e^2 = 0, P = 1$)	2.062	2.648	2.191
Linear trend model ($H_0: P = 0$)	11.154	3.466	1.645
	Soybean basis (T=35)		
Random walk model with drift ($H_0: \sigma_e^2 = 0, P = 1$)	1.794	2.953	1.946
Linear trend model ($H_0: P = 0$)	8.565	2.983	1.501
	Money stock (T=100)		
Random walk model with drift ($H_0: \sigma_e^2 = 0, P = 1$)	1.432	2.869	2.013
Linear trend model ($H_0: P = 0$)	8.542	1.899	1.707
	Stock price (T=118)		
Random walk model with drift ($H_0: \sigma_e^2 = 0, P = 1$)	2.182	2.964	2.302
Linear trend model ($H_0: P = 0$)	10.519	1.379	1.288
	Total Employment (T=99)		
Random walk model with drift ($H_0: \sigma_e^2 = 0, P = 1$)	1.408	3.546	3.159
Linear trend model ($H_0: P = 0$)	5.509	1.892	1.572

Note: *T* is a number of observations.

Optimal Length of Moving Average

Based on the Poisson jump and Bernoulli jump processes with an autocorrelation moment condition, we simulated stochastic series. We select two examples; Oklahoma wheat basis having negative autocorrelation at lag one and stock prices having positive autocorrelation at lag one. We first apply a simple moving average method for actual wheat basis series, stock prices, and simulated stochastic series and then measure the accuracy of forecasts, root mean squared errors (RMSE). The simulated series by the proposed models involves the existence of permanent shocks in the series.

Tables 14 and 16 present length of the moving average, using historical series and stochastic series from Poisson jump and Bernoulli jump models, respectively. In table 14, the 2-year moving average has the lowest RMSE for wheat, corn and soybean harvest basis forecasts from actual series. From table 14, the RMSE with stochastic series gives relatively large errors. Thus, we compute RMSE with adjusted models and the RMSE provide with smaller errors. Table 15 presents the results of RMSE for wheat, corn and soybean harvest basis with adjusted models. For money stock, stock prices and total employment forecasts, the previous year has the lowest RMSE. With the simulated series, the last year is the lowest RMSE for wheat corn and soybean harvest basis forecasts, and the last-year is the lowest RMSE for money stock, stock prices and total employment forecasts. In all the series, the shorter length of moving average is preferred. That is, the effects of permanent shocks dominate and the optimal length of moving average to use when forecasting is small to quickly respond to the permanent shocks. In wheat, corn and soybean bases, the optimal length of moving average (2-years) from historical series differs slightly from the optimal length of moving average (last-year) from the estimated model. In addition, for stock prices of table 16, the simulated series for both Poisson-jump and Bernoulli-jump models report that the optimal length of moving average to use forecasts is 2-years. The reason is that the probability of occurrence of permanent shocks is relatively small (Table 10) and size of permanent shocks is smaller than that of temporary shocks. With these results, the effects of permanent shocks is important for forecasting a series and the effects of temporary shocks is essential for accurate forecasting as well.

Table 14. RMSE of Simple Moving Average Models for Grain Basis

Years	Actual Data	Poisson-Jumps	Bernoulli-Jump
Oklahoma wheat basis	(*T=71*)	(*T=10,000*)	(*T=10,000*)
1-year	0.19174	0.23584	0.24891
2-year	0.16633	0.30265	0.31298
3-year	0.16692	0.35421	0.36441
4-year	0.16772	0.39903	0.40866
5-year	0.16744	0.43972	0.44830
Illinois corn basis	(*T=38*)	(*T=10,000*)	(*T=10,000*)
1-year	0.25214	0.34872	0.32080
2-year	0.24831	0.45016	0.40946
3-year	0.24907	0.52930	0.47855
4-year	0.26457	0.59787	0.53856
5-year	0.26152	0.65966	0.59276
Illinois soybean basis	(*T=35*)	(*T=10,000*)	(*T=10,000*)
1-year	0.24381	0.35319	0.33176
2-year	0.23937	0.43109	0.43250
3-year	0.244165	0.50567	0.51033
4-year	0.24858	0.56960	0.57746
5-year	0.25347	0.67880	0.63715

Note: RMSE is the root mean squared error. The lowest RMSE suggests the optimal length for series forecasts. T is a number of observations.

Table 15. RMSE of Simple Moving Average Models for Grain Basis with Adjusted Model

Years	Actual Data	Adjusted Poisson-Jumps	Adjusted Bernoulli-Jump
Oklahoma wheat basis	(*T=71*)	(*T=10,000*)	(*T=10,000*)
1-year	0.19174	0.23584	0.24891
2-year	0.16633	0.30265	0.31298
3-year	0.16692	0.35421	0.36441
4-year	0.16772	0.39903	0.40866
5-year	0.16744	0.43972	0.44830
Illinois corn basis	(*T=38*)	(*T=10,000*)	(*T=10,000*)
1-year	0.25214	0.34872	0.32080
2-year	0.24831	0.45016	0.40946
3-year	0.24907	0.52930	0.47855
4-year	0.26457	0.59787	0.53856
5-year	0.26152	0.65966	0.59276
Illinois soybean basis	(*T=35*)	(*T=10,000*)	(*T=10,000*)
1-year	0.24381	0.35319	0.33176
2-year	0.23937	0.43109	0.43250
3-year	0.244165	0.50567	0.51033
4-year	0.24858	0.56960	0.57746
5-year	0.25347	0.67880	0.63715

Note: RMSE is the root mean squared error. The lowest RMSE suggests the optimal length for series forecasts. T is a number of observations.

Table 16. RMSE of Simple Moving Average Models for Macroeconomic Variables

Years	Actual Data	Poisson-Jumps	Bernoulli-Jump
Money stock	(*T=100*)	(*T=10,000*)	(*T=10,000*)
1-year	0.08591	0.09153	0.09111
2-year	0.12411	0.11697	0.11622
3-year	0.16125	0.14634	0.14529
4-year	0.19719	0.17699	0.17555
5-year	0.23293	0.20817	0.20643
Stock prices	(*T=118*)	(*T=10,000*)	(*T=10,000*)
1-year	0.16035	0.21391	0.21145
2-year	0.19582	0.20379	0.19959
3-year	0.22326	0.20876	0.20427
4-year	0.24737	0.21786	0.21325
5-year	0.26815	0.22922	0.22499
Total employment	(*T=99*)	(*T=10,000*)	(*T=10,000*)
1-year	0.03890	0.03660	0.04632
2-year	0.05077	0.04626	0.06148
3-year	0.06081	0.05560	0.07716
4-year	0.07030	0.06464	0.09296
5-year	0.07877	0.07355	0.10881

Note: RMSE is the root mean squared error. The lowest RMSE suggests the optimal length for series forecasts. T is a number of observations.

Tables 17 and 18 show the effects of jump size and frequency on optimal length of moving average and a case where there is no time trend, through the simulated series for wheat basis and stock prices. The jump frequency is increased by 0.1 units from 0 through 1, and the size of jumps is changed by an estimates of jump variance. From table 17, in the Poisson jump process for wheat harvest basis, frequency is zero and the size of jump changes, 3-year is the optimal length. However, when the jump frequency and the size of jumps become larger, the one year is the optimal length. From table 18, for simulated stock prices of the Poisson jump process, when the jump frequency is 0 and the size of jumps is 0, the optimal length of moving average is the 2-year. Figure 7-14 visually shows the changes of optimal length of moving average.

If there are no structural breaks then the longer moving average is the optimal length. However, in the developed models, we include a time trend in the series. If there is no time trend, we obtain the longest length as optimal from tables 17 and 18 and figures 9, 11, 13 and 15 as well. For the Bernoulli-jump process, tables 17 and 18 also present the changes of optimal length of moving average according to changes in a jump probability (P) and a jump size (σ_B^2), respectively.

Through the simulated stochastic series based on the permanent-jump and temporary diffusion model for both a Poisson-jump process and a Bernoulli-jump process, clearly the optimal length of moving average is sensitive to the jump frequency and the size of permanent shock.

Table 17. Optimal Length of Moving Average for Wheat Basis Series According to Changes in Mean of Jump Probability and Jump Size from Permanent-Jump Processes

Poisson-Jump	Probability (λ) of Jumps										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Jump size $\left(\sigma_J^2\right) < 0.1315$											
Optimal length (N)	3	2	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) < 0.1315$ and No time trend $(\gamma) = 0$											
Optimal length (N)	10	10	5	4	3	3	2	2	2	2	2
Jump size $\left(\sigma_J^2\right) = 0.1315$											
Optimal length (N)	3	1	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) = 0.1315$ and No time trend $(\gamma) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) > 0.1315$											
Optimal length (N)	3	1	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) > 0.1315$ and No time trend $(\gamma) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1
Bernoulli-Jump	Probability (P) of One Jump										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Jump size $(\sigma_B^2) < 0.1536$											
Optimal length (N)	3	2	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) < 0.1536$ and No time trend $(\beta) = 0$											
Optimal length (N)	10	8	5	4	3	3	2	2	2	2	2
Jump size $(\sigma_B^2) = 0.1536$											
Optimal length (N)	3	1	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) = 0.1536$ and No time trend $(\beta) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) > 0.1536$											
Optimal length (N)	3	1	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) > 0.1536$ and No time trend $(\beta) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1

Figure 8. Optimal Length of Moving Average for Simulated Wheat Basis Series from Permanent-Poisson Jump Process

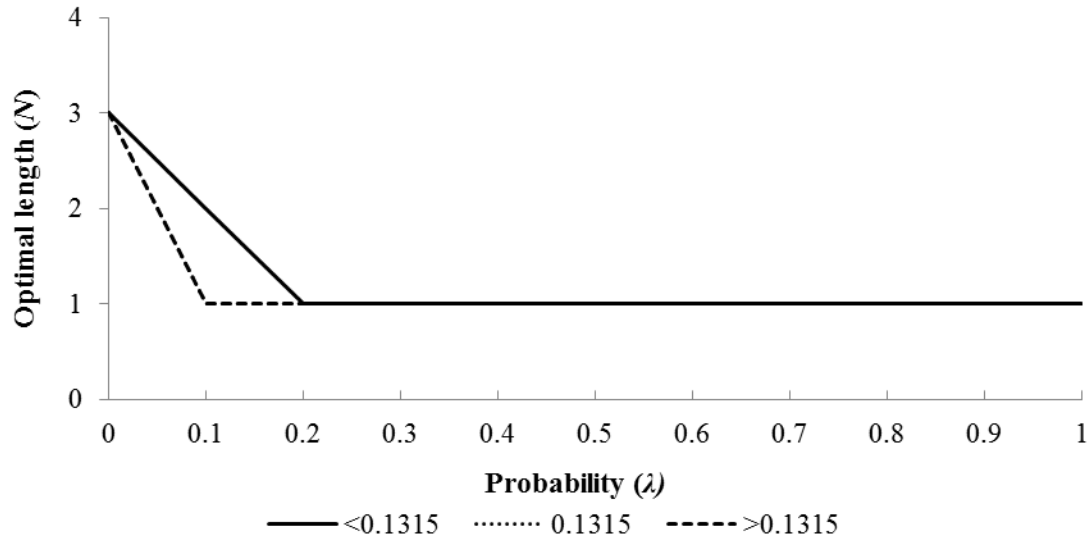


Figure 9. Optimal Length of Moving Average for Simulated with No Trend for Wheat Basis Series from Permanent-Poisson Jump Process

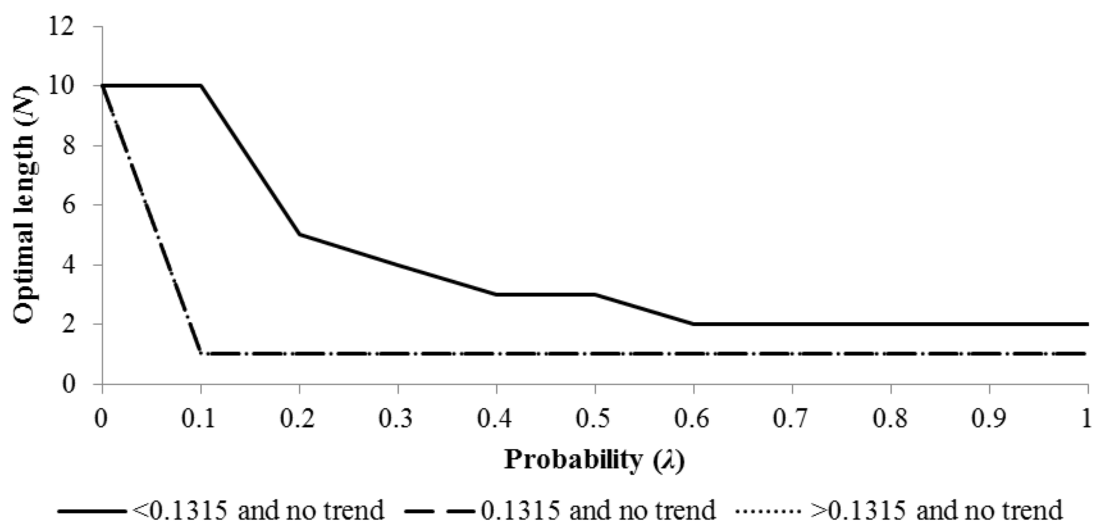


Figure 10. Optimal Length of Moving Average for Simulated Wheat Basis Series from Permanent-Bernoulli Jump Process

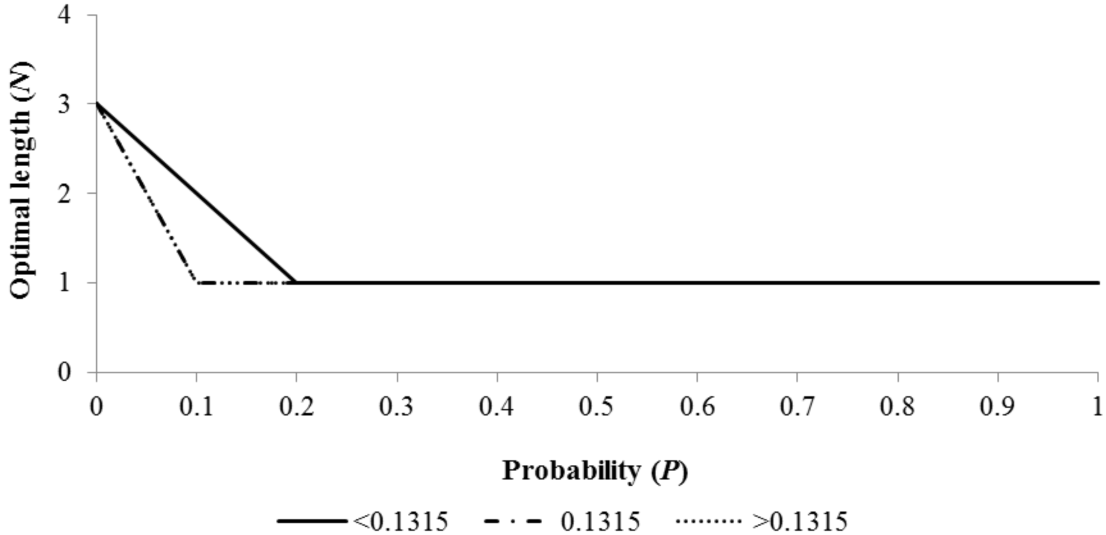


Figure 11. Optimal Length of Moving Average for Simulated with No Trend for Wheat Basis Series from Permanent-Poisson Jump Process

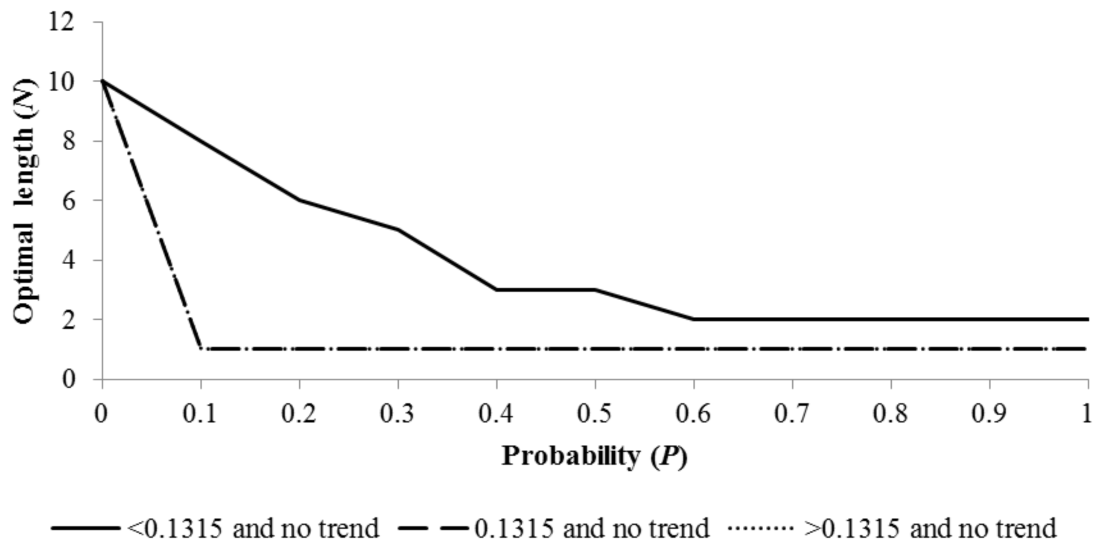


Table 18. Optimal Length of Moving Average for Stock Prices According to Changes in Mean of Jump Probability and Jump Size from Permanent-Jump Processes

Poisson-Jumps	Probability (λ) of Jumps										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Jump size $\left(\sigma_J^2\right) < 0.0006$											
Optimal length (N)	2	1	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) < 0.0006$ and No time trend $(\gamma) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) = 0.0006$											
Optimal length (N)	2	1	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) = 0.0006$ and No time trend $(\gamma) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) > 0.0006$											
Optimal length (N)	2	1	1	1	1	1	1	1	1	1	1
Jump size $\left(\sigma_J^2\right) > 0.0006$ and No time trend $(\gamma) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1
Bernoulli-Jump	Probability (P) of One Jump										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
Jump size $(\sigma_B^2) < 0.0011$											
Optimal length (N)	2	1	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) < 0.0011$ and No time trend $(\beta) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) = 0.0011$											
Optimal length (N)	2	1	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) = 0.0011$ and No time trend $(\beta) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) > 0.0011$											
Optimal length (N)	2	1	1	1	1	1	1	1	1	1	1
Jump size $(\sigma_B^2) > 0.0011$ and No time trend $(\beta) = 0$											
Optimal length (N)	10	1	1	1	1	1	1	1	1	1	1

Figure 12. Optimal Length of Moving Average for Simulated Stock Prices from Permanent-Poisson Jump Process

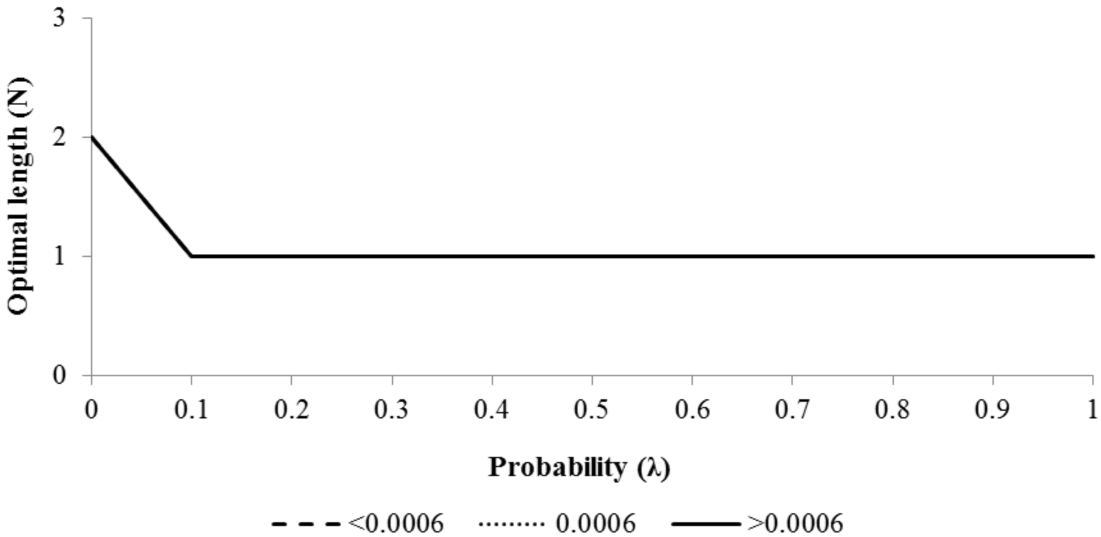


Figure 13. Optimal Length of Moving Average for Simulated with No Trend for Stock Prices from Permanent-Poisson Jump Process

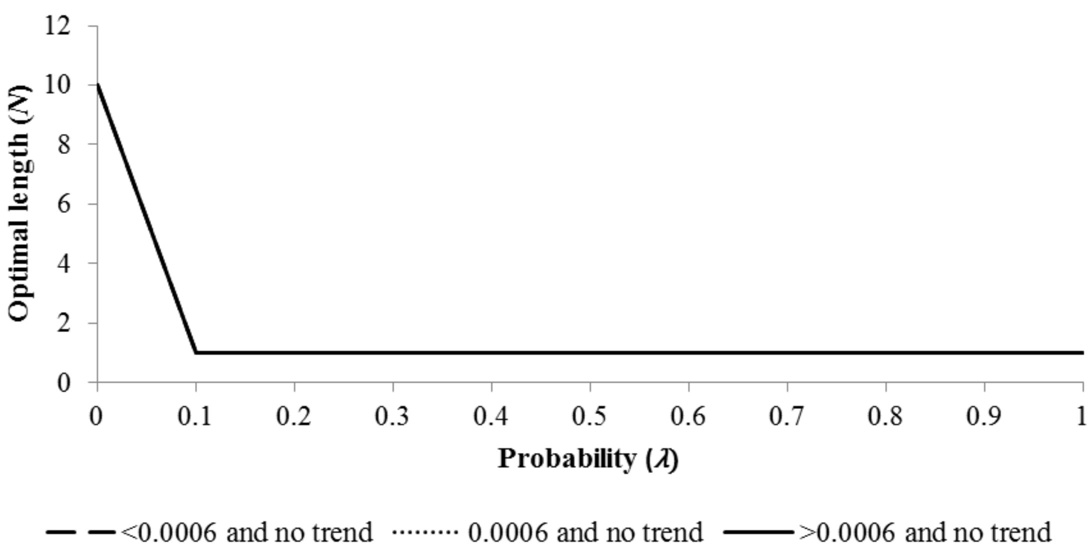


Figure 14. Optimal Length of Moving Average for Simulated Stock Prices from Permanent-Bernoulli Jumps Process

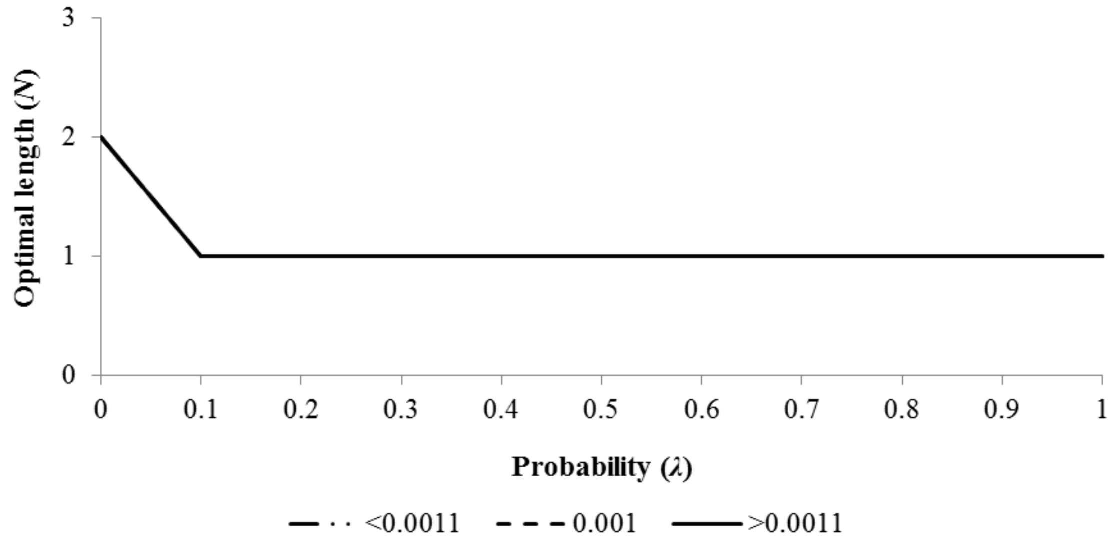
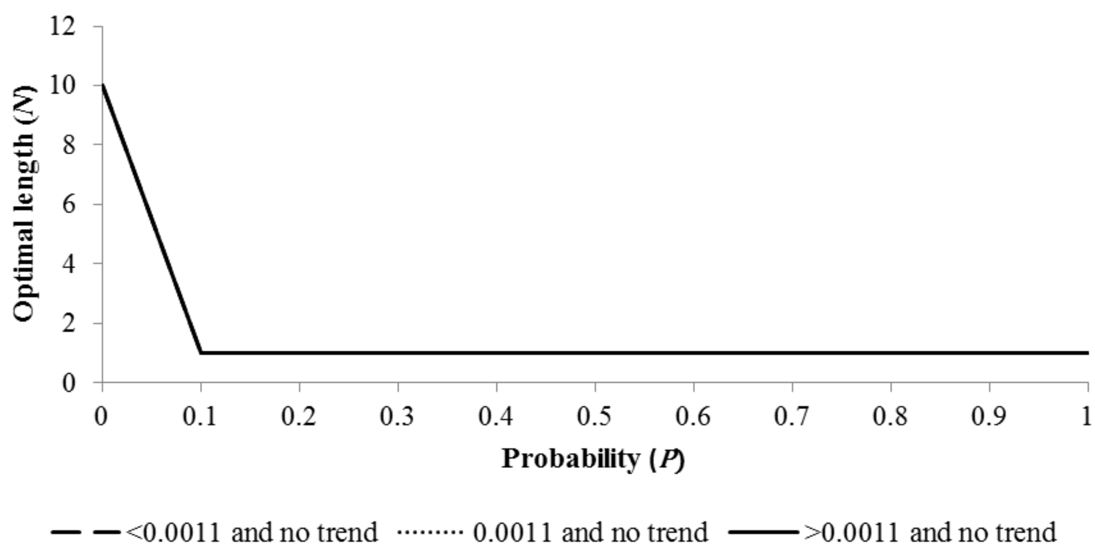


Figure 15. Optimal Length of Moving Average for Simulated with No Trend for Stock Prices from Permanent-Bernoulli Jumps Process



ARIMA Models with Outliers

The ARIMA models with outliers are estimated in order to compare the performance with the developed models. Detection and identification of outliers is important process in ARIMA models because outliers can lead to biased parameter estimation and poor forecasts. Based on the general ARIMA specification procedure, we estimate ARIMA models with first differences. With the ARIMA models, we detect level shift (LS) outliers causing the level or mean changes of the series and transient change (TC) outliers to match the concept of temporary shocks.

Tables 19-24 report the results of the ARIMA model with outliers. From table 19, under the requirement of specific types of outliers such as level shift and transient changes, the model finds two outliers in Oklahoma harvest wheat basis. One outlier occurs in 2007-2008 crop year and has a temporary effect on the series. The other one occurs in 2010-2011 year and has a permanent effect on the series. When we compare the findings of outliers with the graph of the series from figure 1, the findings tends to follow visual inspection. Table 20 reports the results of the ARIMA model with outliers for Illinois harvest corn basis, the model finds a transient outlier in 2010- 2011. From table 21 presenting the results with Illinois soybean harvest basis, the model finds no outliers. In figure 3, there are no extraordinary observations near the end of the period, and the ARIMA model with outlier does not detect any outlier in soybean basis. With wheat, corn and soybean bases for harvest, detection of outliers in this method matches the visual inspection. The value of Akaike information criterion (AIC) that measures the relative goodness of fit of a model is reported in each table. From the AIC value, ARIMA models with outliers always fit data better than conventional ARIMA model.

Table 19 ARIMA with Outliers for Oklahoma Harvest Wheat Basis

Parameters	Estimates	*P*-value	Years of Outliers	Types
ARIMA(0,1,1)				
MA(1)	0.764	<.0001		
AIC	-58.176			
ARIMA (0,1,1) with LS and TC				
MA(1)	0.846	<.0001		
Outlier 1	0.254	0.0002	2011	Level Shift
Outlier 2	-0.433	<.0001	2007	Transient Change
AIC	-80.668			

Table 20 ARIMA with Outliers for Illinois Harvest Corn Basis

Parameters	Estimates	*P*-value	Years of Outliers	Types
ARIMA(0,1,1)				
MA(1)	0.737	<.0001		
AIC	1.236			
ARIMA (0,1,1) with LS and TC				
MA(1)	0.769	<.0001		
Outlier 1	0.505	0.0002	2010	Transient Change
AIC	-8.894			

Table 21 ARIMA with Outliers for Illinois Harvest Soybean Basis

Parameters	Estimates	*P*-value	Years of Outliers	Types
ARIMA(0,1,1)				
MA(1)	0.810	<.0001		
AIC	-4.391			
No Outliers				

Tables 17-19 present the results of ARIMA models with outliers for three macroeconomic series out of Nelson and Plosser's (1982) fourteen data sets. With Nelson and Plosser (1982) data sets, Perron tested the unit-root hypothesis against the alternative hypothesis of trend stationarity with a break to estimate the impacts of either the Great Crash of 1929 or the 1973 oil-price shock. He assumed that each series has only one shock having a permanent effect either at 1929 or at 1973. With money stock, stock prices and total employment series, we allow finding level shift outliers and transient outliers. Table 17 repots the results of money stock and the model finds two transient outliers with one occurring between 1918 and 1919 and another between 1931 and 1932. However, the model did not find a permanent shock at either 1929 or 1973. For the stock prices, the model finds two transient outliers and one level shift outlier in table 18. The level shift outlier occurred between 1921 and 1922 and it is not at the Great Crash or at the oil price shock. From the total employment series, we find one level shift outlier during 1891-1892 and one transient outlier after the level shift, however, findings of outliers in ARIMA model with outliers are not consistent with Perron's assumption that there exist a permanent shock either at 1929 or at 1973.

The ARIMA models with outliers fit data better than the conventional ARIMA model. The method tends to capture outliers near the end of the observation period for wheat and corn basis. For money stock, stock prices and total employment series, the method does not capture the expected structural breaks such as the Great Crash of 1929 and the 1973 oil price shock.

Table 22 ARIMA with Outliers for Money Stock

Parameters	Estimates	*P*-value	Years of Outliers	Types
ARIMA(2,1,0)				
AR(1)	0.839	<.0001		
AR(2)	-0.012	0.9061		
AIC	-316.539			
ARIMA (2,1,0) with LS and TC				
AR(1)	0.763	<.0001		
AR(2)	0.083	0.4189		
Outlier 1	0.052	0.1636	1919	Transient Change
Outlier 2	-0.092	0.0134	1932	Transient Change
AIC	-119.457			

Table 23 ARIMA with Outliers for Stock Prices

Parameters	Estimates	*P*-value	Years of Outliers	Types
ARIMA (0,1,0) with LS and TC				
Outlier 1	0.204	0.1851	1922	Level Shift
Outlier 2	-0.341	0.0050	1931	Transient Change
Outlier 3	-0.294	0.0562	1937	Transient Change
AIC	-103.176			

Table 24 ARIMA with Outliers for Total Employment

Parameter	Estimates	*P*-value	Years of Outliers	Types
ARIMA(1,1,0)				
AR(1)	0.435	<.0001		
AIC	-376.994			
ARIMA (1,1,0) with LS and TC				
AR(1)	0.306	0.0018		
Outlier 1	-0.075	0.0170	1893	Transient Change
Outlier 2	0.018	0.0002	1892	Level Shift
AIC	-388.329			

Calibration Tests

We look at how well the developed model fits data, we compare the cumulative density function (CDF) between an actual time series and a stochastic series. Table 25 reports the results of the Kolmogorov-Smirnov (K-S) test of the empirical CDFs of harvest wheat basis and stock price under the theoretical CDFs of each simulated series based on the proposed models and competing models. We choose harvest wheat basis and stock price as each example for the proposed model with negative autocorrelation and the adjusted model with positive autocorrelation. First, we simulate observations based on the permanent-jump and temporary shocks models, the conventional ARIMA model, and the ARIMA model with outliers, in 1000 replications with the sample size as corresponding actual series. The null hypothesis is that the observed cumulative distribution function for an actual series is from a specific theoretical distribution, which is estimated by the developed models and the competing models. The *D* statistic is the maximum distance between the observed and theoretical cumulative distribution functions. The critical values for the K-S statistic are from the statistical table.

Table 25 presents the results of K-S test between empirical CDF of an actual series and theoretical CDF of the specified distribution models. For the empirical CDF of harvest wheat basis, we reject the null hypothesis that the empirical CDF of the harvest wheat basis is from the theoretical CDF of the conventional ARIMA model at the 5% significance level. For the theoretical CDF of the ARIMA model with outliers, the null hypotheses could be not rejected at the 5% significance level in both wheat basis and stock price. For the theoretical CDFs of the Poisson-jump model and the Bernoulli-jump model, we fail to reject the null hypothesis at the 5% significance level, respectively. The results of K-S test show that the harvest wheat basis and stock price more reasonably are from the mixed distributions of Poisson (Bernoulli) and normal than the normal distribution assumed in convential ARIMA models.

Figures 16-22 graphically display the empirical CDF and the theoretical CDF for harvest wheat basis and stock price. From the results of table 25, we found we fail to reject the null hypothesis that the empirical CDF of the harvest wheat basis is from the theoretical CDF of ARIMA with outliers at the 5% significance level, however, we can see the obvisous difference of the maximum distance between the empirical CDF of an actual series and theoretical CDF of ARIMA with outliers and the empirical CDF of an actual series and theoretical CDF of the proposed models from figures.We impose a Poisson-jump process on permanent shocks and a normal distribution on temporary shocks to find a better stochastic time-series model. According to the K-S, we conclude that the model combined with a Poisson-jump (or Bernoulli-jump) and normal distribution processes reflect the features of data reasonably well.

Table 25. Kolmogorov-Smirnov Test between Empirical CDF of an Actual Series and Theoretical CDF of the Specified Distributions

Data Series	*D* statistic	Critical Values of K-S test
Wheat Basis		
ARIMA	0.209	0.161
ARIMA with LS and TC	0.186	0.161
Poisson-jump process	0.077	0.161
Bernoulli-jump process	0.078	0.161
Stock Prices		
ARIMA with LS and TC	0.182	0.125
Poisson-jump process	0.053	0.125
Bernoulli-jump process	0.077	0.125

Note: The null hypothesis is that the two data sets are from the same distribution.

Figure 16. K-S Test of Empirical CDF of Wheat Basis under Theoretical CDF of ARIMA model

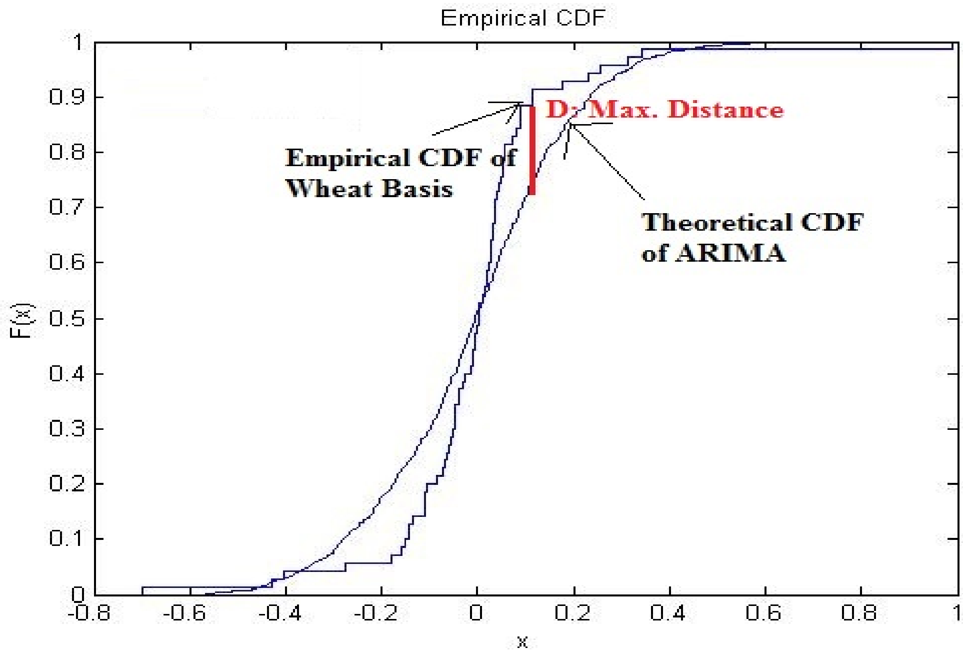


Figure 17. K-S Test of Empirical CDF of Wheat Basis under Theoretical CDF of ARIMA model with Outliers

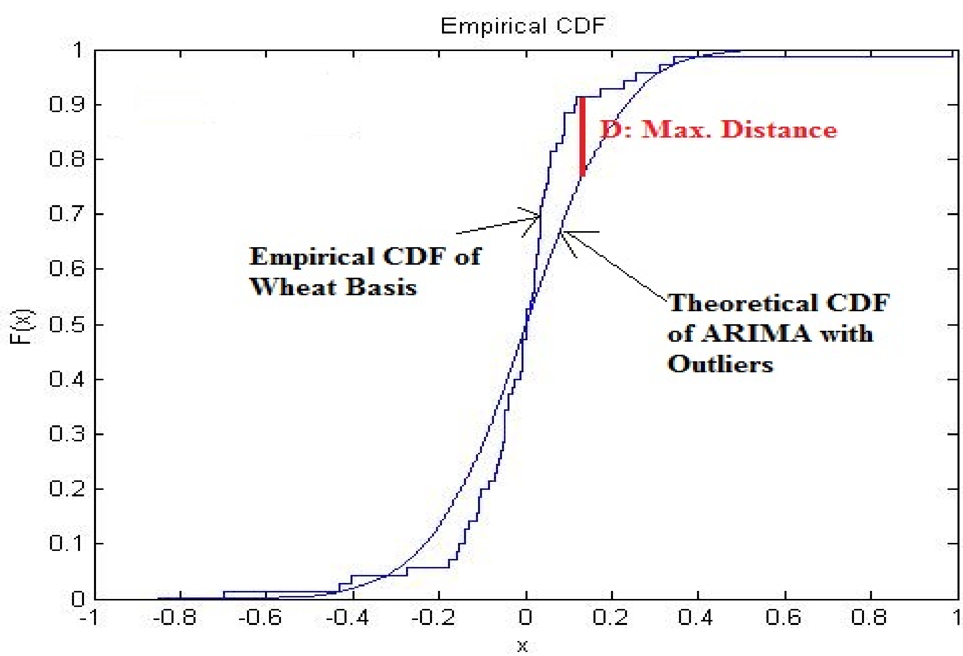


Figure 18. K-S Test of Empirical CDF of Wheat Basis under Theoretical CDF of Poisson-Jump Process

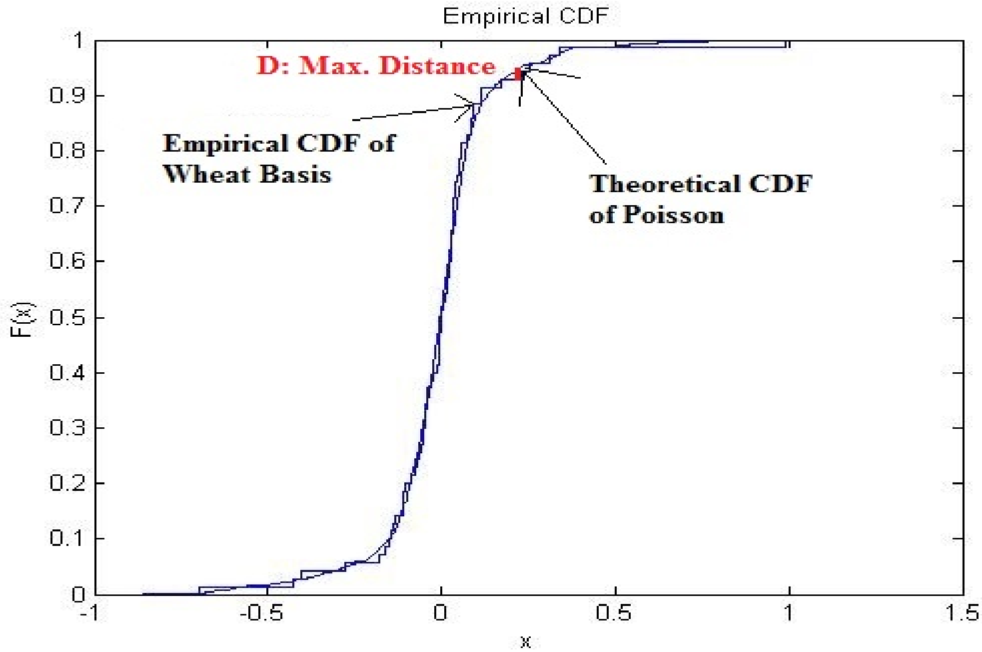


Figure 19. K-S Test of Empirical CDF of Wheat Basis under Theoretical CDF of Bernoulli-Jump Process

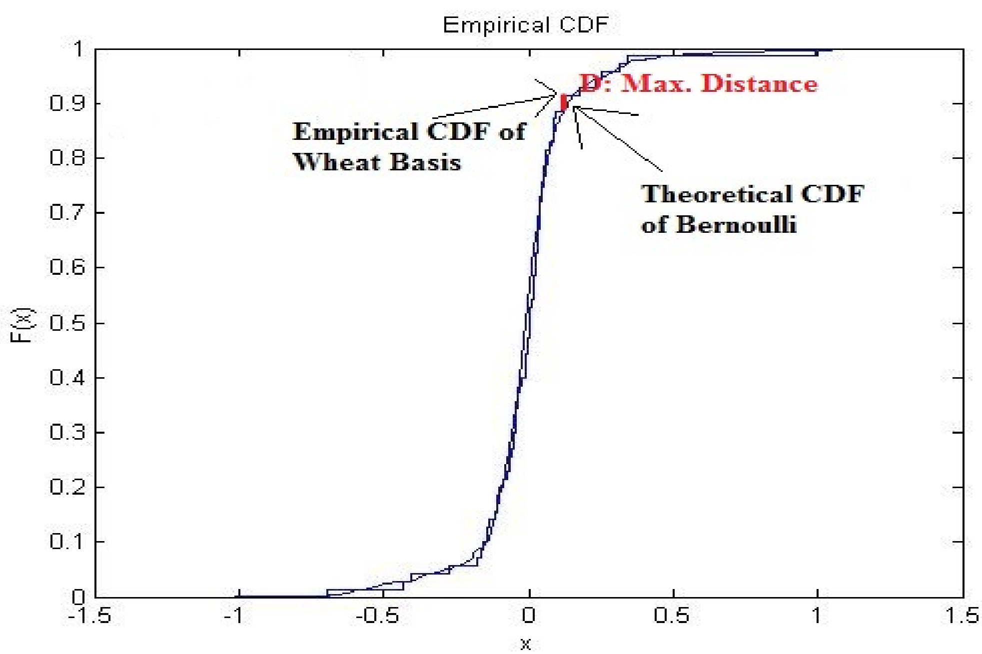


Figure 20. K-S Test of Empirical CDF of Stock Prices under Theoretical CDF of ARIMA model with Outliers

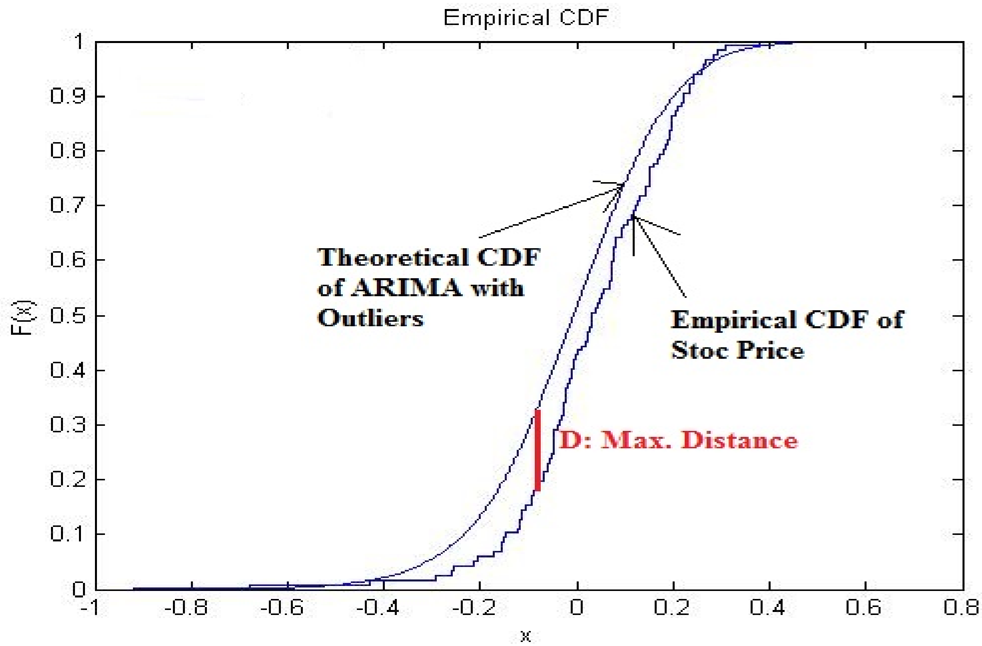


Figure 21. K-S Test of Empirical CDF of Stock Price under Theoretical CDF of Poisson-Jump Process

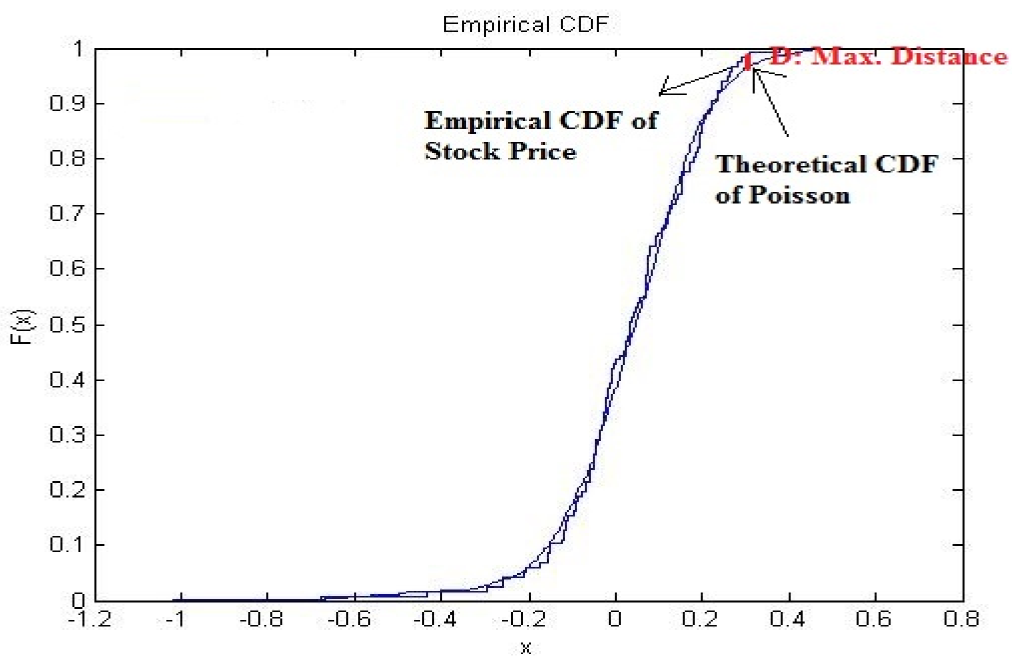
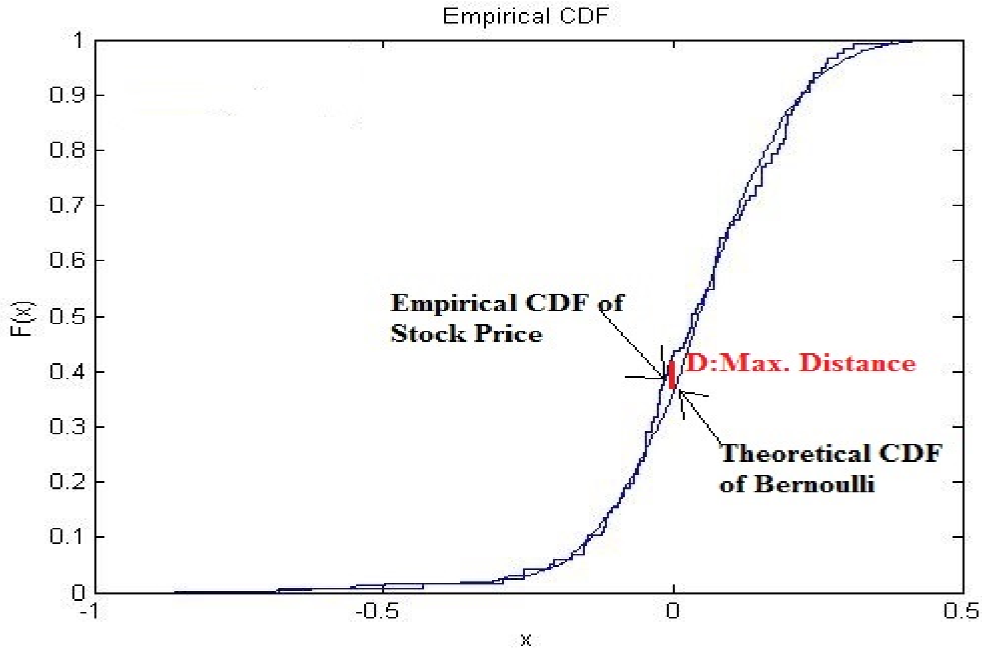


Figure 22. K-S Test of Empirical CDF of Stock Price under Theoretical CDF of Bernoulli-Jump Process



CHAPTER VI

Summary and Conclusion

Many researchers recognize the weaknesses of current unit-root assumptions used in time-series models. The presence of structural breaks in time-series modeling has changed the tests for the unit root hypothesis favored in many time series. Irregular permanent shocks related to structural breaks in a series are a possible explanation of the leptokurtic distributions of many financial time series. Permanent shocks are modeled by specifying a probability distribution rather than by indicator variables. Thus, we take a different approach to build a new time-series model that treats permanent shocks and temporary shocks, differently. We impose a Poisson-jump process and a Bernoulli-jump process to reflect permanent shocks and impose a normal distribution to represent temporary shocks. Oklahoma hard red winter wheat basis for harvest, Illinois corn basis and soybean basis for harvest, money stock, stock prices and total employment and total unemployment rate macroeconomics series are used to estimate the developed model and the relative impacts of permanent shocks related to structural breaks and of transitory shocks. To test the developed model we select four out of Nelson and Plosser's (1982) fourteen series.

With their data sets, many researchers have reported the importance of structural breaks in time-series analyses by several methods. Their studies depend on whether the time that structural breaks occur is assumed known or it is estimated and on the number of structural breaks. In the study, however, we estimate the probability of permanent shocks and not the time points of the breaks in non-stationary series and stationary series. Total unemployment rate is a stationary series while the other six series are non-stationary series.

A temporary shock in the developed model induces negative autocorrelation in the differenced series due to overdifferencing the temporary shocks. We also derive a model for a time series having positive autocorrelation since some time-series have positive autocorrelation. Maximum likelihood estimation (MLE) and generalized method of moments (GMM) with an additional moment condition about autocorrelation are applied to estimate the developed models. Since we have autocorrelation created by overdifferencing temporary shocks, MLE estimation has a computational difficulty to compute numerical optimization. GMM estimation is an alternative approach to handle autocorrelation. An advantage of GMM estimation is that we can have more moment equations than parameters. In the study, we add an additional moment condition about autocorrelation with the GMM procedure. Monte Carlo methods are used to estimate the standard errors of estimates since GMM standard errors are biased. The Bernoulli-jump model has an advantage of encompassing several classic time-series models such as a random walk model with drift and a linear time trend model. With the Bernoulli-jump model, the random walk model cannot be rejected but the linear time trend model is rejected. The critical values of the test statistics are computed using Monte Carlo methods. For a linear time trend model, the computed Wald test critical values are relatively bigger than those obtained assuming an asymptotic chi-squared distribution.

From the developed models, most shocks are permanent except for stock prices as shown by estimated jump variance being relatively bigger than the estimated variance of temporary shocks. Thus, a shorter moving average is preferred to forecast. Based on the results of developed models, we determine the optimal length of moving average to use for forecasts. Two years is the optimal length to use for forecasting wheat basis, corn basis and soybean basis using the actual data, while with the estimated models, one year is the optimal length. For money stock, stock prices and total employment series, the last year is the optimal length for both the actual series and the simulated series. In addition, when jump frequency and size of jump become larger, the optimal length of moving average is the previous year for the time series. That is, the presence of permanent shocks as well as the probability and the variance of permanent shocks clearly reduce the optimal length of moving average. After permanent shocks associated with structural breaks, a shorter moving average is the best for forecasting.

To evaluate the developed model, the autoregressive integrated moving average (ARIMA) model with outliers is selected as a competing model. The ARIMA models with outliers describe a series better than the conventional ARIMA models according to values of AIC. However, outlier detection with money stock, stock prices and total employment series did not detect a level shift outlier either at the 1929 Great crash or at the 1973 oil price shock.

Based on the developed models and the competing model, we test how well the developed models are calibrated to actual series, using the K-S test. Through the K-S test, we provide some implications of the permanent-jump and temporary-diffusion model that specify distributions to better describe permanent and temporary shocks. When we compare the ARIMA models with outliers and the permanent-jump and temporary-diffusion models proposed in the paper, the empirical density curve of proposed models matches a series better than that of the competing ARIMA models with outliers.

There are several potential improvements that could be considered in future research. First, we can concern the robustness of the developed model. Even though quasi-maximum likelihood estimation maximizes the assumed log-likelihood function that does not consider autocorrelation, future research should develop a formal proof of consistency of maximum likelihood estimation without considering autocorrelation. The GMM estimation including an additional moment equation is used to deal with autocorrelation. Second, the additional moment equation can cause over identification problems in GMM estimation and thus the *J*-statistics is suggested for a well specified overidentified model. Third, the proposed models can be expanded by adding a diffusion process for temporary shocks. Finally, we can consider to estimate the proposed model with the multivariate time series because a time series can be influenced by one or more variables.

REFERENCES

Andersen, T. G. and B. E. Sorensen. 1996. " GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study." *Journal of Business & Economic Statistics* 14:328-352.

Akgiray, V. and G. Booth. 1988. "Mixed Diffusion-Jump Process Modeling of Exchange Rate Movements." *The Review of Economics and Statistics* 70:631-637.

Andersen, T.G., H.J. Chung, and B.E. Sorensen. 1999. "Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study." *Journal of Econometrics* 91:61-87.

Baldwin, E. D., and T. R. Smith. 2011. "Soybean Movements in the United States: Interregional Flow Patterns and Transportation Requirements in 1985." Agricultural Experiment Station, Southern Cooperative Series 251, University of Illinois.

Balke, N. S. 1993. "Detecting Level Shifts in Time Series." *Journal of Business & Economic Statistics* 11:81-92.

Ball, C. A., and W. N. Torous. 1985. " On Jumps in Common Stock Prices and Their Impact on Call Option Pricing." *Journal of Finance* 40:155-173.

Banerjee, A., and G. Urga. 2005. "Modeling Structural Breaks, Long Memory and Stock Market Volatility: An Overview." *Journal of Econometrics* 129:1-34.

Beveridge, S., and C. R. Nelson. 1981. " A New Approach to Decomposition of Economic Time series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle. *Journal of Monetary Economics* 7: 151-174.

Bollerslev, T. 1985. "Generalized Autoregressive Conditional Heteroskedasticity." *Journal of Econometrics* 31:307-327.

Bond, S., A. Hoeffler, and J. Temple. 2001. "GMM Estimation of Empirical Growth Models." CEPR Discussion Papers 3048, C.E.P.R Discussion Papers.

Burnside,.C. and Eichenbaum, M. 1996. "Small-Sample Properties of GMM-Based Wald Tests." *Journal of Business & Economic Statistics* 14:294-308.

Campbell, J. Y., and N.G. Mankiw. 1987. "Permanent and Transitory Components in Macroeconomic Fluctuations." *The American Economic Review* 77:111-117.

Chang, I.,G. C. Tiao, and C. Chen. 1988. " Estimation of Time Series Parameters in the Presence of Outliers" *Technometrics* 30:193-20.

Chen, B., and Y. Hong. 2012. "Testing for Smooth Structural Changes in Time Series Models via Nonparametric Regression." *Econometrica* 80:1157-1183.

Chen, C. and L. M. Liu 1993. "Joint Estimation of Model Parameters and Outlier Effects in Time Series." *Journal of American Statistical Association* 88:284-297.

Chow. G. C. 1960. "Tests of Equality between Sets of Coefficients in Two Linear Regressions." *Econometrica* 28:591-605.

Clarida, R. H., and M. P. Taylor. 2003. "Nonlinear Permanent-Temporary Decomposition in Macroeconomics and Finance." *The Economic Journal* 113:125-139.

Commandeur, J.J. F., and S. J. Koopman. 2007. "An Introduction to State Space Time Series Analysis." Oxford: Oxford University Press.

Commandeur J.J. F., S. J. Koopman, and M. Ooms. 2011. "Statistical Software for State Space Methods." *Journal of Statistical Software* 41:1-18.

Dhuyvetter, K.C., and T.L. Kastens. 1998. "Forecasting Crop Basis: Practical Alternatives." Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference, pp.49-67.

Dickey, D.A., and W.A. Fuller. 1979. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association* 74:427-431.

Duffie, D., J. Pan, and K. Singleton. 2000. "Transform Analysis and Asset Pricing for Affine Jump-Diffusions." *Econometrica* 68:1343-1376.

Durbin, J. and S. J. Koopman. 2012. *Time Series Analysis by State Space Methods,* 2nd. ed. Oxford: Oxford University Press.

Dykema, A., N.L. Klein, and G. Taylor. 2002. "The Widening Corn Basis in South Dakota: Factors Affecting and the Impact of the Loan Deficiency Payment." Paper presented at WAEA annual meeting, Long Beach, California, 28-31 July.

Engle, R. F. and Smith, A. 1999. "Stochastic Permanent Breaks." *The Review of Economics and Statistics* 81:553-574.

Fan, X., A Felsovalyi, S. A. Sivo, and S. C. Keenan. 2001. SAS for Monte Carlo Studies: A Guide for Quantitative Researchers. SAS Institute Inc., Cary, NC, USA.

Farmdoc. "Illinois Regional Cash Price Database." University of Illinois at Urbana Champaign, 2013. Online. Available at http://farmdoc.illinois.edu/manage/uspricehistory/us_price_history.html.

Fountis, N. G. and D.A. Dickey. 1989. "Testing for a Unit Root Nonstationarity in Multivariate Autoregressive Time Series." *The Annals of Statistics* 17:419-428.

Fox, A.J. 1972. "Outliers in Time Series." *Journal of the Royal Statistical Society* 34:240-363

Franses, P. H. and Haldrup, N. 1994. "The Effects of Additive Outliers on Tests for Unit Roots and Cointegration." *Journal of Business & Economic Statistics* 12:471-478.

Gallant, A.R., and G.E. Tauchen. 1996. "Which Moments to Match?" *Econometric Theory* 12:657-681.

Greene, W. H. 2011. *Econometric Analysis*. 7th ed. New Jersey: Prentice-Hall.

Hall, J. A., B. W. Brorsen, and S. H. Irwin, 1989. "The Distribution of Futures Prices: A Test of the Stable Paretian and Mixture of Normals Hypotheses." *Journal of Financial and Quantitative Analysis* 24(1):105-116.

Hamilton, J.D. 1994. *Time Series Analysis*, 1st. ed. New York: Princeton University Press.

Hansen, B. E. 1996. "Inference When a Nuisance Parameter Is Not Identified Under the Null Hypothesis." *The Econometric Society* 64:413-430.

Hansen, B. E. 2001."The New Econometrics of Structural Changes: Dating Breaks in U.S. Labor Productivity." *The Journal of Economic Perspectives* 4:117-128

Hansen, B. E and K.D. West. 2002. "Generalized Method of Moments and Macroeconomics." *Journal of Business and Economic Statistics* 20:460-469.

Hansen, L.P. 1982. “Large Sample Properties of Generalized Method of Moments Estimators.” *Econometrica* 50:1029-1054.

Hansen, L. P., and K. Singleton. 1982. “Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models.” *Econometrica* 50:239-266.

Hansen, L. P., and R. J. Hodrick. 1980. “Forward Exchange Rates as Optimal Predictors of Future Spot Rates: an Econometric Analysis.” *Journal of Political Economy* 88:829-853.

Harris, D. 1999. “ GMM estimation of Time Series Models.” In L. Matyas, ed. *Generalized Method of Moments Estimation*. Cambridge University Press, pp:149-169

Hatchett, R.B. 2009. “Optimal Length of Moving Average to Use When Forecasting Basis.” M.S. thesis, Oklahoma State University.

Hatchett, B. R., B. W. Brorsen, and K. B. Anderson. 2010. “Optimal Length of Moving Average to Forecast Futures Basis.” *Journal of Agricultural and Resource Economics* 5:18-33.

Heijmans, R.D.H and J.R. Magnus 1986. “Consistent Maximum-Likelihood Estimation with Dependent Observations: The General (Non-normal) Case and the Normal Case.” *Journal of Econometrics* 32:253-285.

Hilliard, J.E., and J.A. Reis. 1998. "Valuation of Commodity Futures and Options under Stochastic Convenience Yields, Interest Rates, and Jump Diffusions in the Spot." *Journal of Financial and Quantitative Analysis* 33:61-86.

Hnatkovska, V., V. Marmer, and Yao Tang. 2012. “Comparison of Misspecified Calibrated Models: the Minimum Distance Approach.” *Journal of Econometrics* 169:131-138.

Hull, J. 2005. *Options, Futures, and Other Derivatives Securities*. 6th. ed. New Jersey: Prentice Hall.

Jagannathan, R., G. Skoulakis, and Z. Wang. 2002. “Generalized Method of Moments: Applications in Finance.” *Journal of Business & Economic Statistics* 20 470-481.

Jiang, B., and M. Hayenga. 1997. “Corn and Soybean Basis Behavior and Forecasting: Fundamental and Alternative Approaches.” Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Proceedings of the NCR-134 Conference, pp.125-140.

Koekebakker, S., and G. Lien. 2004. “Volatility and Price Jumps in Agricultural Futures Prices Evidence from Wheat Options.” *American Journal of Agricultural Economics* 86:1018-1031.

Kou, S. G. 2002. “A Jump-Diffusion Model for Option Pricing.” *Management Science* 48:1086-1101.

Kou, S. G. 2008. “Jump-Diffusion Models for Asset Pricing in Financial Engineering.” *Handbooks in Operations Research and Management Science* 15:73-116.

Lee, B. S. 1995. “The Response of Stock Prices to Permanent and Temporary Shocks to Dividends.” *Journal of Financial and Quantitative Analysis* 30:1-22.

Lee, J., and M.C. Strazicich. 2003. “Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks.” *Review of Economics and Statistics* 85:1082-1089.

Lumsdain, L. R., and D. H. Papell. 1997. “Multiple Trend Breaks and the Unit-Root Hypothesis.” *The Review of Economics and Statistics* 79:212-218

Nelson, C. P., and C.I. Plosser. 1982 “Trends and Random Walks in Macroeconomic Time Series.” *Journal of Monetary Economics* 10:139-162.

Newey, W. K. and K.D. West. 1987. “A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix.” *Econometrica* 55:703-708.

Nunes, L. C., P. Newbold., and C. M. Kuan. 1996. “Spurious Number of Breaks.” *Economics Letters* 50:175-178.

Nunes, L. C. and P. Newbold., and Chung-Min Kuan. 1997. “Testing for Unit Roots with Breaks: Evidence on the Great Crash and the Unit Root Hypothesis Reconsidered.” *Oxford Bulletin of Economics and Statistics* 4:435-448.

Martin, L., J.L. Groenewegen, and E. Pidgeon. 1980. “Factors Affecting Corn Basis in Southwestern Ontario.” *American Journal of Agricultural Economics* 62:107-112.

Melino, A. and S. Turnbull. 1990. “Pricing Foreign Currency Options with Stochastic Volatility." *Journal of Econometrics* 72:219-255.

Merton, R.C. 1976. “Option Pricing When Underlying Stock Returns Are Discontinuous.” *Journal of Financial Economics* 3:125-144.

Perron, P. 1989. “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis.” *Econometrica* 57:1361-1401.

Pesaran, M. H., and B. Pesaran. 1993. “A Simulation Approach to the Problem of Computing Cox’s Statistic for Testing Nonnested Models.” *Journal of Econometrics* 57:377-392.

Pesaran, M.H., D. Pettenuzzo, and A. Timmermann. 2006. “Forecasting Time Series Subject to Multiple Structural Breaks.” *The Review of Economic Studies* 73:1057 -1084.

Pierce, D. 1971. “Least Squares Estimation in the Regression Model with Autoregressive Moving Average Error.” *Biometrika* 58:299-312.

Plosser, C. I., and G. W. Schewert 1978 “Money, Income, and Susspots:Measuring Economic Relationships and the Effects of Differencing.” *Journal of Monetary Economics* 4:637-660.

Proietti, T. 2008. “Missing Data in Time Series: A Note on the Equivalence of the Dummy Variables and the Skipping Approaches.” *Statistics & Probability Letters* 78:257-264.

Purcell, W.D., and S.R. Koontz. 1998. *Agricultural Futures and Options: Principles and Strategies* 2nd ed. New Jersey: Prentice Hall.

Sanchez, M. J. and D. Pena. 2003. “The identification of Multiple Outliers in ARIMA models.” *Communications in Statistics-Theory and Methods* 32:1265-1287.

Sanders, D.R., and M.R. Manfredo. 2006. “Forecasting Basis Levels in the Soybean Complex: A Comparison of Time Series Methods.” *Journal of Agricultural and Applied Economics* 38:513-523.

Stephens, M. A. 1976. “Asymptotic Results for Goodness-of-Fit Statistics with Unknown Parameters.” *Annals of Statistics* 4:357-369.

Stock, J. H. and M. W. Watson. 1988. “Testing for Common Trends.” *Journal of the American Statistical Association* 83:1097-1107.

Tomek, W. G., and R. J. Myers 1993. “Empirical Analysis of Agricultural Commodity Prices: a View Point.” *Applied Economic Perspectives and Policy* 15:181-202

Trivez, F. J. 1993. “Level Shifts, Temporary Changes and Forecasting.” *Journal of Forecast* 14:543-550.

Tsay, R.S. 1986. “Time Series Model Specification in the Presence of Outliers.” *Journal of the American Statistical Association* 81:132-141.

Tsay, R.S. 1988. “Outliers, Level Shifts, and Variance Changes in Time Series.” *Journal of Forecasting* 7:1-20.

Zivot, E. and D. W. K. Andrews. 1992. “Further Evidence on the Great Crash, the Oil Price Shock, and the Unit-Root Hypothesis.” *Journal of Business and Economic Statistics* 10:251-270.

APPENDICES

In A1, five moment equations for the Bernoulli-jump process are computed as:

A1-1

$$\bar{m}_1 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial ll\left(\hat{\tilde{\boldsymbol{\xi}}}, \Delta series_t\right)}{\partial \hat{\beta}}\right)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\sum_{j=0}^{1} \frac{\tilde{P}^{\hat{j}} \cdot \left(1-\tilde{P}\right)^{1-j} \cdot (\Delta series_t - \hat{\beta} - j \cdot \hat{\mu}_B) \cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j \cdot \tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j \cdot \tilde{\sigma}_B^2\right)}} \cdot \sqrt{2}}{(4 \cdot \hat{\sigma}_e^2 + 2 \cdot j \cdot \hat{\sigma}_B^2) \cdot (j)\sqrt{\pi(2 \cdot \hat{\sigma}_e^2 + j \cdot \hat{\sigma}_B^2)}}}{\sum_{j=0}^{1} \frac{1}{2} \cdot \frac{\tilde{P}^{\hat{j}} \cdot \left(1-\tilde{P}\right)^{1-j} \cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j \cdot \tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j \cdot \tilde{\sigma}_B^2\right)}} \cdot \sqrt{2}}{\sqrt{\pi(2 \cdot \hat{\sigma}_e^2 + j \cdot \hat{\sigma}_B^2)}}}\right),$$

A1-2

$$\bar{m}_2 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial ll\left(\tilde{\xi}, \Delta series_t\right)}{\partial \hat{\mu}_B}\right)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\sum_{j=0}^{1} \frac{\tilde{P}^{\hat{j}} \cdot \left(1-\tilde{P}\right)^{1-j} \cdot (\Delta series_t - \hat{\beta} - j \cdot \hat{\mu}_B) \cdot j \cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j \cdot \tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j \cdot \tilde{\sigma}_B^2\right)}} \cdot \sqrt{2}}{(4 \cdot \hat{\sigma}_e^2 + 2 \cdot j \cdot \hat{\sigma}_B^2) \cdot (j)\sqrt{\pi(2 \cdot \hat{\sigma}_e^2 + j \cdot \hat{\sigma}_B^2)}}}{\sum_{j=0}^{1} \frac{1}{2} \cdot \frac{\tilde{P}^{\hat{j}} \cdot \left(1-\tilde{P}\right)^{1-j} \cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j \cdot \tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j \cdot \tilde{\sigma}_B^2\right)}} \cdot \sqrt{2}}{\sqrt{\pi(2 \cdot \hat{\sigma}_e^2 + j \cdot \hat{\sigma}_B^2)}}}\right),$$

A1-3

$$\bar{m}_3 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial ll(\hat{\xi}, \Delta series_t)}{\partial \hat{\sigma}_e^2}\right)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\left(\left(\left(\sum_{q=0}^{Q}\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot(\Delta series_t - \hat{\beta} - j\cdot\hat{\mu}_B)\cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j\cdot\tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j\cdot\tilde{\sigma}_B^2\right)}}\cdot\sqrt{2}}{(2\cdot\hat{\sigma}_e^2 + 2\cdot j\cdot\hat{\sigma}_B^2)\cdot(j)\sqrt{\pi(2\cdot\hat{\sigma}_e^2 + j\cdot\hat{\sigma}_B^2)}} - \frac{1}{2}\right.\right.\right.$$

$$\left.\cdot\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j\cdot\tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j\cdot\tilde{\sigma}_B^2\right)}}\cdot\sqrt{2}}{(\pi(2\cdot\hat{\sigma}_e^2 j\cdot\hat{\sigma}_B^2))^{3/2}}\right)$$

$$\left./\left(\sum_{j=0}^{1}\frac{1}{2}\cdot\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j\cdot\tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j\cdot\tilde{\sigma}_B^2\right)}}\cdot\sqrt{2}}{\sqrt{\pi(2\cdot\hat{\sigma}_e^2 + j\cdot\hat{\sigma}_B^2)}}\right)\right),$$

A1-4

$$\bar{m}_4 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial ll(\hat{\xi}, \Delta series_t)}{\partial \hat{\sigma}_B^2}\right)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\left(\left(\left(\sum_{j=0}^{1}\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot(\Delta series_t - \hat{\beta} - j\cdot\hat{\mu}_B)\cdot j\cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j\cdot\tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j\cdot\tilde{\sigma}_B^2\right)}}\cdot\sqrt{2}}{(2\cdot\hat{\sigma}_e^2 + 2\cdot j\cdot\hat{\sigma}_B^2)\cdot(j)\sqrt{\pi(2\cdot\hat{\sigma}_e^2 + j\cdot\hat{\sigma}_B^2)}} - \frac{1}{2}\right.\right.\right.$$

$$\left.\cdot\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot j\cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j\cdot\tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j\cdot\tilde{\sigma}_B^2\right)}}\cdot\sqrt{2}}{(\pi(2\cdot\hat{\sigma}_e^2 + j\cdot\hat{\sigma}_B^2))^{3/2}}\right)$$

$$\left./\left(\sum_{j=0}^{1}\frac{1}{2}\cdot\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j\cdot\tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j\cdot\tilde{\sigma}_B^2\right)}}\cdot\sqrt{2}}{\sqrt{\pi(2\cdot\hat{\sigma}_e^2 + j\cdot\hat{\sigma}_B^2)}}\right)\right),$$

A1-5

$$\bar{m}_5 = \frac{1}{N}\sum_{t=1}^{N}\left(\frac{\partial ll(\hat{\xi}, \Delta series_t)}{\partial \hat{\lambda}}\right)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\left(\sum_{j=0}^{1}\left(-\frac{1}{2}\cdot\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot\left(1-\tilde{P}\right)\cdot j\cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j\cdot\tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j\cdot\tilde{\sigma}_B^2\right)}}\cdot\sqrt{2}}{\left(1-\tilde{P}\right)\sqrt{\pi(2\cdot\hat{\sigma}_e^2 + j\cdot\hat{\sigma}_B^2)}} + \frac{1}{2}\right.\right.$$

$$\left.\cdot\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot j\cdot e^{\frac{-\left(\Delta series_t - \tilde{\gamma} - q\tilde{\mu}_J\right)^2}{2\left(2\tilde{\sigma}_\varepsilon^2 + q\tilde{\sigma}_J^2\right)}}\cdot\sqrt{2}}{\tilde{P}\cdot\sqrt{\pi(2\cdot\hat{\sigma}_e^2 + j\cdot\hat{\sigma}_B^2)}}\right)$$

$$\left./\left(\sum_{j=0}^{1}\frac{1}{2}\cdot\frac{\tilde{P}^{j}\cdot\left(1-\tilde{P}\right)^{1-j}\cdot e^{\frac{-\left(\Delta series_t - \tilde{\beta} - j\cdot\tilde{\mu}_B\right)^2}{2\left(2\tilde{\sigma}_e^2 + j\cdot\tilde{\sigma}_B^2\right)}}\cdot\sqrt{2}}{\sqrt{\pi(2\cdot\hat{\sigma}_e^2 + j\cdot\hat{\sigma}_B^2)}}\right)\right).$$

In A2, For a Bernoulli-jump process, the computation for theoretical autocorrelation is derived as:

A2-1
$$corr(\Delta series_t, \Delta series_{t-1}) = \frac{cov(\Delta series_t, \Delta series_{t-1})}{\sqrt{var(\Delta series_t)} \cdot \sqrt{var(\Delta series_{t-1})}}$$

$$= \frac{E[(\Delta series_t - \overline{\Delta series})(\Delta series_{t-1} - \overline{\Delta series})]}{var(\Delta series_t)}$$

$$= \frac{E[(e_t - e_{t-1})(e_{t-1} - e_{t-2})]}{var(\Delta series_t)}$$

$$= \frac{E(-e_{t-1} \cdot e_{t-1})}{var(\beta + B_t \cdot J_t + e_t - e_{t-1})}$$

$$= \frac{-var(e_{t-1})}{var(B_t \cdot J_t) + var(e_t) + var(e_{t-1})}$$

$$= \frac{-\sigma_e^2}{P \cdot \mu_B^2 - (P^2 \cdot \mu_B^2) + P \cdot \sigma_B^2 + 2 \cdot \sigma_e^2},$$

where for $var(B_t \cdot J_t)$

$$var(B_t \cdot J_t) = E(B_t)^2 \cdot var(J_t) + var(B_t) \cdot E(J_t)^2 + var(B_t) \cdot var(J_t)$$

$$= P \cdot \mu_B^2 - (P^2 \cdot \mu_B^2) + P \cdot \sigma_B^2.$$

In A3, for the adjusted Bernoulli-jump process, the derivation of theoretical autocorrelation follows as:

A3-1
$$corr(\Delta series_t, \Delta series_{t-1}) = \frac{cov(\Delta series_t, \Delta series_{t-1})}{\sqrt{var(\Delta series_t)} \cdot \sqrt{var(\Delta series_{t-1})}}$$

$$= \frac{E[(\Delta series_t - \overline{\Delta series})(\Delta series_{t-1} - \overline{\Delta series})]}{var(\Delta series_t)}$$

$$= E[(B_t \cdot J_t + \theta \cdot B_{t-1} \cdot J_{t-1} + e_t - e_{t-1})(B_{t-1} \cdot J_{t-1} + \theta \cdot B_{t-2} \cdot J_{t-2} + e_{t-1} - e_{t-2})]/var(\Delta series_t)$$

$$= \frac{E[\theta \cdot (B_{t-1} \cdot J_{t-1})^2 - (e_{t-1})^2]}{var(\beta + B_t \cdot J_t + \theta \cdot B_{t-1} \cdot J_{t-1} + e_t - e_{t-1})}$$

$$= \frac{\theta \cdot E(B_{t-1}J_{t-1})^2 - E(e_{t-1})^2}{var(B_t \cdot J_t) + \theta^2 \cdot var(B_{t-1} \cdot J_{t-1}) + var(e_t) + var(e_{t-1})}$$

$$= \frac{\theta \cdot (P \cdot \mu_B^2 - (P^2 \cdot \mu_B^2) + P \cdot \sigma_B^2) - \sigma_e^2}{(1+\theta^2) \cdot (P \cdot \mu_B^2 - (P^2 \cdot \mu_B^2) + P \cdot \sigma_B^2) + 2 \cdot \sigma_e^2},$$

where for $var(B_t \cdot J_t)$

$$var(B_t \cdot J_t) = E(B_t)^2 \cdot var(J_t) + var(B_t) \cdot E(J_t)^2 + var(B_t) \cdot var(J_t)$$

$$= P \cdot \mu_B^2 - (P^2 \cdot \mu_B^2) + P \cdot \sigma_B^2,$$

where for $\theta^2 \cdot var(B_{t-1}J_{t-1})$

$$\theta^2 \cdot var(B_{t-1}J_{t-1})$$

$$= \theta^2 \cdot E(B_{t-1})^2 \cdot var(J_{t-1}) + var(B_{t-1}) \cdot E(J_{t-1})^2 + \theta^2 \cdot var(B_{t-1}) \cdot var(J_{t-1})$$

$$= \theta^2 \cdot (P \cdot \mu_B^2 - (P^2 \cdot \mu_B^2) + P \cdot \sigma_B^2).$$

In A4, we provide SAS program code for the Bernoulli-jump model without autocorrelation.

```
proc iml;
use a;
read all var{dy} into y;
start f(x) global(y);
n=nrow(y);
Pi=3.14;
/**** b="drift", c="jump mean", d="variance", j="jump variance",
p="probability" ****/
/*Compute moment equation*/
/*F.O.C of given log-likelihood*/
b=x[1]; c=x[2]; d=x[3]; j=x[4]; p=x[5];
m1 = ((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))
+p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)));
m2 = p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))));
m3 = ((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)
+2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)));
m4 = (p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/4)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)));
m5 = (-(1/2)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)));

mm1=sum(m1)/n;
mm2=sum(m2)/n;
mm3=sum(m3)/n;
mm4=sum(m4)/n;
mm5=sum(m5)/n;
m=(mm1//mm2//mm3//mm4//mm5);

/*Compute a weighting matrix*/
w11=sum(m1#m1)/n;
w21=sum(m2#m1)/n;w22=sum(m2#m2)/n;
w31=sum(m3#m1)/n;w32=sum(m3#m2)/n;w33=sum(m3#m3)/n;
w41=sum(m4#m1)/n;w42=sum(m4#m2)/n;w43=sum(m4#m3)/n;w44=sum(m4#m4)/n;
w51=sum(m5#m1)/n;w52=sum(m5#m2)/n;w53=sum(m5#m3)/n;w54=sum(m5#m4)/n;w55
=sum(m5#m5)/n;

 w12=w21;w13=w31;w14=w41;w15=w51;w23=w32;w24=w42;w25=w52;
 w34=w43;w35=w53;w45=w54;
```

```
w=(w11||w12||w13||w14||w15)//
  (w21||w22||w23||w24||w25)//
  (w31||w32||w33||w34||w35)//
  (w41||w42||w43||w44||w45)//
  (w51||w52||w53||w54||w55);
v=ginv(w);
q=m`*v*m;
return(q);
finish  f;
x={-0.00804 0.02086 0.0025063 0.1536 0.2059};
optn=j(1,10,.);
optn[1]=0;/*specify a minimization problem*/
optn[2]=3;/*specify the amount of printed output*/
optn[3]=0;/*specify the scaling of the Hessian matrix(HESCAL)*/
optn[5]=0;/*defines th line-search techniqe for the unconstrianed or
linearly constrained*/
optn[8]=0;/*specify types of differences and how to compute the
difference interval*/
tc=repeat(.,1,12);/*termination criteria that are tested in each
iteration*/
tc[9]=1.e-15;/*9 specifies the absolute function convergence
criterion(ABSFTOL)*/
*tc[12]=1.e-2;/*12 specifies the absolute parameter convergence
criterion(ABSXTOL),
Termination requires a small relative parameter change in consecutive
iterations*/
con={  .       .     1.e-5    1.e-5     1.e-5,
       .       .      .       .          .};
call nlpnra(rc,result, "f", x, optn,con)tc=tc;
xopt=result`;
n=nrow(y);  Pi=3.14;
/*Compute variance-covariance matrix*/
b=xopt[1]; c=xopt[2]; d=xopt[3]; j=xopt[4]; p=xopt[5];
m1 = ((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))
+p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)));
m2 = p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))));
m3 = ((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)
+2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)));
m4 = (p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/4)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)));
m5 = (-(1/2)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
```

```
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)));

mm1=sum(m1)/n;
mm2=sum(m2)/n;
mm3=sum(m3)/n;
mm4=sum(m4)/n;
mm5=sum(m5)/n;
m=(mm1//mm2//mm3//mm4//mm5);

/*Compute a weighting matrix*/
w11=sum(m1#m1)/n;
w21=sum(m2#m1)/n;w22=sum(m2#m2)/n;
w31=sum(m3#m1)/n;w32=sum(m3#m2)/n;w33=sum(m3#m3)/n;
w41=sum(m4#m1)/n;w42=sum(m4#m2)/n;w43=sum(m4#m3)/n;w44=sum(m4#m4)/n;
w51=sum(m5#m1)/n;w52=sum(m5#m2)/n;w53=sum(m5#m3)/n;w54=sum(m5#m4)/n;w55
=sum(m5#m5)/n;

 w12=w21;w13=w31;w14=w41;w15=w51;w23=w32;w24=w42;w25=w52;
 w34=w43;w35=w53;w45=w54;
w=(w11||w12||w13||w14||w15)//
  (w21||w22||w23||w24||w25)//
  (w31||w32||w33||w34||w35)//
  (w41||w42||w43||w44||w45)//
  (w51||w52||w53||w54||w55);

v=ginv(w);

/*F.O.C of each moment equation*/
sm11 = (-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))+(1/8)#(1-
p)#(Y-b)##2#exp(-(1/4)#(Y-b)##2/d)
/(d##2#sqrt(Pi#d))-p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j)))
+2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j))))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
-((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))+p#(Y-b-
c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)#sqrt(Pi#(2#d+j))))##2/((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm12 = (-p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j)))
+2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j))))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))-((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-
b)##2/d)/(d#sqrt(Pi#d))
+p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))#p#(Y-b-c)#exp(-
(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/(((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))##2#(4#d+2#j)#sqrt(Pi#(2#d+j)));
```

```
sm13 = (-(1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
+(1/16)#(1-p)#(Y-b)##3#exp(-(1/4)#(Y-b)##2/d)/(d##3#sqrt(Pi#d))
-(1/8)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)#Pi/(d#(Pi#d)##(3/2))
-4#p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
+4#p#(Y-b-c)##3#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))
-p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/((4#d+2#j)#(Pi#(2#d+j))##(3/2)))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))-((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-
b)##2/d)/(d#sqrt(Pi#d))
+p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)#sqrt(Pi#(2#d+j))))#((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-
b)##2/d)
/(d##2#sqrt(Pi#d))-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi
/(Pi#d)##(3/2)+2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))/((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm14 = (-2#p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
+2#p#(Y-b-c)##3#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))
-(1/2)#p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/((4#d+2#j)#(Pi#(2#d+j))##(3/2)))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))-((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)
/(d#sqrt(Pi#d))+p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)#sqrt(Pi#(2#d+j))))#(p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/4)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm15 = (-(1/4)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))
+(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))-((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-
b)##2/d)/(d#sqrt(Pi#d))
+p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))#(-(1/2)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))/((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
```

```
sm21 = -p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
+2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j))#((1/2)#(1-
p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
-((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))+p#(Y-b-
c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))#p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/(((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))##2#(4#d+2#j)#sqrt(Pi#(2#d+j)));
sm22 = -p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
+2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j))#((1/2)#(1-
p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
-2#p##2#(Y-b-c)##2#(exp(-(Y-b-c)##2/(4#d+2#j)))##2
/((4#d+2#j)##2#Pi#(2#d+j)#((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))##2);
sm23 = -4#p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
+4#p#(Y-b-c)##3#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##3#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))-p#(Y-b-
c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)#Pi/((4#d+2#j)#(Pi#(2#d+j))##(3/2)#((1/2)#(1-
p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))-p#(Y-b-c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)#((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-
b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#p#(Y-b-
c)##2#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/(Pi#(2#d+j))##(3/2))/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2);
sm24 = -2#p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)
```

```
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
+2#p#(Y-b-c)##3#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##3#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
-(1/2)#p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi
/((4#d+2#j)#(Pi#(2#d+j))##(3/2)#((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
-p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#(p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/4)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2);
sm25 = (Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))
-p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#(-(1/2)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2);
sm31 = (-(1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
+(1/16)#(1-p)#(Y-b)##3#exp(-(1/4)#(Y-b)##2/d)/(d##3#sqrt(Pi#d))
-(1/8)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)#Pi/(d#(Pi#d)##(3/2))-4#p#(Y-
b-c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))+4#p#(Y-b-
c)##3#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))-p#(Y-b-c)#exp(-(Y-
b-c)##2
/(4#d+2#j))#sqrt(2)#Pi/((4#d+2#j)#(Pi#(2#d+j))##(3/2)))/((1/2)#(1-
p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
-((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))+p#(Y-b-
c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))#((1/8)#(1-p)#(Y-
b)##2#exp(-(1/4)#(Y-b)##2/d)
/(d##2#sqrt(Pi#d))-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)
+2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))/((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm32 = (-4#p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
```

```
+4#p#(Y-b-c)##3#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))-p#(Y-b-
c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)#Pi/((4#d+2#j)#(Pi#(2#d+j))##(3/2)))/((1/2)#(1-
p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))-p#(Y-b-
c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)#((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-
b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#p#(Y-b-
c)##2#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#p#exp(-(Y-b-
c)##2
/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))/((4#d+2#j)#sqrt(Pi#(2#d+j))
#((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2);
sm33 = (-(1/4)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-b)##2/d)
/(d##3#sqrt(Pi#d))+(1/32)#(1-p)#(Y-b)##4#exp(-(1/4)#(Y-
b)##2/d)/(d##4#sqrt(Pi#d))
-(1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-
b)##2/d)#Pi/(d##2#(Pi#d)##(3/2))+(3/8)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)#Pi##2
/(Pi#d)##(5/2)-16#p#(Y-b-c)##2#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))+8#p#(Y-b-c)##4#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##4#sqrt(Pi#(2#d+j)))-4#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/((4#d+2#j)##2#(Pi#(2#d+j))##(3/2))+(3/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi##2/(Pi#(2#d+j))##(5/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
-((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#p#(Y-b-
c)##2#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))##2
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm34 = (-8#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))
+4#p#(Y-b-c)##4#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##4#sqrt(Pi#(2#d+j)))
-2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/((4#d+2#j)##2#(Pi#(2#d+j))##(3/2))
+(3/4)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi##2/(Pi#(2#d+j))##(5/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))-((1/8)#(1-p)#(Y-b)##2#exp(-
(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#p#(Y-b-
c)##2#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/(Pi#(2#d+j))##(3/2))#(p#(Y-b-c)##2#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/4)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
```

```
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm35 = (-(1/8)#(Y-b)##2#exp(-(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
+(1/4)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
-((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#p#(Y-b-
c)##2#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/(Pi#(2#d+j))##(3/2))#(-(1/2)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)+(1/2)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))/((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm41 = (-2#p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))+2#p#(Y-b-c)##3#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))-(1/2)#p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/((4#d+2#j)#(Pi#(2#d+j))##(3/2)))/((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))-
((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)
/(d#sqrt(Pi#d))+p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)#sqrt(Pi#(2#d+j))))#(p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/4)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm42 = (-2#p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
+2#p#(Y-b-c)##3#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))
-(1/2)#p#(Y-b-c)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/((4#d+2#j)#(Pi#(2#d+j))##(3/2)))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))-p#(Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#(p#(Y-
b-c)##2#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/4)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/(Pi#(2#d+j))##(3/2))/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2);
sm43 = (-8#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))
+4#p#(Y-b-c)##4#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##4#sqrt(Pi#(2#d+j)))
-2#p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/((4#d+2#j)##2#(Pi#(2#d+j))##(3/2))
```

```
+(3/4)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi##2/(Pi#(2#d+j))##(5/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))-((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-
b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#p#(Y-b-
c)##2#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/(Pi#(2#d+j))##(3/2))#(p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/4)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))/((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm44 = (-4#p#(Y-b-c)##2#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##3#sqrt(Pi#(2#d+j)))+2#p#(Y-b-c)##4#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##4#sqrt(Pi#(2#d+j)))-p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/((4#d+2#j)##2#(Pi#(2#d+j))##(3/2))+(3/8)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi##2
/(Pi#(2#d+j))##(5/2))/((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))-(p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/4)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi
/(Pi#(2#d+j))##(3/2))##2/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm45 = ((Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/4)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))/((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
-(p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/4)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))#(-
(1/2)#exp(-(1/4)#(Y-b)##2/d
)/sqrt(Pi#d)+(1/2)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm51 = (-(1/4)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))+(Y-b-
c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))/((1/2)#(1-p)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
-((1/4)#(1-p)#(Y-b)#exp(-(1/4)#(Y-b)##2/d)/(d#sqrt(Pi#d))+p#(Y-b-
c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)#sqrt(Pi#(2#d+j))))#(-(1/2)#exp(-
(1/4)#(Y-b)##2/d)
/sqrt(Pi#d)+(1/2)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
```

```
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm52 = (Y-b-c)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j))))-p#(Y-b-
c)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)#(-(1/2)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)
+(1/2)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
/((4#d+2#j)#sqrt(Pi#(2#d+j))#((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2);
sm53 = (-(1/8)#(Y-b)##2#exp(-(1/4)#(Y-b)##2/d)/(d##2#sqrt(Pi#d))
+(1/4)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/sqrt(Pi#(2#d+j)))-((1/8)#(1-p)#(Y-b)##2#exp(-(1/4)#(Y-
b)##2/d)/(d##2#sqrt(Pi#d))
-(1/4)#(1-p)#exp(-(1/4)#(Y-b)##2/d)#Pi/(Pi#d)##(3/2)+2#p#(Y-b-
c)##2#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/2)#p#exp(-(Y-b-
c)##2
/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))#(-(1/2)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm54 = ((Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))
-(1/4)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))-(p#(Y-b-c)##2#exp(-(Y-b-
c)##2/(4#d+2#j))#sqrt(2)
/((4#d+2#j)##2#sqrt(Pi#(2#d+j)))-(1/4)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)#Pi/(Pi#(2#d+j))##(3/2))#(-(1/2)#exp(-(1/4)#(Y-
b)##2/d)
/sqrt(Pi#d)+(1/2)#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))
/((1/2)#(1-p)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#p#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;
sm55 = -(-(1/2)#exp(-(1/4)#(Y-b)##2/d)/sqrt(Pi#d)+(1/2)#exp(-(Y-b-c)##2
/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2/((1/2)#(1-p)#exp(-(1/4)#(Y-
b)##2/d)/sqrt(Pi#d)
+(1/2)#p#exp(-(Y-b-c)##2/(4#d+2#j))#sqrt(2)/sqrt(Pi#(2#d+j)))##2;


sr11=sum(sm11)/n;sr21=sum(sm21)/n;sr31=sum(sm31)/n;sr41=sum(sm41)/n;sr5
1=sum(sm51)/n;
sr12=sum(sm12)/n;sr22=sum(sm22)/n;sr32=sum(sm32)/n;sr42=sum(sm42)/n;sr5
2=sum(sm52)/n;
sr13=sum(sm13)/n;sr23=sum(sm23)/n;sr33=sum(sm33)/n;sr43=sum(sm43)/n;sr5
3=sum(sm53)/n;
sr14=sum(sm14)/n;sr24=sum(sm24)/n;sr34=sum(sm34)/n;sr44=sum(sm44)/n;sr5
4=sum(sm54)/n;
```

```
sr15=sum(sm15)/n;sr25=sum(sm25)/n;sr35=sum(sm35)/n;sr45=sum(sm45)/n;sr5
5=sum(sm55)/n;

r=(sr11||sr21||sr31||sr41||sr51)//
  (sr12||sr22||sr32||sr42||sr52)//
  (sr13||sr23||sr33||sr43||sr53)//
  (sr14||sr24||sr34||sr44||sr54)//
  (sr15||sr25||sr35||sr45||sr55);

h=r`*v*r;
cov=(1/n)*(ginv(h));

se_b=sqrt(cov[1,1]);
se_c=sqrt(cov[2,2]);
se_d=sqrt(cov[3,3]);
se_j=sqrt(cov[4,4]);
se_p=sqrt(cov[5,5]);

/**** b="drift", c="jump mean", d="variance", j="jump variance",
p="probability" ****/
print b c d j p;
print se_b se_c se_d se_j se_p;
print r;
print cov;
quit;
```

VITA

Lee, Yoonsuk

Candidate for the Degree of

Doctor of Philosophy

Thesis: IMPACTS OF PERMANENT AND TEMPORARY SHOCKS ON OPTIMAL LENGTH OF MOVING AVERAGE TO FORECAST A TIME SERIES

Major Field: Agricultural Economics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Agricultural Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2013.

Graduate from Oklahoma State University, Stillwater, Oklahoma in May 2009; received a Master of Science degree in Agricultural Economics.

Graduate from KangWon National University, ChunCheon, South Korea, February 2003; received a Bachelor of Science degree in Agricultural Economics.

Experience: 09/08-05/13 Graduate Research Assistant, Department of Agricultural Economics, Oklahoma State University