THREE ESSAYS ON PRICE DYNAMICS AND CAUSATIONS AMONG ENERGY MARKETS AND MACROECONOMIC INFORMATION

A Dissertation

by

SUNG WOOK HONG

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee,	David Bessler
	Ximing Wu
Committee Members,	Henry Bryant
	Jianhua Huang
Head of Department,	C. Parr Rosson III

December 2012

Major Subject: Agricultural Economics

ABSTRACT

This dissertation examines three important issues in energy markets: price dynamics, information flow, and structural change. We discuss each issue in detail, building empirical time series models, analyzing the results, and interpreting the findings. First, we examine the contemporaneous interdependencies and information flows among crude oil, natural gas, and electricity prices in the United States (US) through the multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) model, Directed Acyclic Graph (DAG) for contemporaneous causal structures and Bernanke factorization for price dynamic processes. Test results show that the DAG from residuals of out-of-sample-forecast is consistent with the DAG from residuals of within-sample-fit. The result supports innovation accounting analysis based on DAGs using residuals of out-of-sample-forecast. Second, we look at the effects of the federal fund rate and/or WTI crude oil price shock on US macroeconomic and financial indicators by using a Factor Augmented Vector Autoregression (FAVAR) model and a graphical model without any deductive assumption. The results show that, in contemporaneous time, the federal fund rate shock is exogenous as the identifying assumption in the Vector Autoregression (VAR) framework of the monetary shock transmission mechanism, whereas the WTI crude oil price return is not exogenous. Third, we examine price dynamics and contemporaneous causality among the price returns of WTI crude oil, gasoline, corn, and the S&P 500. We look for structural break points and then build an econometric model to find the consistent sub-periods having

stable parameters in a given VAR framework and to explain recent movements and interdependency among returns. We found strong evidence of two structural breaks and contemporaneous causal relationships among the residuals, but also significant differences between contemporaneous causal structures for each sub-period.

DEDICATION

To my family, for their love and support

ACKNOWLEDGMENTS

First and foremost, I express heartfelt gratitude to my advisor and co-advisor, Dr. Bessler and Dr. Wu. As independent thinkers, they set a high standard for their students to emulate. This dissertation could not have been written without their support and guidance.

I thank Dr. Bryant and Dr. Huang, whose lengthy discussions inspired me to develop my research in the appropriate direction, and who motivated me to become an active researcher. Their valuable advice always came at just the right time.

I thank my friends and colleagues in College Station for their support and the opinions offered from diverse backgrounds of knowledge. I am lucky to have such good friends and I will miss the enjoyable times we spent together.

Most important, I acknowledge my family. Thank you for always being a constant source of love, concern, and strength.

NOMENCLATURE

AHE	Average hourly earnings
COB	California-Oregon
EMP	Employment and hours
EXR	Exchange rates
FFR	US federal fund rate
HSS	Housing starts and sales
INT	Interest rates
MON	Money and credit quantity aggregates
OIV	Orders and real inventories
OUT	Real output and income
PJM	Pennsylvania-New Jersey-Maryland
PRI	Price indexes
R_{COB}	Price returns of COB Electricity firm on peak
R_{CORN}	Price returns of corn
R_{GAS}	Price returns of Henry Hub natural gas
$R_{GASOLINE}$	Price returns of gasoline
R_{OIL}	Price returns of Dated Brent crude oil
R_{PJM}	Price returns of PJM Electricity firm on peak
$R_{S\&P500}$	Price returns of S&P 500
R_{WTI}	Price returns of WTI crude oil

$R_{VAR}\_R_{COB}$	Residuals in VAR model for the price returns of COB Electricity firm on peak
$R_{VAR}\_R_{GAS}$	Residuals in VAR model for the price returns of Henry Hub natural gas
$R_{VAR}\_R_{OIL}$	Residuals in VAR model for the price returns of Dated Brent crude oil
$R_{VAR}\_R_{PJM}$	Residuals in VAR model for the price returns of PJM Electricity firm on peak
$R_{VDG}\_R_{COB}$	Standardized residuals in VAR-DCC-GARCH model for the price returns of COB Electricity firm on peak
$R_{VDG}\_R_{GAS}$	Standardized residuals in VAR-DCC-GARCH model for the price returns of Henry Hub natural gas
$R_{VDG}\_R_{OIL}$	Standardized residuals in VAR-DCC-GARCH model for the price returns of Dated Brent crude oil
$R_{VDG}\_R_{PJM}$	Standardized residuals in VAR-DCC-GARCH model for the price returns of PJM Electricity firm on peak
SPR	Spreads
STO	Stock prices
UEMP	Unemployment rate
WTI	West Texas Intermediate crude oil

TABLE OF CONTENTS

LIST OF FIGURES

Page

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

The dynamic relations and causality among energy markets and macroeconomic information are one of the most interesting topics in empirical economic studies. Recent advances in econometric techniques have stimulated many researchers to investigate the effects of changes in energy prices (e.g., crude oil, gasoline, natural gas and electricity) on macroeconomic and financial indicators.

However, few have studied the specific impacts on energy markets resulting from dynamic information flows among various sectors of the US economy and structural breaks. Some research has found strong evidence of contemporaneous correlations, dynamic relations and information flows among energy prices, macroeconomic and financial indicators (Barsky and Kilian 2002; Barsky et al. 2004; Bernanke et al. 1997; Hamilton 1983, 1996, 2003; Hoover and Perez 1994).

Therefore, this dissertation provides information on establishing dynamic market processes based on an econometric framework that address price dynamics and causal relationships among US energy markets and macroeconomic information. The dissertation is presented as three essays in Chapters II, III and IV. Each self-contained essay includes an introduction, methodology and conclusion. Chapter V summarizes the results and gives concluding comments.

Chapter II investigates the interdependencies and information flows among crude oil, natural gas, and electricity prices in the US by using the multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) model which we consider superior to using separate univariate models. We perform forecasting exercises considering 200 one-step-ahead forecasts, reclusively and evaluate forecasting performance. Studying stability in the modeling structure by comparing the variance-covariance structures in residuals from within-sample-fit and out-of-sample-forecast allows us to investigate the appropriateness of a standard Directed Acyclic Graphs (DAG) application in time series analysis with respect to modeling new information. Subsequently, we assess whether the DAG structures from residuals of within-sample-fit show the same patterns as the DAG structures from residuals of out-of-sample-forecast.

Chapter III inductively infers the contemporaneous information flows without any deductive information and investigates the structural economic shocks transmission mechanism under the FAVAR framework. We use a two-step procedure to show that the co-movement of these time series over time is adequately described in terms of a number of unobserved latent factors and the US federal fund rate or WTI crude oil price return. First, we extract common factors from a large macroeconomic dataset of the US economy using the method suggested by Stock and Watson (2002a, 2002b) and Bernanke et al. (2005). Second, we estimate the parameters governing their joint dynamics with the US federal fund rate and WTI crude oil price return series in two FAVAR models. Third, we identify the contemporaneous causal structures among innovations based on the residuals of the estimated FAVAR models using the Directed Acyclic Graph (DAG) model. Fourth, we derive and interpret the impulse response functions for each augmented factor and two considered variables and decompose the forecast error variance for each factor into the parts attributable to each of a set of innovations processes in the FAVAR models. Finally, we perform forecasting exercises considering 35 one-step-ahead forecasts, reclusively. This exertion is accompanied by comparing the forecasting performances between the estimated FAVAR and univariate Autoregression (AR) models to check for analytical robustness.

Chapter IV investigates the variations of contemporaneous causal structures among energy, agricultural, and financial markets by identifying structural changes in their dynamic relationships. The finding of structural change allows us to produce one sample before and one sample after the identified change point. However, it is widely known that the accurate directions and magnitude of the linkages are difficult to capture since their dynamic relationships are varied by time and they strengthen/weaken during crisis periods. Thus, careful estimation of the price dynamics is paramount in order to identify structural changes contemporaneous linkages. Bearing this in mind, we build an econometric model to examine whether crude oil, gasoline, corn and the US stock market index are linked contemporaneously and how their relationships change through time and across markets.

Chapter V summarizes the results of the three chapters and lists the key findings. Also, we discuss the shortcomings of this dissertation and suggestions for further research.

CHAPTER II

IDENTIFICATION OF CONTEMPORANEOUS CAUSAL STRUCTURE ON ENERGY MARKETS: WITHIN-SAMPLE-FIT VS. OUT-OF-SAMPLE-FORECAST

2.1. Introduction

Multivariate time series models that include causal relationships among variables and information about random shocks can significantly improve forecasting ability. The common approach for identifying causal structure derives either from economic theory, or the researcher's knowledge of the data (Stock and Watson 2001). The widely accepted concept proposed by Granger (1969), "X_t is cause of Y_t with respect to other series Z_t", known as Granger causality, describes the relationships between time series in the forecasting framework. However, this concept neglects mention of possible contemporaneous causality/correlation between X_t and Y_t (Granger 1988; Lütkepohl and Reimers 1992). Thus, Swanson and Granger (1997) suggested using the residuals from a Vector Autoregression (VAR) model to test for vanishing difference of product of correlation or partial correlation among variables. The Directed Acyclic Graph (DAG) approach, which is based on the graph theory developed by Spirtes et al. (2000) and Pearl (2000) identifies the contemporaneous causal inferences among the variables with relative ease by testing the conditional independence on the residuals (Bessler and Lee 2002; Bessler and Yang 2003; Demiralp and Hoover 2003; Moneta 2004, 2008; Swanson and Granger 1997).

However, evidence of within-sample predictability does not hold in out-of-sample predictability (Granger 1980). Thus, the evaluating predictability and forecasting performance of time series models has remained a crucial issue particularly for economics and econometrics. In general, the predictability tests and methods for evaluating forecasting performance are based on the within-sample-fit of a model and/or the out-of-sample-forecast obtained from a sequence of recursive regression. For within-sample-fit, the entire sample is used in fitting the model, whereas an out-of-sample-forecast attempts to mimic the data constraints (Chatfield 2001). Numerous studies examine the test of predictability and power of forecasting performance (Clemen 1989; Clements and Hendry 1993; Diebold and Lopez 1996; Diebold and Mariano 1995; Granger 1989; Harvey et al. 1997).

As an extension of these stylized facts, our interest lies in assessing whether the causal structure based on residuals from within-sample-fit is the same as the causal structure based on residuals from out-of-sample-forecast. Typically, residuals from within-sample-fit represent the difference between actual value and expected value based on past information, and residuals from out-of-sample-forecast represent the difference between actual value and predicted value based on present and past information (Engle 2001). The usual way to infer contemporaneous causality is to use residuals from within-sample-fit because they are easier to identify and there is less computational burden. If the proposition holds that causal flows based on both residuals from within-sample-fit and out-of-sample-forecast exhibit consistency, we can be confident in the out-of-sample-forecast and its causal results. For example, Kim and

Bessler (2007), who assessed causal relationship on the US equity market by using the Vector Error Correction (VEC) model, claimed that the DAG constructed based on residuals from within-sample-fit is consistent with the DAG based on residuals from out-of-sample-forecast.

In this chapter, we investigate and address the interdependencies and information flows among crude oil, natural gas, and electricity prices in the US. Since the energy crises in the 1970s, a number of studies have looked at the economic impacts of the high volatility of crude oil prices and oil price shocks (Hamilton 1983; Hickman et al. 1987; Jones et al. 2004; Kilian 2008; Mork and Hall 1980; Rasche and Tatom 1977). In recent years, including 2012, dramatic increases in price volatility have even prompted some legislators to call for investigations into the possibility of oil and gas price speculation and market manipulation (Cantwell 2012).

In general, energy prices, such as crude oil, natural gas, and electricity, are often characterized by high volatility, strong mean-reversion, and abrupt and unanticipated upward price jumps or spikes which quickly decay (Blanco and Soronow 2001). However, price volatility is still insufficiently defined and there is no widely accepted definition of adequate volatility modeling and measurement. Thus, whether conditional volatility (expected volatility) and volatility shocks (unexpected volatility) in a specific energy commodity market influence volatility in other commodity markets is a crucial question for diversification of economic issues on market integration.

From this perspective, we investigate the relationships among these series by using the multivariate generalized autoregressive conditional heteroscedasticity

(MGARCH) model, which we consider superior to working with separate univariate models. We perform forecasting exercises considering 200 one-step-ahead forecasts, reclusively. This exertion is accompanied by the evaluation of forecasting performance. Subsequently, we assess whether the DAG structures from standardized residuals of within-sample-fit show the same patterns from standardized residuals of out-of-sample-forecast.

The contributions of this study to the literature on causal modeling and energy markets is two-fold: (1) determining the direction of causalities among the prices in US oil, natural gas, and electricity markets; and (2) whether the information flows between residuals from within-sample and out-of-sample forecast reveal consistency

The remainder of Chapter II is organized as follows. Section 2 discusses VAR, Dynamic Conditional Correlation (DCC) GARCH models and DAG specifications. Section 3 describes the data used in the analysis and presents summary statistics and basic non-stationary test results. Section 4 discusses the empirical analysis of daily price returns of Dated Brent crude oil, Henry Hub natural gas, PJM electricity firm on peak and COB electricity firm on peak. Section 5 concludes.

2.2. Methodology

This section introduces the basic concept of volatility modeling, i.e., decomposing a given time series into predictable and unpredictable parts. Although volatility modeling methods are available in the literature, our focus is on the VAR and MGARCH models. We describe DAG, the basic framework and causal searching algorithm of graphical modeling.

2.2.1. Vector Autoregressive (VAR) Model

Following Sims (1980), a basic VAR model consists of a set of *N* endogenous variables $Y_t = (y_{1t}, y_{2t}, \cdots, y_{nt})$ for $n = 1, 2, \cdots, N$ and $t = 1, 2, \cdots, T$. Therefore, we define a VAR (*p*) process as:

$$Y_t = \mathrm{c} + \sum_{i=1}^{p} \Gamma_i Y_{t-i} + \eta_t \tag{2.1}$$

where Γ_i are $(N \times N)$ coefficient matrices for $i = 1,2, \cdots, p$ and η_t is a *N*-dimensional process with $E(\eta_t) = 0$ and time invariant positive definite covariance matrix $E(\eta_t \eta_t^T) = \Sigma_\eta$. For any given variable, we estimate the coefficients of a VAR (*p*) process using by Ordinary Least Squares (OLS) separately for each time series.

A VAR approach consists of generating stationary time series with time invariant means, variances and covariance structure, given sufficient starting values. However, the reduced form VAR above (equation (2.1)) does not allow for contemporaneous dependent relationships. One way of allowing contemporaneous dependency is to multiply a matrix Γ_0 on both sides of equation (2.1):

$$\Gamma_0 Y_t = d + \sum_{i=1}^{p} \Gamma_{\mathrm{i}}^{*} Y_{t-i} + \varepsilon_t \tag{2.2}$$

where a matrix Γ_0 represents the causal dependency of each variable on its contemporaneous counterparts, which is upper triangular with a unit diagonal; ε_t can be expressed as $\Gamma_0 \eta_t$, which is the diagonal matrix.

Fitting a VAR approach for modeling the conditional mean equation of the price returns in energy markets is a natural extension of this methodology in line with our research question. We now describe the relevant specifications of a multivariate GARCH model.

2.2.2. Multivariate GARCH Model

Consider a stochastic vector process $\{Y_t\}$ with dimension $N \times 1$. We denote that ζ_{t-1} is the information set generated by the observed series $\{Y_t\}$ up to and including time $t-1$. Formally, we assume that $\{Y_t\}$ is conditionally heteroscedasticity.

We express the standard multivariate GARCH framework with no linear dependence structure in $\{Y_t\}$ as:

$$Y_t|\zeta_{t-1} \sim U(\mu_t, H_t) \tag{2.3}$$

where $U(\mu_t, H_t)$ is an un-specified multivariate distribution with time dependent mean μ_t and time dependent variance-covariance matrix H_t.

More specifically, we define the standard multivariate GARCH framework as:

$$Y_t = \mu_t + \varepsilon_t \tag{2.4}$$

where μ_t[1] is the predictable conditional mean vector with respect to the information set ζ_{t-1}, and ε_t is the unpredictable error term, given the information set ζ_{t-1}, which we write as:

$$\varepsilon_t = H_t^{1/2}\eta_t \tag{2.5}$$

where H_t is the conditional variance-covariance matrix of Y_t which we write as:

$$H_t = \left[h_{ijt}\right] \tag{2.6}$$

where H_t is the $N \times N$ positive definite and symmetric matrix.

Also, we assume that η_t is an identically independent distributed (i.i.d.) random $N \times 1$ vector such that:

$$E(\eta_t) = 0 \text{ and } E(\eta_t\eta_t') = I_N \tag{2.7}$$

where I_N is an identity matrix of order *N*.

Now, we need to specify the conditional covariance matrix H_t, while noting that how we parameterize it will produce rather different results. Numerous attempts described in the literature have given rise to two general classes of models, namely, modeling conditional covariance matrix H_t directly (VEC[2] model and BEKK[3] model), and modeling conditional correlation matrix indirectly (constant conditional correlation (CCC[4]) model and dynamic conditional correlation (DCC) model). We also focus on the DCC MGARCH model because there are fewer parameters to be estimated and it is easier to use the numerical optimization for obtaining the convergence.

---

[1] In this dissertation, μ_t in equation (2.3) is the equivalent of term $(d + \sum_{i=1}^{p} \Gamma_i^* Y_{t-i})$ in equation (2.2).
[2] VEC-GARCH model is a generalization of the univariate GARCH model by Bollerslev et al. (1988).
[3] BEKK-GARCH model can be viewed as a restricted version of VEC-GARCH model which is the Baba-Engle-Kraft-Kroner (BEKK) defined in Engle and Kroner (1995).
[4] CCC-GARCH model is the simplest multivariate correlation model by Bollerslev (1990).

2.2.2.1. DCC-GARCH Model

Bollerslev (1990) developed the CCC model which we use to estimate the correlation of MGARCH models indirectly. However, the assumption that conditional correlations are constant over time is not realistic in practice. Subsequently, numerous econometricians and researchers have tried to generalize Bollerslev's CCC model. Engle (2002) proposed the DCC GARCH model and Engle and Sheppard (2001) extended it to accommodate large time varying covariance matrices. Our challenge is to transform the constant correlation matrix R to its time-varying counterpart R_t. We define Engle's dynamic correlation structure as:

$$H_t = D_t R_t D_t \tag{2.8}$$

where D_t is the conditional standard deviation matrix that can be expressed as $D_t = diag(\sigma_{1,t}, \sigma_{2,t}, \cdots, \sigma_{N,t})$ and R_t is the time-varying correlation matrix. We estimate $\sigma_{i,t}$ as:

$$\sigma_{it}^2 = w_{it} + \sum_{i=1}^{q} \alpha_{i,j} \varepsilon_{i,t-j}^2 + \sum_{j=1}^{p} \beta_{i,j} \sigma_{i,t-j}^2 \tag{2.9}$$

where $i, j = 1, 2, \cdots, N$, and $\sigma_{ij,t} = \rho_{ij} \sigma_{it} \sigma_{jt}$ for $i \neq j$.

We express the time-varying correlation matrix R_t as:

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1} \tag{2.10}$$

$$Q_t = (1 - \sum_{m=1}^{M} \lambda_{1m} - \sum_{n=1}^{N} \lambda_{2n}) \bar{Q} + \sum_{m=1}^{M} \lambda_{1m} (\eta_{t-m} \eta_{t-m}') + \sum_{n=1}^{N} \lambda_{2n} Q_{t-n} \tag{2.11}$$

where $\bar{Q} = E[\eta_t \eta_t']$, α_m and β_n are scalars such that $\sum_{m=1}^{M} \alpha_m + \sum_{n=1}^{N} \beta_n < 1$. Also, $\eta_t \sim U(0, R_t)$ is a $N \times 1$ vector of residuals standardized by their conditional standard deviation with typical element $\eta_{it} = \frac{\varepsilon_{it}}{\sqrt{h_{ii,t}}}$ which we obtain when estimating the univariate GARCH volatility models. We express Q_t^* as:

$$Q_t^* = \begin{bmatrix} \sqrt{q_{11,t}} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{q_{22,t}} & 0 & \cdots & 0 \\ 0 & 0 & \sqrt{q_{33,t}} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{q_{NN,t}} \end{bmatrix} \tag{2.12}$$

Thus, $Q_t^* = [q_{ii,t}^*] = [\sqrt{q_{ii,t}}]$ is a diagonal matrix with the square root of the i^{th} diagonal element of Q_t on its i^{th} diagonal position.

The typical element of R_t will be in a form such as $\rho_{it} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}}$. Engle and Sheppard (2001) established that the positive definiteness of Q_t will necessarily and sufficiently ensure the positive definiteness of R_t, which is validity condition of R_t as a correlation matrix. They used the unconditional variance-covariance matrix of the standardized residuals to replace the matrix $\bar{Q}$ when estimating the parameters, which is in line with standard univariate GARCH results. That is, simple variance-covariance matrix of $\hat{\bar{Q}} = \frac{\sum_{t=1}^{T} u_t u_t'}{T}$ serves as the estimator of $\bar{Q}$. This simplification invokes the concept of variance targeting introduced by Engle and Mezrich (1996), which assumes that in the long run the process of Q_t will approach the sample variance-covariance matrix $\hat{\bar{Q}}$. Even though variance targeting is achieved in this context, we cannot guarantee the positive definiteness of the variance-covariance matrix H_t. Hafner and Franses (2003) proposed a generalized DCC model to ensure the positive definiteness of the H_t matrix while sacrificing the variance targeting. Whether to choose variance targeting depends on the complexity of the model estimation. However, we do not expect major differences in these two categories. Similarly, we impose correlation targeting when necessary.

2.2.2.2. The Estimation Procedure of DCC-GARCH Model

Engle and Sheppard (2001) proved that the two-step estimator is consistent, and thus we apply it. We estimate the univariate GARCH model, and use the results as input to estimate the correlation parameters. One considerable assumption for this estimation method is the distribution of the standardized residuals. Assuming the unknown residual series η_t is the multivariate normal distribution, we apply the Maximum Likelihood Estimator (MLE) properties, and when the multivariate normality assumption does not hold, we apply the Quasi-Maximum Likelihood Estimator (QMLE) properties.

Let the standardized residual, η_t, assume multivariate Gaussian distributed. Then detail the joint distribution of $\eta_1, \eta_2, \cdots, \eta_T$ as:

$$f(\eta_t) = \prod_{t=1}^{T} \frac{1}{2\pi^{\frac{n}{2}}} exp\left\{-\frac{1}{2}\eta_t^T \eta_t\right\} \tag{2.13}$$

where $E(\eta_t) = 0$, $E(\eta_t \eta_t^T) = I$, and $t = 1, 2, \cdots, T$. Estimate the Maximum Likelihood function for $\varepsilon_t = H_t^{1/2}\eta_t$ by:

$$L(\theta) = \prod_{t=1}^{T} \frac{1}{2\pi^{\frac{n}{2}}|H_t|^{\frac{1}{2}}} exp\left\{-\frac{1}{2}\varepsilon_t^T H_t^{-1}\varepsilon_t\right\} \tag{2.14}$$

where θ is the model's parameters. Next, divide parameter θ into two groups:

$$(\phi, \psi) = (\phi_1, \phi_2, \cdots, \phi_n, \psi) \tag{2.15}$$

where $\phi_1 = (\alpha_{0i}, \alpha_{1i}, \cdots, \alpha_{qi}, \beta_{1i}, \cdots, \beta_{pi})$ are the parameters of the univariate GARCH model for the i^{th} price returns[5], and $\psi = (\lambda_1, \lambda_2)$ are the parameters of the correlation structure in equation (2.10).

Transforming the logarithm of equation (2.13) and substituting $H_t = D_t R_t D_t$ gives the log-likelihood function:

$$\begin{aligned} ln\big(L(\theta)\big) &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}(nln(2\pi) + ln(|H_t|) + \varepsilon_t^T H_t^{-1}\varepsilon_t) \\ &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}(nln(2\pi) + ln(|D_t R_t D_t|) + \varepsilon_t^T D_t^{-1} R_t^{-1} D_t^{-1}\varepsilon_t) \qquad (2.16) \\ &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}(nln(2\pi) + 2ln(|D_t|) + ln(|R_t|) + \varepsilon_t^T D_t^{-1} R_t^{-1} D_t^{-1}\varepsilon_t) \end{aligned}$$

The estimation of the correctly specified log-likelihood is difficult, but recall that the DCC-GARCH model is designed to allow for two-stage estimation. In the first stage, we estimate parameter ϕ of the univariate GARCH models for each price return. The likelihood used in the first step results in replacing R_t with the identity matrix I_N, which results in the quasi-likelihood function. Thus, we rewrite equation (2.15) as:

$$\begin{aligned} ln\big(L_1(\phi)\big) &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}(nln(2\pi) + 2ln(|D_t|) + ln(|R_t|) + \varepsilon_t^T D_t^{-1} R_t^{-1} D_t^{-1}\varepsilon_t) \\ &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}(nln(2\pi) + 2ln(|D_t|) + ln(|I_N|) + \varepsilon_t^T D_t^{-1} I_N D_t^{-1}\varepsilon_t) \\ &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}(nln(2\pi) + 2ln(|D_t|) + \varepsilon_t^T D_t^{-1} I_N D_t^{-1}\varepsilon_t) \qquad (2.17) \\ &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}\left(nln(2\pi) + \sum_{i=1}^{n}\left[ln(h_{it}) + \frac{\varepsilon_{it}^2}{h_{it}}\right]\right) \\ &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}\left(nln(2\pi) + \sum_{i=1}^{n}\left[ln(h_{it}) + \frac{\varepsilon_{it}^2}{h_{it}}\right]\right) \end{aligned}$$

[5] In this study, i is equal to four (i.e., crude oil, natural gas, and two electricity) market price returns.

Having estimated the parameter set $\phi = (\phi_1, \phi_2, \cdots, \phi_n)$, we can estimate the conditional variance h_{it} for each price return and also estimate $u_t = D_t^{-\frac{1}{2}}\varepsilon_t$ and $\bar{Q} = E[u_t, u_t^T]$.

After the first step, we cannot reveal the parameters λ_1 and λ_2. Thus, in the second step, we estimate parameter ψ using the correctly specified log-likelihood in equation (2.15), given parameter ϕ. Since D_t is constant when conditioning on the parameters from the first step, we exclude the constant terms and maximize equation (2.15) as:

$$\begin{aligned} ln\big(L_2(\psi)\big) &= -\frac{1}{2}\sum_{t=1}^{T}(nln(2\pi) + 2ln(|D_t|) + ln(|R_t|) + \varepsilon_t^T D_t^{-1} R_t^{-1} D_t^{-1} \varepsilon_t) \\ &= -\frac{1}{2}\sum_{t=1}^{T}(nln(2\pi) + 2ln(|D_t|) + ln(|R_t|) + \varepsilon_t^T R_t^{-1} \varepsilon_t) \qquad (2.18) \\ &= -\frac{1}{2}\sum_{t=1}^{T}(ln(|R_t|) + \varepsilon_t^T R_t^{-1} \varepsilon_t) \end{aligned}$$

2.2.3. Directed Acyclic Graphs (DAG)

This graphical approach is based on the graph theory that statistically inferred information about the probability distribution of the estimated residuals can be helpful in identifying the causal relationships among variables. Identification occurs by testing the conditional independence on the residuals (Bessler and Lee 2002; Bessler and Yang 2003; Demiralp and Hoover 2003; Moneta 2004, 2008; Swanson and Granger 1997).

The common approach is the directed acyclic graph (DAG) developed by Pearl (2000) and Spirtes et al. (2000), which shows the direction of information flows using directed edges among a set of variables. According to Pearl (1995), a DAG of causality has two parts: a certain number of nodes $(X_1, X_2, X_3, \cdots, X_n)$, and directed/undirected edges among nodes. Generally, a DAG exhibits acyclic patterns in a graph-like format. Each node X_i on the graph is expressed as a non-parametric structural equation $X_i = f_i(pa_i, \varepsilon_i)$, where pa_i are the parents of X_i on the graph and the ε_i are mutually independent. The non-parametric structural equations with $X_i = f_i(pa_i, \varepsilon_i)$, where X_i can be replaced by x_i, gives the distribution of the variables.

Mathematically, we write:

$$Pr(x_1, x_2, x_3, \cdots, x_n) = Pr \prod_{i=1}^{n} (x_i | pa_i) \tag{2.19}$$

where $Pr(\cdot)$ is the joint probability of variables $x_1, x_2, x_3, \cdots, x_n$, and pa_i are parent nodes (variables) of x_i meaning that pa_i links with x_i as a direct causal relation.

For example, consider the four variables x_1, x_2, x_3, and x_4 in figure 2.1, which "graphs" their causal relationships, i.e., x_2 causes x_3 and so on.

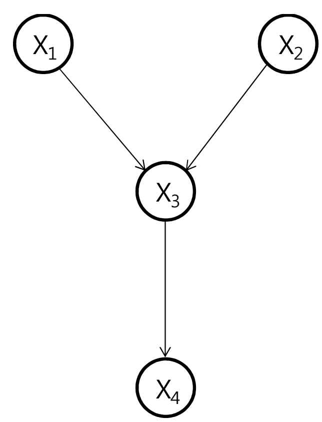


Figure 2.1. Example of a DAG and contemporaneous causal structures

We express the relationships as a functional form of the probability distribution product by:

$$Pr(x_1, x_2, x_3, x_4) = Pr(x_1)Pr(x_2)Pr(x_3|x_1, x_2)Pr(x_4|x_3) \tag{2.20}$$

The rule for interpretation tells us that variable x_1 and x_2 are independent if conditioning on some subset of $\{x_3,\ x_4\}$, since they are not connected, whereas variables x_3 and x_4 are dependent even if conditioning on any subset of $\{x_1,\ x_2\}$, since they are connected. Specifically, Pearl (1985) proposed the concept of d-separation as a graphical pattern of the conditional independence relations determined by a DAG. Under this simple concept, we say that two variables are d-separated when a third variable blocks the information flow between them. We can easily conceptualize d-separation by the three basic patterns of causal relationships: causal chains, causal forks and causal inverted forks. The pattern of causal chains represents the "$x_1 \rightarrow x_2 \rightarrow x_3$" of causal relationships among three variables, i.e., x_1 and x_3 are each dependent, but both are independent conditional on x_2. The pattern of causal forks represents the "$x_1 \leftarrow x_2 \rightarrow x_3$" of information flows among the three variables, i.e., x_1 and x_3 are each dependent, but both are independent conditional on x_2. The pattern of causal inverted forks represents the "$x_1 \rightarrow x_2 \leftarrow x_3$" relationship, i.e., x_1 and x_3 are each independent, but both are dependent conditional on x_2.

In practice, the goal of most graphical modes is to locate the most appropriate undirected/directed edges that represent the dependence structure from a given dataset. An edge between two nodes (variables) occurs if and only if the two corresponding nodes (variables) are dependent, even if conditioning on every subset of the remaining nodes (variables). A statistical test (Spirtes et al. 2000) analyzes dependency of two nodes (variables), given a set of other nodes (variables). Causal search algorithms find the direction of information flows by using a statistical measure of independence, conditional correlation, checking the systematical patterns of conditional independence and dependence, and then working backward to the allowed causal relationships (Hoover 2005).

The PC algorithm[6], the most widely used in the literature (Hoover 2005; Kim and Bessler 2007), is the greedy or structurally restricted approach introduced by Bernanke et al. (2005), who incorporated d-separation into the algorithm. The PC algorithm starts from connecting complete undirected edges for all variables in the graph and then it recursively deletes edges between variables based on conditional independence (zero correlation or partial correlation) decisions using Fisher's Z[7] statistics. The output of the PC algorithm is the pattern of causal flows containing both undirected and directed edges. The undirected edge indicates the ambiguous direction of arrow. However, in discovering that the PC algorithm frequently omits edges when sample size are small (less than 200 observations), Spirtes et al. (2000) suggested that the significance level used for Fisher's Z test should increase as the sample size decreases. They recommended 20% significance level for less than 100 observations and a 10%

[6] "PC" stands for Peter Spirtes and Clark Glymour who invented the algorithm in 1991.

[7] The Fisher's Z statistic is $Z[\rho(i,j|k),n] = \left(\frac{1}{2}\sqrt{n-|k|-3}\right) \times ln\left(\frac{|1+\rho(i,j|k)|}{|1-\rho(i,j|k)|}\right)$, where $\rho(i,j|k)$ is the sample correlation between i and j conditional on k, $|k|$ is the number of conditional variables in k, and n is the number of observations used to estimate the correlation. The null hypothesis is that conditional correlation is equal to zero.

significance level for sample sizes between 100 and 300. Therefore, we use a 10% significance level for the PC algorithm.

We begin by generating the DAGs of within-sample-fit and out-of-sample-forecast using the PC algorithm in the software project TETRAD IV, which represent the direction of the contemporaneous causal structure among the price returns of US energy markets. The next section describes our dataset.

2.3. Data Description

Our dataset consists of the daily spot prices from the Bloomberg database from May 3, 2004 to December 30, 2011, excluding all public holidays for all markets simultaneously. We select this time period because it represents a continuous series of data from the newest observation up to 2000th observation in total for each price series. The crude oil and natural gas prices are the daily spot prices of Dated Brent crude oil and Henry Hub, respectively, and the electricity prices are the firm peak daily spot prices of the PJM (Pennsylvania-New Jersey-Maryland) and COB (California-Oregon) electricity markets. Table 2.1 gives the summary statistics and figure 2.2 gives the plots of the price series. As shown in figure 2.2, each price series exhibits high volatility and potential heteroscedasticity. To account for these two issues, we use log-transformed data for all estimations by using a robust estimator. The robust estimator computes a heteroscedasticity consistent estimate of the asymptotic covariance matrix of the estimated parameters (Greene 2007).

We denote the first difference of log transformed price series as a measure of the price returns of Dated Brent crude oil, Henry Hub natural gas, PJM Electricity firm on peak and COB Electricity firm on peak as R_{OIL}, R_{GAS}, R_{PJM}, and R_{COB}. Table 2.2 gives the summary statistics and figure 2.3 gives the time series plots of the daily returns. We observe that it is easy to find volatility clustering in energy price return series.

We start by analyzing the dynamic behavior of each univariate series, which serves to facilitate the multivariate modeling and the understanding of multivariate dynamics. In table 2.2, we note that the Dated Brent crude oil market experiences positive mean returns unlike the other energy markets. Based on the magnitude of the unconditional standard deviations, the PJM market is more volatile. Both Henry Hub and COB generate positive skewness and very high kurtosis, whereas both Dated Brent and PJM exhibit negative skewness and relatively small kurtosis.

We test for the presence of a unit root for the log transformed prices and price returns of each market. Since a series with a unit root is non-stationary with an infinite unconditional variance, it is not possible to generalize it to other time periods. Table 2.3 shows the Dickey-Fuller test and augmented Dickey-Fuller test statistics; the log level prices of both Dated Brent and Henry Hub fail to reject the null hypothesis of a unit root at the 10% significant level, whereas PJM and COB suggest that both series are stationary in log levels. However, all price returns, i.e., first differencing of the logarithm of the price series, result in rejecting the null hypothesis at the 1% significance level, indicating stationary.

Table 2.1
Summary statistics on the logarithms of daily energy prices

Energy price series	Dated Brent Crude Oil ($/BBL)	Henry Hub Natural Gas ($/MMBTU)	PJM Electricity Firm on Peak ($/MWh)	COB Electricity Firm on Peak ($/MWh)
Mean	1.85	0.77	1.76	1.70
Standard Deviation	0.143	0.159	0.160	0.160
Variance	0.021	0.025	0.026	0.026
Minimum	1.51	0.26	1.36	0.98
Maximum	2.16	1.19	2.48	2.45
Skewness	-0.028	0.182	0.519	-0.123
Kurtosis	2.303	2.685	3.160	3.525

Table 2.2
Summary statistics on the daily returns of energy prices

Returns of energy prices	Dated Brent Crude Oil ($/BBL)	Henry Hub Natural Gas ($/MMBTU)	PJM Electricity Firm on Peak ($/MWh)	COB Electricity Firm on Peak ($/MWh)
Mean	0.00024	-0.00014	-0.00005	-0.00008
Standard Deviation	0.00980	0.01853	0.11302	0.04663
Variance	0.00010	0.00034	0.01277	0.00217
Minimum	-0.05345	-0.11087	-0.47807	-0.55027
Maximum	0.05944	0.13011	0.44771	0.52189
Skewness	-0.00619	0.51102	-0.04639	0.36034
Kurtosis	6.07486	10.34786	4.45832	30.86017

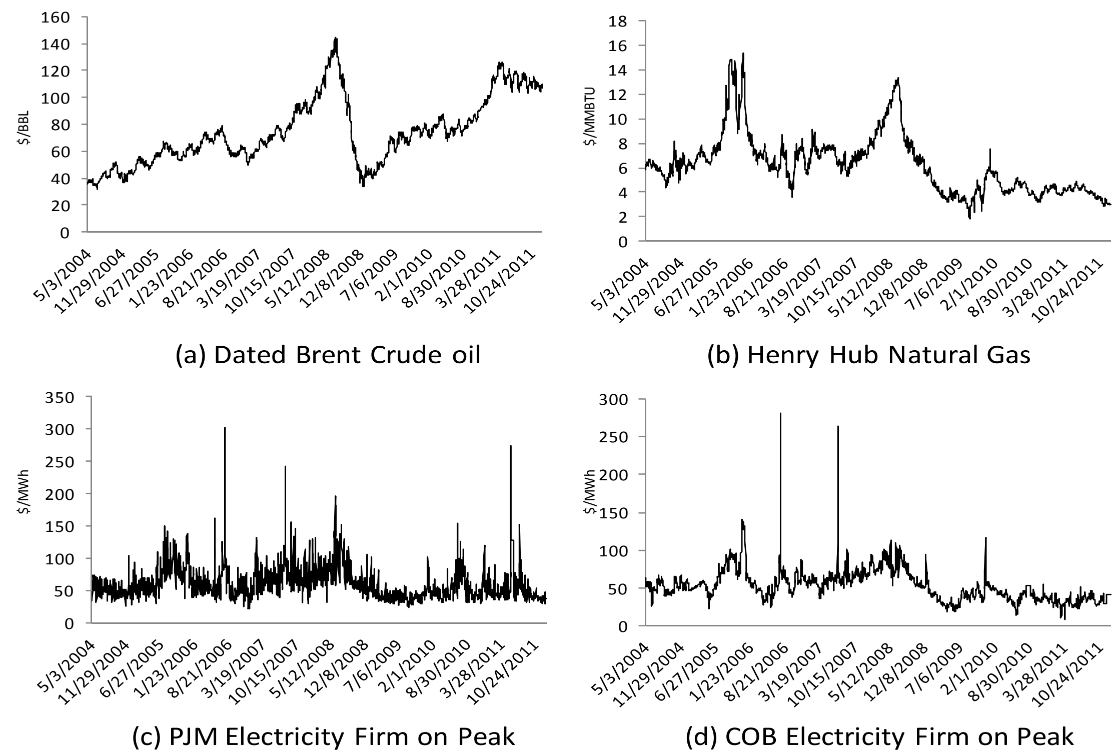


Figure 2.2. Plots of the price series

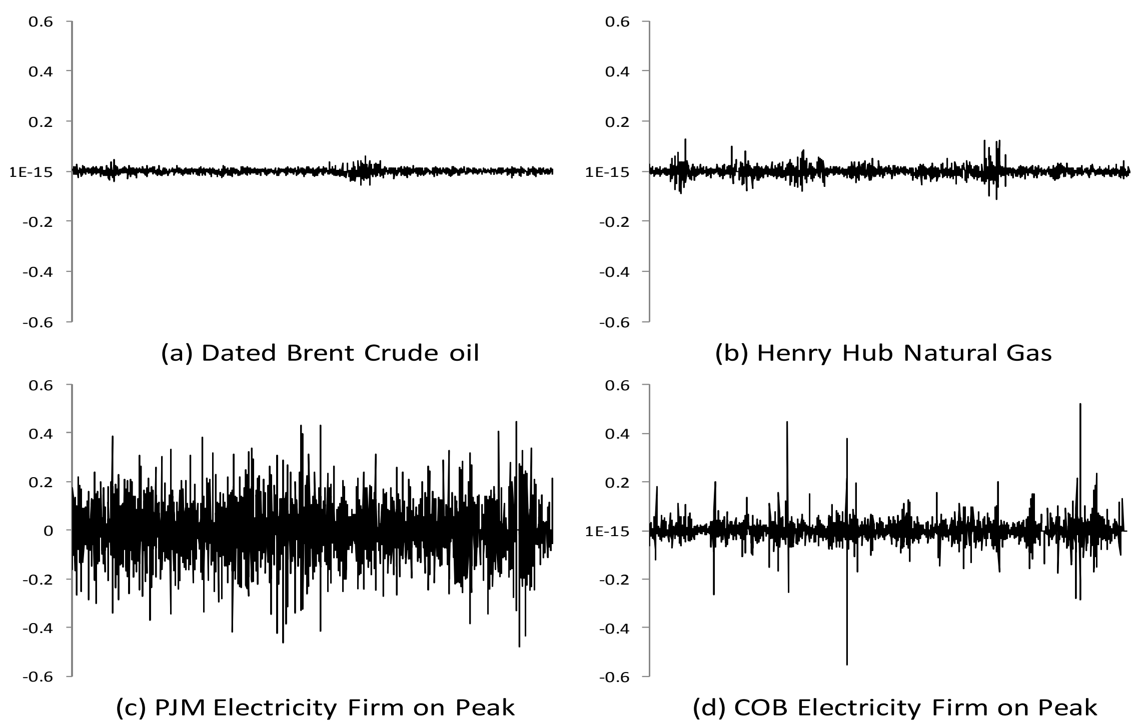


Figure 2.3. Plots of the daily price returns

Table 2.3
Tests for non-stationary of energy price series

Energy price series	DF Test		ADF Test (k)[a]	
	Log Level	First Difference	Log Level	First Difference
Dated Brent Crude Oil	-1.96	-44.36*	-1.89	-25.68*
Henry Hub Natural Gas	-2.29	-43.06*	-2.10	-28.91*
PJM Firm on Peak Electricity	-16.86*	-59.91*	-10.93*	-37.31*
COB Firm on Peak Electricity	-6.58*	-45.76*	-5.71*	-31.79*

Note: * indicates 1% significance level; the critical value is -3.51 at the 1% significance level; [a] indicates the number of lag determined by optimal lag order selection criteria.

2.4. Empirical Results

Here, we present the empirical results from our DCC-GARCH model fitted to the data. We assume that our GARCH (1, 1) model is parsimonious, i.e., we use a DCC specification of MGARCH (1, 1) which allows for dynamic conditional correlations among the price returns. Then we implement standardized innovation analysis.

2.4.1. Vector Autoregression Results

We conduct a preliminary data analysis by using the maximum likelihood estimation procedure of Johansen (1991) to construct a VAR (p) process. We determine the optimal lag-length based on loss information criteria, i.e., Akaike, Schwarz, and Hannan and Quinn losses. Table 2.4 shows the somewhat ambiguous results: SIC suggests $p = 2$, HQIC suggests $p = 4$, and AIC suggests $p = 6$ as an optimal lag order. For specification of the VAR process, AIC is appropriate information criterion for monthly data, HQIC is appropriate criterion for quarterly data and SIC is universally applicable (Haan and Levin 2000). Following SIC, we select $p = 2$ as an optimal lag order since we have daily data with 2000 observations.

Having chosen the most parsimonious specification, we proceed to fit the VAR (2) model to the four-variate log price returns of the time series. Table 2.5 gives the estimated parameters and robust standard errors.

Table 2.4
VAR optimal lag-length determination

Lag Order	Akaike Information Criterion (AIC)	Schwarz Information Criterion (SIC)	Hannan and Quinn Information Criterion (HQIC)
0	-10.5720	-10.5606	-10.5678
1	-10.5971	-10.5399	-10.5761
2	-10.7084	-10.6055*	-10.6706
3	-10.7539	-10.6053	-10.6993
4	-10.7714	-10.5770	-10.6999*
5	-10.7786	-10.5386	-10.6904
6	-10.8042*	-10.5184	-10.6991

Note: * indicates the most appropriate lag order for the model; information criteria used to identify the optimal lag-length (p) of a VAR process are $AIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{2n}{T}\right)$, $SIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{nlnT}{T}\right)$, and $HQIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{2nln(lnT)}{T}\right)$, where $\widehat{\Omega}_p$ is the maximum likelihood estimate of variance-covariance matrix of Ω, p is the proposed lag-length, n is the number of variables and T is the sample size.

Table 2.5
VAR (2) model estimation results for energy price returns

Parameters	$R_{OIL,}$ $(i = 1)$		$R_{GAS,}$ $(i = 2)$		$R_{PJM,}$ $(i = 3)$		$R_{COB,}$ $(i = 4)$	
	Coefficient	Std. Err	Coefficient	Std. Err	Coefficient	Std. Err	Coefficient	Std. Err
d_i	0.00079	0.00055	-0.00012	0.00107	0.00042	0.00338	-0.00052	0.00255
$\gamma_{1,i1}$	-0.03738*	0.02369	0.19209***	0.04382	0.12342	0.14418	0.16236	0.10906
$\gamma_{1,i2}$	0.02836**	0.0123	-0.05627**	0.02467	0.21873***	0.07743	-0.11481**	0.05857
$\gamma_{1,i3}$	-0.00241	0.00384	0.02220***	0.00744	0.00090	0.02335	-0.01352	0.01766
$\gamma_{1,i4}$	-0.00958*	0.00517	0.03374***	0.01002	-0.06775**	0.03145	-0.03006	0.02379
$\gamma_{2,i1}$	0.01626	0.02375	-0.01648	0.04394	0.19740	0.1446	0.27237**	0.10934
$\gamma_{2,i2}$	-0.00317	0.01263	-0.10234***	0.0245	0.03110	0.07688	0.02403	0.05815
$\gamma_{2,i3}$	-0.00760**	0.00384	-0.0003	0.00744	-0.24883***	0.02335	0.01007	0.01766
$\gamma_{2,i4}$	-0.00239	0.00519	-0.00902	0.01006	-0.06703**	0.03158	-0.18049***	0.02390

Diagnostics tests	R_{OIL}	R_{GAS}	R_{PJM}	R_{COB}
R^2	0.007	0.034	0.068	0.041
RMSE	0.024	0.046	0.143	0.109
χ^2	13.364*	63.509***	133.165***	76.864***
Log-likelihood	9774.98			
# of observations	1800			

Note: *, **, and ***, indicate 10%, 5%, and 1% significance levels, respectively.

The VAR (2) model has 36 parameters, of which 17 are significant. Note that both Henry Hub and COB price returns are negatively impacted by their own lags, whereas both Dated Brent and PJM price returns are positively impacted by their own lags at the 5% significance level.

In the next step we use the Ljung-Box Q to determine whether the residuals of the VAR (2) model are white noise by investigating the autocorrelation and the square of each residual for the 8, 16, and 24 lags, respectively. Table 2.6 reveals that all residuals and squared residuals[8] are highly significant (all have p-values of less than .01) with the exception of the residuals in Dated Brent at lags 8 and 12. The Ljung-Box Q statistics suggest a strong conditional heteroscedasticity in all residuals in energy price returns, i.e., once volatility increases it tends to persist for a certain period of time. Also, we use the Breusch-Pagan and White test statistics for detecting heteroscedasticity. Test results indicate the residuals of each series from the VAR (2) model has heteroscedasticity problems. Therefore, the GARCH process is a plausible candidate for modeling their time series behaviors.

Next, we fit a suitable MGARCH model to the residuals (ε_{1t}, ε_{2t}, ε_{3t}, ε_{4t}) of the VAR (2) model for the price returns of Dated Brent crude oil, Henry Hub natural gas, PJM Electricity firm on peak and COB Electricity firm on peak.

---

[8] The squared price returns can be viewed as a proxy for the variance of the series.

Table 2.6
Ljung-Box Q and Heteroscedasticity test statistics for residuals in VAR (2) model for the price returns

Residuals of energy price returns in VAR (2) model	VAR			
	R_{VAR}_R_{OIL}	R_{VAR}_R_{GAS}	R_{VAR}_R_{PJM}	R_{VAR}_R_{COB}
JB	1.638 (0.441)	64.805 (0.000)	7.405 (0.025)	961.119 (0.000)
Q(8)	8.718 (0.367)	81.45 (0.000)	195.7 (0.000)	88.07 (0.000)
Q(12)	13.02 (0.368)	119.7 (0.000)	203.5 (0.000)	106.1 (0.000)
Q(24)	38.55 (0.030)	186.2 (0.000)	237.3 (0.000)	145.6 (0.000)
$Q^2(8)$	404 (0.000)	343.7 (0.000)	65.05 (0.000)	193.8 (0.000)
$Q^2(12)$	986.6 (0.000)	463.8 (0.000)	71.14 (0.000)	194.1 (0.000)
$Q^2(24)$	1461.3 (0.000)	637.2 (0.000)	111.4 (0.000)	219.1 (0.000)
Heteroscedasticity Test				
Breusch-Pagan	0.0224	0.1514	0.0089	0.0202
White	0.0003	0.1604	0.0152	0.0664

Note: $Q(k)$ and $Q^2(k)$ are the Ljung-Box Q test statistics for serial correlation of k lags of the original and squared price returns, respectively; under the null hypothesis of no serial correlation, the Q-statistics follow the chi-squared distribution with k degrees of freedom; p-values are in parentheses.

2.4.2. DCC-GARCH Results

For the DCC-GARCH model, we fit equations (2.9) ~ (2.11) to the residuals of the VAR (2) model. Recall that we have chosen to adopt the most parsimonious specification with $m = 1$ and $n = 1$. Table 2.7 reports the DCC-GARCH (1, 1) estimates. Fourteen parameters are significant at the 1% significance level. We can clearly interpret the model's correlation structure, i.e., there is a non-constant interaction of the four time series with respect to conditional correlation, and this correlation impacts the current correlation with a lag of one. We note that this interaction effect would be neglected if the four time series of VAR residuals were each modeled with a univariate GARCH model in isolation.

In table 2.8, which shows the diagnostic test statistics of the fitted model, we note that the range of residuals is now closer to what we expect from a standard normal distribution; the standardized residuals of each series from the DCC-GARCH (1, 1) model do not exhibit autocorrelation, since the squared residuals remain in the range of the critical values. The Breusch-Pagan and White test statistics also indicate that the standardized residuals of each series from the DCC-GARCH (1, 1) model do not show a heteroscedasticity problem. Therefore, we conclude that the residuals of the DCC-GARCH (1, 1) model satisfy the necessary white noise properties.

Table 2.7
DCC-GARCH (1, 1) model estimation results for energy price returns

Parameters		Estimated Coefficients	
		Coefficient	Std. Err
R_{OIL}	w_t	0.00001***	0.00000
	α_t	0.05723***	0.00013
	β_t	0.91920***	0.00028
R_{GAS}	w_t	0.00003***	0.00000
	α_t	0.11349***	0.00027
	β_t	0.87647***	0.00025
R_{PJM}	w_t	0.00087***	0.00000
	α_t	0.15652***	0.00088
	β_t	0.81457***	0.00088
R_{COB}	w_t	0.00071***	0.00000
	α_t	0.34329***	0.00517
	β_t	0.65671***	0.00869
Correlation parameters	λ_1	0.01334***	0.00001
	λ_2	0.95989***	0.00017
# of observations	2000		

Note: *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.

Table 2.8
Diagnostic test statistics for the standardized residuals in DCC-GARCH (1, 1) model

Residuals of energy price returns	DCC-GARCH			
	R_{VDG}_R_{OIL}	R_{VDG}_R_{GAS}	R_{VDG}_R_{PJM}	R_{VDG}_R_{COB}
JB	24.79917 (0.000)	0.516821 (0.772)	5.97415 (0.050)	75.76827 (0.000)
Q_{12}	11.41920 (0.493)	5.02030 (0.957)	74.97320 (0.000)	36.22970 (0.000)
Q^2_{12}	4.67810 (0.968)	11.4388 (0.492)	10.5938 (0.056)	6.3199 (0.891)
Heteroscedasticity Test				
Breusch-Pagan	0.8248	0.2602	0.1788	0.4135
White	0.9756	0.3441	0.3744	0.6904

Note: Residuals from the DCC-GARCH model are the standardized residuals, $R_{VDG}\_R = \varepsilon/h^{0.5}$; p-values are in parentheses; the lags of Ljung-Box Q statistics over standardized residuals and squared residuals are in subscript.

Figure 2.4 graphs the time-varying variance-covariance matrix between the energy price returns of the four commodities estimated from the DCC-GARCH (1, 1) model.

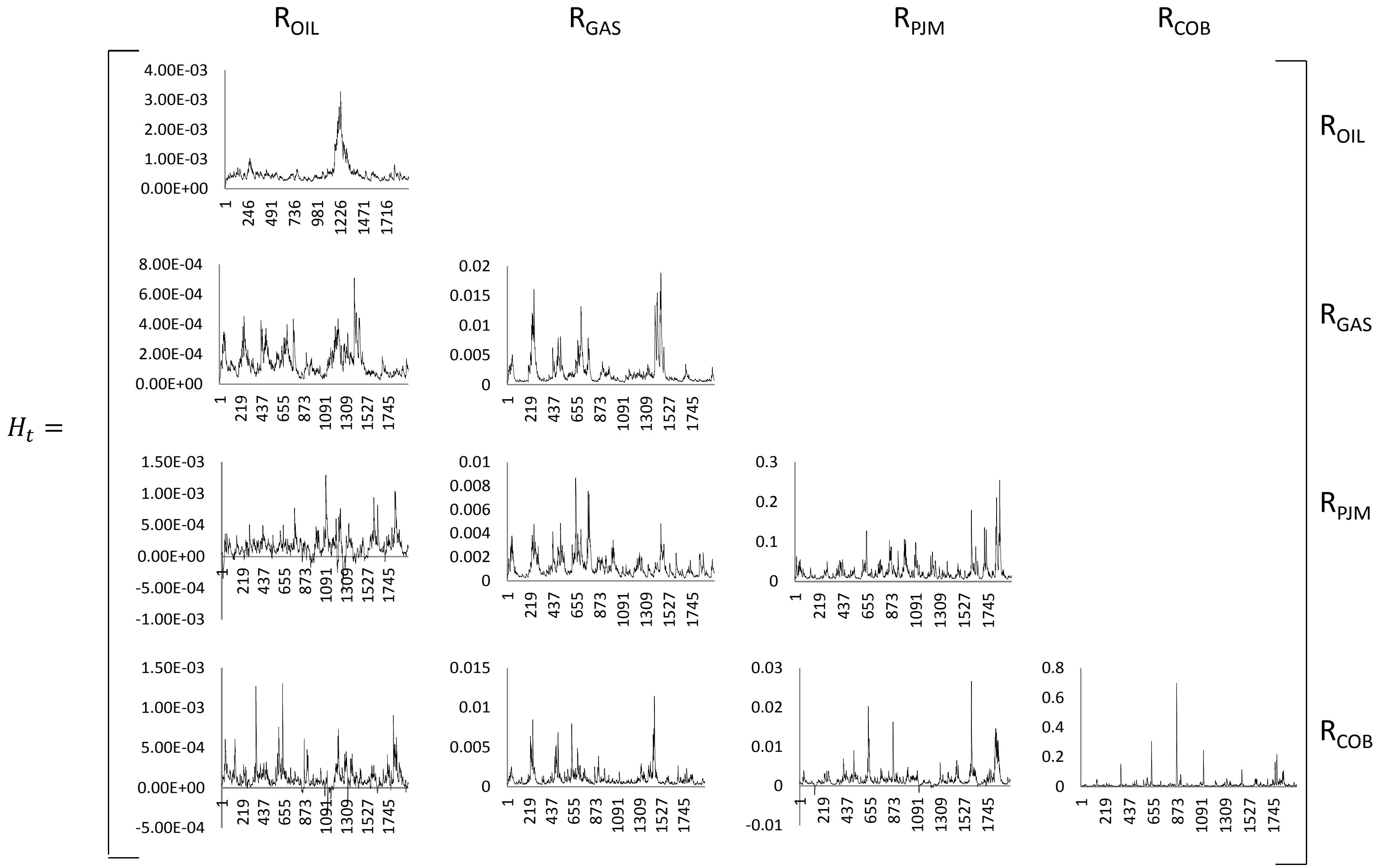


Figure 2.4. Time-varying variance-covariance matrix from DCC-GARCH (1, 1) model

2.4.3. Forecast and Forecasting Performance Evaluation

Based on the estimated VAR-DCC-GARCH model, we adopt two types of forecasting procedures. The first type is within-sample-fit. To perform this procedure, we use the entire dataset of 2000 trading days. The second type is out-of-sample-forecast. To perform this procedure, we use the last 200 observations (March 1, 2011 to December 31, 2011). The out-of-sample forecasting procedure is as follows. We estimate the models 200 times based on last 200 samples of 2000 observations. We use November 10, 2003 to February 28, 2011 to forecast the covariance matrix of March 1, 2011 based on the estimated model for the first sample. We use November 10, 2003 to March 1, 2011 to forecast the covariance matrix for March 2, 2011 based on the estimated model for the second sample. These estimation and forecasting steps can be repeated 200 times for the available sample and we produce the 200 one-step-ahead covariance matrix forecasts and residuals. To compare results, we only use 200 residuals from March 1, 2011 to December 31, 2011 as residuals of within-sample-fit.

We calculate the out-of-sample-forecast residuals as the difference between the forecasted and the actual price returns of the four commodities. Based upon these exercise results, we have two forecast performance measure statistics, Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE), as summarized in table 2.9.

Table 2.9
Summary of forecast performance of within-sample-fit and out-of-sample-forecast

	MSE		MAPE	
	Within-sample	Out-of-sample	Within-sample	Out-of-sample
R_{OIL}	0.959987	0.960569	72.36324	72.03272
R_{GAS}	0.788752	0.79122	68.55013	66.9236
R_{PJM}	0.974272	1.038579	75.66105	74.30997
R_{COB}	1.244264	1.253367	80.05075	80.87663

Note: $\text{MSE} = \frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \hat{Y}_t\right)^2$; $\text{MAPE} = \frac{100}{T}\sum_{t=1}^{T}\left|Y_t - \hat{Y}_t\right|$; a lower loss measure indicates a higher forecasting power.

The MSEs of the within-sample-fit are 0.960, 0.789, 0.974 and 1.244, respectively for each variable. We observe that they are all less than the MSEs of the out-of-sample-forecast. The MAPEs of the within-sample-fit are 72.36, 68.55, 75.66 and 80.05, respectively for each variable. We observe that they are greater than the MAPEs of the out-of-sample-forecast with the exception of the price return for COB. We conclude that the forecast performance measure statistics are poorer than we had expected. In other words, the MSEs of within-sample-fit for all variables generally exhibit smaller values compared to the out-of-sample-forecast, yet fail to show superiority in MAPE cases. This is plausible since our model is essentially based on generalized least squares estimators that minimize the sum of squared residuals. Therefore, the MSEs of within-sample-fit have to be smaller. However, this stylized fact does not necessarily hold for MAPE, unless we use minimum absolute deviation estimations.

We also note that the two performance measures do not provide a statistical significance of the similarity/difference between the within-sample and out-of-sample-forecast. Therefore, we use the DM test (Diebold and Mariano 1995) and the forecast encompassing test (Harvey et al. 1997) to further examine the out-of-sample predictability. Table 2.10 summarizes the test statistics.

From the DM test statistics, we cannot reject the hypothesis of equality of forecast errors between within-sample-fit and out-of-sample-forecast for the price returns of Dated Brent, Henry Hub, and COB at any significance level, whereas the price returns of PJM can reject the null hypothesis at the 5% significance level. Therefore, we say that the within-sample-fit and out-of-sample-forecast for the price returns of Dated Brent, Henry Hub, and COB statistically perform similarly. However, the *p*-value for PJM indicating that the hypothesis of equality of forecast errors between within-sample-fit and out-of-sample-forecast cannot be rejected at the 10% significance level implies that the within-sample-fit and out-of-sample-forecast statistically perform similarly.

Table 2.10
Summary of test statistics of DM and forecast encompassing tests for within-sample-fit and out-of-sample-forecast

	DM		Forecast Encompassing			
			Dependent variable			
			$e_{forecast,t}$		$e_{fit,t}$	
	test statistics	*p*-value	$\hat{\lambda}$	*p*-value	$\hat{\lambda}$	*p*-value
R_{OIL}	0.0386	0.9692	0.523	0.383	0.477	0.426
R_{GAS}	0.0526	0.9581	0.511	0.011	0.490	0.015
R_{PJM}	1.7332	0.0831	0.913	0.000	0.087	0.719
R_{COB}	0.3560	0.7218	0.641	0.106	0.359	0.364

Note: Both tests are based on the null hypothesis of no difference in the accuracy (equal predictive ability) between within-sample-fit and out-of-sample-forecast.

In the DM test, the null hypothesis of equal forecast accuracy is tested based on $E(d_t) = 0$, where E is expectation operator and $d_t = e^2_{fit,t} - e^2_{forecast,t}$; variables $e_{fit,t}$ and $e_{forecast,t}$ are forecast errors generated by within-sample-fit and out-of-sample-forecast, respectively. The DM test statistic is $DM = [\hat{V}(\bar{d})]^{-1/2}\bar{d}$, where $\bar{d}$ is the sample mean of d_t, $\hat{V}(\bar{d})$ is the sample variance of $\bar{d}$ asymptotically estimated by $T^{-1}\left[\gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k\right]$, and $\gamma_k$is the kth autocovariance of d_t estimated from $T^{-1}\sum_{t=k+1}^{T}\left(d_t - \bar{d}\right)\left(d_{t-k} - \bar{d}\right)$. Under the null hypothesis, these statistics follow an asymptotic standard normal distribution.

In the forecast encompassing test, the test determines weights based on the covariance between the errors from out-of-sample-forecast, $e_{forecast,t}$, and the difference between the errors of the out-of-sample-forecast and within-sample-fit, $e_{forecast,t} - e_{fit,t}$. If the covariance is not equal to zero, then information can be gained and a composite forecast can be built. This is tested based on $e_{forecast,t} = \lambda\left(e_{forecast,t} - e_{fit,t}\right) + \varepsilon_t$, where ε_t is a composite forecast error. The null hypothesis is $\lambda = 0$. If the null is true, then the out-of-sample-forecast encompasses the within-sample-fit. The actual test involves an OLS regression of $e_{forecast,t}$ on $\left(e_{forecast,t} - e_{fit,t}\right)$, but we use *t*-test of $\hat{\lambda}$ for forecast encompassing.

For the forecast encompassing test, the hypothesis is that if $\hat{\lambda}$ is significantly different from zero when the dependent variable is the residuals of the out-of-sample-forecast/within-sample-fit, then the within-sample-fit/out-of-sample-forecast encompasses the out-of-sample-forecast/within-sample-fit. From the results, the within-sample-fit/out-of-sample-forecast does not encompass the out-of-sample-forecast/within-sample-fit at the 10% significance level in the price returns of Dated Brent (*p*-values are 0.383 and 0.426) and COB (*p*-values are 0.106 and 0.364) cases, nor does the within-sample-fit/out-of-sample-forecast encompass the out-of-sample-forecast/within-sample-fit at the 1% significance level in the price returns of Henry Hub (*p*-values are 0.011 and 0.015, respectively). However, the within-sample-fit encompasses the out-of-sample-forecast (*p*-value is 0.000), whereas the out-of-sample-forecast does not encompass the within-sample-fit (*p*-value is 0.719) for PJM. We conclude that the within-sample-fit is superior in forecasting ability. In general, these forecast encompassing test results are consistent with the findings of the equality tests.

2.4.4. Variance-covariance matrices of residuals from within-sample-fit and out-of-sample- forecast

It is well known that unconditional covariance and correlation coefficients vary significantly over time and that MGARCH can capture the dynamic conditional (time-varying) variance-covariance matrices (Engle and Sheppard 2001; Tse and Tsui 2002). We can also capture and explore the dynamic interactions among the four price returns with the DCC-GARCH framework. However, this means that the DAG patterns between variables will be sensitive to the time and sign of the change in the price returns resulting from information flows. Moreover, we can identify the contemporaneous causal structures from the variance-covariance matrices of innovations which are generated from the estimated time series model. These structures are based on the results of the DAGs (Bessler and Lee 2002; Bessler and Yang 2003; Pearl 2000; Spirtes et al. 2000; Swanson and Granger 1997). Following them, we generate two constant variance-covariance matrices from the standardized residuals of within-sample-fit and out-of-sample-forecast for March 1, 2011 to December 31, 2011 to obtain the robust DAG patterns for the four price returns. Doing so allows us to fulfill our objective of comparing the DAG patterns between within-sample-fit and out-of-sample-forecast. In addition, we can test the equality of the variance-covariance matrices between the standardized residuals from within-sample-fit and out-of-sample-forecast from the obtained variance-covariance matrices. Figure 2.5 details the generated variance-covariance matrices.

From the above two variance-covariance matrices, we observe that all elements of variance-covariance matrix ($\Sigma_{within-sample-fit}$) for the standardized residuals from within-sample-fit are less than the elements of the variance-covariance matrix ($\Sigma_{out-of-sample-forecast}$) for the standardized residuals from the out-of-sample-forecast.

$$
\boldsymbol{\Sigma}_{\textit{within-sample-fit}} = \begin{array}{c}
\begin{array}{cccc} R_{OIL} & R_{GAS} & R_{PJM} & R_{COB} \end{array} \\
\begin{bmatrix} 0.9553 & & & \\ 0.0994 & 0.7877 & & \\ 0.0849 & 0.1025 & 0.9735 & \\ 0.0309 & 0.0465 & 0.1382 & 1.2418 \end{bmatrix}
\end{array}
\begin{array}{c} \\ R_{OIL} \\ R_{GAS} \\ R_{PJM} \\ R_{COB} \end{array}
$$

$$
\boldsymbol{\Sigma}_{\textit{out-of-sample-forecast}} = \begin{array}{c}
\begin{array}{cccc} R_{OIL} & R_{GAS} & R_{PJM} & R_{COB} \end{array} \\
\begin{bmatrix} 0.9557 & & & \\ 0.0994 & 0.7903 & & \\ 0.0916 & 0.1061 & 1.0375 & \\ 0.0393 & 0.0594 & 0.1630 & 1.2523 \end{bmatrix}
\end{array}
\begin{array}{c} \\ R_{OIL} \\ R_{GAS} \\ R_{PJM} \\ R_{COB} \end{array}
$$

Figure 2.5. The variance-covariance matrices of the standardized residuals from the within-sample-fit and out-of-sample-forecast

We use the Box M statistic derived by Box (1949) to test for homogeneity based on the likelihood ratio test. The null hypothesis is given by:

$$H_0: \Sigma_{within-sample-fit} = \Sigma_{out-of-sample-forecast} \tag{2.22}$$

Table 2.11 summarizes the test statistics for the within-sample-fit and out-of-sample-forecast. From the results, we find that the test statistic value is 0.2370. This suggests that we fail to reject the null hypothesis at the 1% significance level (*p*-value is 1.000), since the critical value is 23.21. We conclude that the two covariance matrices from the within-sample-fit and out-of-sample-forecast are the same.

Table 2.11
Summary of test statistics of the Box M test for within-sample-fit and out-of-sample-forecast

	Box M test statistics	Critical Value at 1% significance level	p-value
Box M test	0.2370	23.21	1.000

Note: We adopt the Box M test when the sample size is small following Mardia and Kent (1979). The Box M test statistic is generated by $M = \gamma \sum_{i=1}^{g} (n_i - 1) log \left| S_u S_{u_i}^{-1} \right|$, where $\gamma = 1 - \frac{2k^2+3k-1}{6(k+1)(g-1)} \left(\sum_{i=1}^{g} \frac{1}{n_i - 1} - \frac{1}{n-g} \right)$, $S_u = \frac{n}{n-g} S$, $S_{u_i} = \frac{n_i}{n_i - g} S_i$, $S = \sum_{i=1}^{g} \frac{n_i S_i}{n}$ is the pooled covariance matrix, g is the number of groups with non-singular covariance matrices, $n = n_1 + n_2 + \cdots + n_g$ is the number of the total sample size, n_i is number of the sample size for deriving sample covariance matrix S_i, k is the dimension of the covariance matrix, and $i = 1, 2, \cdots, g$. The Box-M test statistic is asymptotically distributed as a chi-square distribution with the degree of freedom, $k(k+1)/2$.

2.4.5. Comparing DAGs between residuals from within-sample-fit and out-of-sample- forecast

From the standardized residuals of the within-sample-fit and out-of-sample-forecast, we obtain two identical DAGs from TETRAD IV's PC algorithms at the 10% significance level. Figure 2.6 shows the DAG patterns.

Figure 2.6 (a) and (b) indicate no strong contemporaneous causal relationships among the four price returns. However, we find one (undirected) edge between the price returns of PJM and COB in the DAGS, whereas the price returns of Dated Brent and Henry Hub are revealed as independent.[9]

[9] These contemporaneous causal structures are significantly different from the results of Mjelde and Bessler (2009). Possible reasons are that 1) these graphs are generated by using only recent short time periods (March 1, 2011 to December 31, 2011, 200 observations); and 2) Mjelde and Bessler (2009) considered weather effects by using weekly price data, whereas we do not include weather effects by using daily data. We also find strong information flows between electricity prices and gas as well as between gas and oil by using long historical data (November 2003 to December 2011). However, the difference between causal structures of U.S energy market is not a controversial issue for this dissertation, since this

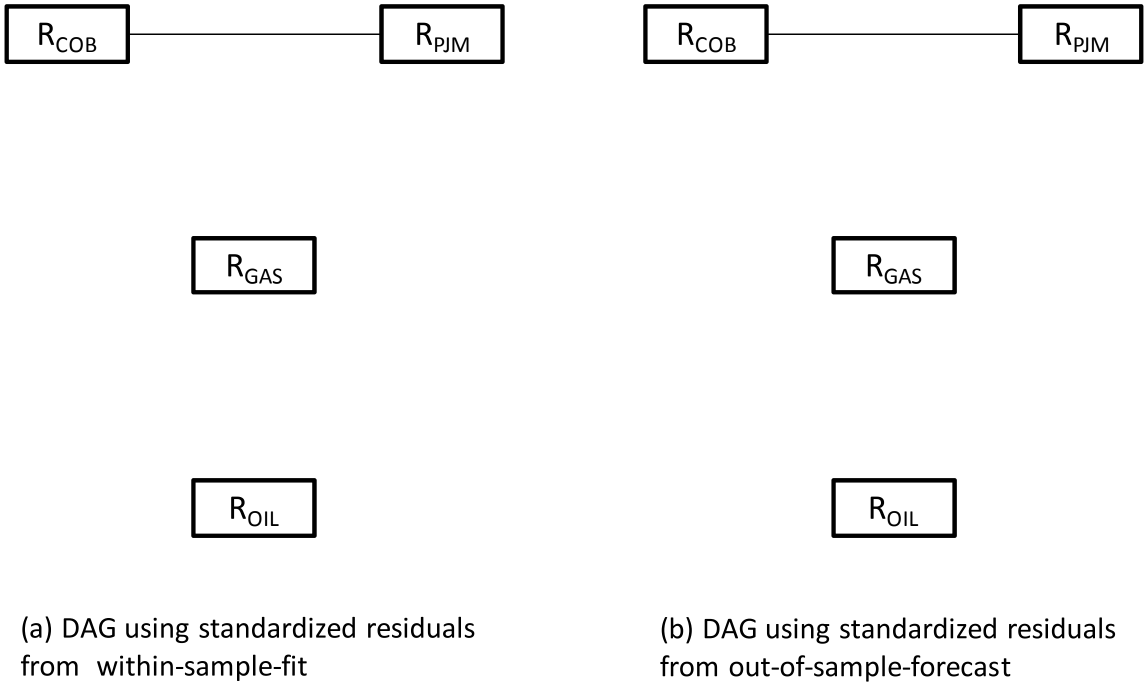


Figure 2.6. DAG patterns from the standardized residuals in the within-sample-fit and out-of-sample-forecast

The undirected edge implies that the algorithm cannot completely determine the direction of information flows from the 200 standardized residuals from the within-sample-fit and out-of-sample-forecast.

Comparing the DAG patterns shows that both contemporaneous causal structures have the same pattern. From this salient finding, we can emphasize forecasting performances as well as having confidence in the within-sample-fit and contemporaneous causal flow results.

---

dissertation focuses on the consistency of contemporaneous causal relationships between within-sample-fit and out-of-sample-forecast.

2.5. Conclusion

Using the residuals from time series models for identifying contemporaneous causal structures, Chapter II examined how dynamic price information flows among US energy market spot prices. We estimated a causal model for the price dynamics for contemporaneous relationships by using the daily price returns of Dated Brent crude oil, Henry Hub natural gas, PJM Electricity firm on peak and COB Electricity firm on peak. We found that the Dated Brent crude oil and Henry Hub natural gas price are non-stationary, whereas PJM and COB Electricity firm on peak prices are stationary. We also found high volatility characteristics and obvious heteroscedastic problems. Hence, we concluded that the VAR-DCC-GRACH model was appropriate for this study.

Using the VAR-DCC-GARCH models, we assessed the standardized residuals from within-sample-fit and out-of-sample-forecast for modeling information flows in contemporaneous time. These processes and comparisons of forecast performance and variance-covariance matrices implied that the within-sample-fit and out-of-sample-forecast statistically performed similarly, whereas the within-sample-fit model generally outperformed the out-of-sample-forecast and contained small elements in variance-covariance matrix. We used the PC algorithm in TETRAD IV to demonstrate that the contemporaneous causal structures on standardized residuals from the two methods for calculating standardized residuals (within-sample-fit versus out-of-sample-forecast) showed the same patterns in both DAGS. Further, we found that the price returns of PJM and COB revealed an ambiguous direction of information flows from the given information.

We conclude that our hypothesis is correct: there is no difference in causal flows based on the standardized residuals from within-sample-fit and out-of-sample-forecast. Moreover, the test results for homogeneity of variance-covariance matrices and forecast performance (accuracy) support this conclusion. Therefore, we have confidence in the results of out-of-sample-forecast and its contemporaneous causal structure

CHAPTER III

INFERRING CONTEMPORANEOUS CAUSALITY USING A FACTOR AUGMENTED VECTOR AUTOREGRESSION (FAVAR) MODEL

3.1. Introduction

This chapter looks at US federal monetary policy and oil price shocks from the perspective of their causal relationships to overall economic activity. We discuss how to test for inferring the contemporaneous causal structures between the federal fund rate and a large information set, and between West Texas Intermediate (WTI) crude oil price returns and a wide range of macroeconomic and financial datasets based on a graphical causal model.

Specifically, we examine the well-known identifying assumption in the Vector Autoregression (VAR) framework of the monetary shock transmission mechanism, particularly, the unobserved latent factors that do not respond to monetary policy innovations. Bernanke et al. (2005) used this setup for the identification of the Factor Augmented Vector Autoregression (FAVAR) model. This assumption says that a monetary policy shock is orthogonal in contemporaneous time to other economic variables. In other words, the US macroeconomic and financial indicators do not respond contemporaneously to realization of the monetary policy shock.

Further, we focus on a second well-known identifying assumption, which is that an oil price shock is exogenous in contemporaneous time. In other words, the oil price shock is not caused by the systematic responses to variations in the state of the economy.

Many studies have employed VAR approaches to identify the exogenous effects of oil shocks and to estimate their effects (Barsky and Kilian 2002; Barsky et al. 2004; Bernanke et al. 1997; Hamilton 1983, 1996, 2003; Hoover and Perez 1994). This literature, however, has not reached a consensus on how these shocks affect the economy.

Inferring the contemporaneous causal structure from innovations of reduced form FAVAR models by using the graphical causal model can help us to validate the typical views regarding oil and monetary shock transmission mechanisms. Furthermore, in order to identify transmission mechanism of those structural shocks, this chapter discusses two common approaches to innovation accounting: the generating impulse response functions (IRFs) and the forecast error variance decompositions under the correlation structure of innovations in the FAVAR model. From this innovation accounting analysis, we can learn whether or not the price puzzle[10] (Sims 1992) appears in our FAVAR model.

A number of studies explain how to infer the underlying causal structure and measure the structural economic shocks' transmission mechanism using the structural equation model (SEM) and VAR frameworks. The VAR framework proposed by Sims (1980) is widely used because it provides the possibility of inferring causal flow (structures) from data using statistical properties without too much *a priori* theory and/or

[10] The price puzzle says that the response of prices to a monetary policy shock is sometimes contrary to economic theory. When monetary policy shocks are identified with innovation in the federal fund rate, the responses of output and money supply are correct, as a monetary tightening (an increase in federal fund rate) is associated with a fall in the money supply and output. However, the response of the price level is incorrect, as monetary tightening is associated with an increase in the price level rather than decrease (Sims 1992).

information from the data. The VAR framework also allows easy identification of the IRFs and forecast error variance decompositions of considered variables.

The standard VAR approach which typically contains only six to eight variables (Bernanke et al. 2005), does not allow coverage of whole datasets. Subsequently, this small scale leads to the problem of omitting variables that contain information about the structural economic shocks in the VAR analysis. Leeper et al. (1996) attempted a larger VAR framework using Bayesian priors. They tried to contain thirteen and eighteen variables in their analysis, but found that increasing the number of variables induced low efficiency of estimation. Moreover, using less than twenty variables is still insufficient than the hundreds of time series actually used as macroeconomic and financial indicators. Thus, we say that the VAR framework suffers from the curse of dimensionality.

Latent factor models provide a possible solution by summarizing the information embedded in a large dataset into a small number of factors and applying them to conventional econometric models. Stock and Watson (2002a, 2002b) developed a dynamic factor model which uses principal component analysis (PCA) to extract information from a large dataset. They applied the model in forecasting, and claimed that forecasts based on dynamic factor models show better performance compared to Autoregressive (AR) models, VAR models, and leading indicator models. Bernanke and Boivin (2003), who estimated the policy reaction function for the US Federal Reserve, concluded that using large datasets improves forecasting accuracy.

Recently, Bernanke et al. (2005) suggested a FAVAR model to incorporate a large amount of information in the VAR framework without including too many variables. Basically, they combined the FAVAR approach with the standard VAR framework and latent factor analysis. To estimate the FAVAR model, they suggested a two-step approach and the Bayesian method based on Gibbs sampling. These two approaches produced similar qualitative results; however, the two-step approach tended to produce more reasonable IRFs. This two-step approach summarizes large amounts of information about the macroeconomic and financial indicators by a small number of estimated factors using the Stock and Watson (2002a) method and then incorporates them into the FAVAR framework.

Due to these merits, the FAVAR model has become popular. It outperforms other time series models in forecasting and analyzing the structural economic shocks' transmission mechanisms (Bianchi et al. 2009; Boivin et al. 2009; Eickmeier et al. 2011; Forni and Gambetti 2010; Forni et al. 2000, 2003, 2004; Forni and Lippi 2001; Forni and Reichlin 1998; Gilchrist et al. 2009; Helbling et al. 2010; Kwon 2007; Lagana and Mountford 2005; Lagana and Sgro 2011; Moench and Ng 2011; Ng and Moench 2009; Stock and Watson 2002a, 2002b, 2005).

For these reasons, we inductively infer the contemporaneous information flows without any deductive information and investigate the structural economic shocks transmission mechanism under the FAVAR framework. We use a large dimension dataset of the US economy which is confined to 126 macroeconomic and financial time series. We show how the co-movement of these time series over time is adequately described in terms of a number of unobserved latent factors and the US federal fund rate or WTI crude oil price returns. Our two-step procedure first extracts the common factors from our large dataset following Stock and Watson (2002a, 2002b) and Bernanke et al. (2005). Second, we estimate the parameters governing their joint dynamics with the US federal fund rate and WTI crude oil price return series in each FAVAR model. Then, we identify the contemporaneous causal structures among innovations based on the residuals of our two estimated FAVAR models by using the Directed Acyclic Graph (DAG) model. We also derive and interpret the IRFs with respect to each of the augmented factors and two considered variables, and decompose the forecast error variance for each factor into the parts attributable to each of a set of innovations processes in the FAVAR model. Finally, we perform forecasting exercises considering 35 one-step-ahead forecasts, reclusively. Comparing forecasting performances between the estimated FAVAR and univariate AR models for the US federal fund rate and WTI crude oil price returns allows us to check analytical robustness.

The remainder of this chapter is structured as follows. Section 2 presents the FAVAR model, estimation, determination of number of factors, parameterization, and DAG method. Section 3 describes the dataset. Section 4 provides the two estimated models for the US federal fund rate and WTI crude oil price returns and their empirical implementations. Section 5 discusses the results of the contemporaneous causal structures, IRFs, forecast error variance decompositions, and forecast accuracies for both models. Section 6 concludes.

3.2. Methodology

This section explains how we use the FAVAR model proposed by Bernanke et al. (2005). We describe the model and provide some discussion of the adopted estimation methodology. in We also present the estimation, determination of number of factors, parameterization, and the DAG.

3.2.1. Factor Augmented Vector Autoregression (FAVAR) Model

Let F_t be a $K \times 1$ vector of unobservable factors which can summarize most of the information contained in X_t which is an $N \times 1$ stationary time series variable observed for $t = 1, 2, \cdots, T$. Y_t is an $M \times 1$ observable variable and is a subset of X_t.

We interpret F_t as the unobserved latent variables that affect many macroeconomic and financial indicators. We extract them from observations on the large information set in X_t. The number of informational time series, *N*, is large and may be larger than *T*, the number of time periods; we assume it to be much larger than *K*+*M*. We also assume that the information set is related to the unobservable factors, F_t, and the observable variables, Y_t. The FAVAR model is given by:

$$X_t = \Lambda^f F_t + \Lambda^y Y_t + \varepsilon_t \tag{3.1}$$

where Λ^f is $N \times K$ matrix of factor loadings of the factors, and Λ^y is $N \times M$ matrix of factor loadings of the observable variables. The error term ε_t has mean zero and a variance covariance matrix Σ, which is assumed to be diagonal.

Equation (3.1) is the dynamic factor model proposed by Stock and Watson (2002a). It implies that X_t is estimated by both unobservable factors and observable variables. Thus, variables (F_t and Y_t) can be correlated, and X_t is governed by a dynamic process including lagged values. Consequently, the FAVAR state equation represents the joint dynamics of F_t and Y_t; therefore, we rewrite equation (3.1) as:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \Phi(L) \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + e_t \tag{3.2}$$

where $\Phi(L)$ is a conformable polynomial in the lag operator L of finite order p. The error term e_t is expressed as $e_t = Au_t$. Specifically, the time t reduced form shock e_t consists of the time t structural shock u_t with contemporaneous relations represented through matrix A.

Next, we consider the following finite order VAR (p) approximation of the unobserved state dynamics:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \sum_{i=1}^{p} \phi_i \begin{bmatrix} F_{t-1} \\ Y_{t-1} \end{bmatrix} + Au_t \tag{3.3}$$

3.2.2. Estimation of the FAVAR Model

In the latent factor model, we describe the variability among observed variables (large dimension macroeconomic and financial indicators) by a small number of unobservable factors F_t, and model the variables as linear combinations of latent factors and innovation terms. We extract the unobservable factors F_t from the large dimension observed variables to measure joint variations in a data set, while the innovation term is covered by the part of variability which cannot be explained by latent factors.

Generally, we use PCA to extract the latent factors from the large dimension macroeconomic and financial indicators. PCA performs an orthogonal transformation onto the original dataset, which takes into account all variability of the variables. We then order this new set by variance, or consider a variance-maximizing rotation of variable space. We compute the principal components by seeking a matrix V consisting of the set of all eigenvectors of covariance matrix C such that $V^{-1}CV = D$, where D is the diagonal matrix of the eigenvalues of C.

To estimate the FAVAR model equation (3.3), we use a two-step PCA approach. In the first step, we estimate the common components $C_t = (F_t, Y_t)$ using the first K+M principal components of X_t. The first step of the estimation does not exploit the fact that Y_t is observed. However, it allows us to obtain $\hat{F}_t$ as the part of the space covered by $\hat{C}_t$ which is not covered by Y_t. In the second step, we estimate FAVAR equation (3.3) by using $\hat{F}_t$ replacing F_t. It imposes few distributional assumptions and allows for some degree of cross correlation in the idiosyncratic error term e_t (Stock and Watson 2002a).

For analytical purposes, we interpret the innovation of WTI crude oil price returns as oil shocks. Thus, the WTI crude oil price return is an observed variable Y_t, and the other variables are a subset of the large information set X_t.

However, for our federal fund rate model, any of the linear combinations underlying $\hat{C}_t$ could involve the monetary policy instrument, i.e., federal fund rate, which is observed variable Y_t. Under this condition, it would be invalid to estimate $\hat{C}_t$ and Y_t within the VAR framework. Therefore, we remove the dependence of $\hat{C}_t$ on the monetary policy instrument. Hence, we apply the two-step procedure proposed by Bernanke et al. (2005) for the model of federal fund rate distinct from the model of WTI crude oil price returns. This procedure starts from identifying variables in X_t that are not related to the monetary policy shock, and is described in detail as follows.

Bernanke et al. (2005) classified all variables in the large information set X_t as fast-moving or slow-moving variables. They claimed that there is high collinearity between the fast-moving variables and any policy shock, arguing that the fast-moving variables in X_t are highly sensitive to policy shocks, fast structural shocks, and contemporaneous information, such as financial news and economic data release. By their logic, monetary policy shock should account for the information contained in the fast-moving variables, and the slow-moving variables, e.g., unemployment rates and price indexes, are assumed to be unaffected in the month after the shock. See Appendix A for a classification of the variables.

Since the slow-moving variables are not related to the monetary policy shock contemporaneously, neither are the common components we extract from them. Thus, we express $\hat{C}_t$ as:

$$\hat{C}_t = \beta^{slow}\hat{F}_t^{slow} + \beta^{Y}Y_t + v_t \tag{3.4}$$

We remove the dependence of $\hat{C}_t$ on the monetary policy instrument to get the factors $\hat{F}_t$ for the FAVAR equation (3.3) as:

$$\hat{F}_t = \hat{C}_t - \hat{\beta}^{Y}Y_t \tag{3.5}$$

where $\hat{C}_t$ are the extracted principal components from X_t, and $\hat{\beta}^{Y}$ is the coefficient estimated from equation (3.4).

Next, we estimate FAVAR equation (3.3) which consists of $\hat{F}_t$ and Y_t in the same manner as our model of the WTI crude oil price returns. We use the federal fund rate and WTI crude oil price return as the only observed variable Y_t for each model, which is ordered last in the FAVAR model.

3.2.3. Determination of the Number of Factors

Bai and Ng (2002), who provided the econometric theory of the determination of number of factors, demonstrated that the dynamic factor in equation (3.1) always has the static factor representation in equation (3.2), where the VAR framework characterizes the dynamics of F_t. In the same paper, they stated that including more factors in the latent factor model leads to an increase in the statistical fit of the dataset, but may give rise to the dimensionality problem, whereas too few latent factors cause insufficient problems of incorporated information. Thus, they proposed various information criteria for selecting the number of factors. We follow the Bai and Ng procedure as described below.

First, we use PCA to estimate the static factors, noting that *r* is consistently selected using one of the six variants of information criteria developed in Bai and Ng (2002). All the criteria are asymptotically equivalent, but their small sample properties vary due to different specifications of the penalty term. The most widely used criterion and one of the best in terms of performance in simulations is:

$$IC_i(r) = ln\big(V(r,F)\big) + rg_i(N,T) \tag{3.6}$$

$$PCP_i(r) = V(r,F) + r\bar{\sigma}^2 g_i(N,T) \tag{3.7}$$

where $V(r,F)$ can be detailed as $V(r,F) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it} - \Lambda_i F_t)^2$, $g_i(N,T)$ is a penalty function which can be detailed as $g_i(N,T) = \left(\frac{N+T}{NT}\right) ln(min\{N,T\})$, and $\bar{\sigma}^2$ is equal to $V(r_{max}, F)$ for pre-specified value r_{max}. We allow a maximum of number of 20 factors, i.e., $r_{max} = 20$, and apply the four penalty functions $g_i(N,T)$, where $i = 1, \cdots, 3$ proposed by Bai and Ng (2002). Note that we choose this number to minimize the value of information criteria.

The penalty imposed by the second term i.e., $g_i(N,T)$, in equations (3.6) and (3.7), which is an increasing function of *N* and *T* as well as the number of factors, serves to counter-balance the minimized residual sum of squares by effecting an optimal trade-off between goodness of fit and over-fitting. Evidently, the criterion can be viewed as an extension of the Akaike information criterion (AIC) with consideration for the additional cross-sectional dimension to the time series.

3.2.4. Directed Acyclic Graphs (DAG)

The identification of underlying innovations has become a crucial issue in the analysis of the VAR framework (Lütkepohl 1999; Sims 1980). One concern is that if the innovations are not orthogonal, contemporaneous relationships exist among them. In general, the procedure used to identify u_t in equation (3.3) assumes that matrix A in equation (3.3) is a lower triangular matrix, i.e., using a Choleski decomposition of $\sum_t = AA'$. In this procedure, the elements of $u_t = A^{-1}e_t$ depend recursively on the elements of the observation vector e_t in equation (3.2). With this type of orthogonalization, u_t will depend on the ordering of variables in e_t. However, choosing appropriate orthogonal transformations of underlying innovations is often difficult to justify economically, since we assume *a priori* economic theory what the contemporaneous causal order should do. Unfortunately, formal economic theory is rarely decisive about causal order. Moreover, this contemporaneous causal structure is important to perform the innovation accounting analysis in the VAR framework and should be embedded for orthogonal innovations prior to analysis.

Therefore, we first apply the DAG approach (Swanson and Granger 1997) for identifying underlying innovations. Using this approach, we can inductively infer the contemporaneous causal information (in the form of restriction on matrix A in equation (3.3)) from the data. Then, we use the Sims-Bernanke decomposition[11] (Bernanke 1986; Sims 1986) to embed our inferred contemporaneous causal restrictions.

We can inductively infer the contemporaneous causal information (in the form of restriction on matrix A in equation (3.3)) required for identification in the FAVAR framework from data based on the graphical causal model or the DAG approach. The graphical causal model approach using the graph theory identifies the contemporaneous causal relationships among a set of observational or non-experimental data. The basic idea is that statistically inferred information about the probability distribution of the estimated residuals can be helpful in identifying the causal relationships among variables. In the graphical model approach, we can identify the contemporaneous causal inferences among the variables with relative ease, by testing the conditional independence on the

[11] The procedure of Sims-Bernanke decomposition suggested by Bernanke (1986) and Sims (1986) provide over-identified restrictions based on the theoretical background among the variables and relax the Cholesky ordering which imposes just-identified restrictions.

residuals (Bessler and Lee 2002; Bessler and Yang 2003; Demiralp and Hoover 2003; Moneta 2004, 2008; Swanson and Granger 1997). The directed acyclic graph (DAG) that is the graphical method using the graph theory has been developed by Pearl (2000) and Spirtes et al. (2000).

Contemporaneous causal relationship can be defined by causal search algorithms, which are based on statistical measures of independence and conditional correlation, and checking the patterns of conditional independence and dependence between variables. In general, the PC algorithm suggested by (Spirtes et al. 2000) is widely used (Hoover 2005; Kim and Bessler 2007). This algorithm assesses particular independence and conditional independence using the null hypothesis test. Another popular algorithm is the Greedy Equivalence Search (GES) algorithm provided by Meek (1997), then discussed and well established by (Chickering 2002, 2003). It does not require causal sufficiency, Markov condition and faithfulness, or an appropriate significance level. Therefore, we use it to generate our DAG patterns.

Starting from the premise that all variables are independent, this algorithm searches the contemporaneous causal relationships using the Bayesian scoring criterion of Schwarz loss in sequence expressed as:

$$S(G,D) = lnPr\left(D\middle|\hat{\theta},G^k\right) - \frac{h}{2}lnT \tag{3.8}$$

where Pr is the probability distribution, $\hat{\theta}$ is the maximum-likelihood estimate of the unknown parameters, D is the data available to researchers, G is DAGs, h is the number of free parameters (not equal to zero) of G, and T is the number of observations. The scoring criterion considers the trade-off between fit represented by $lnPr\left(D\middle|\hat{\theta},G^k\right)$ and parsimony modeled by the term $\frac{h}{2}lnT$.

By comparing the Bayesian scoring among all possible equivalence classes, the GES algorithm selects an equivalence class with the maximum score, meaning that it selects the best fit model among the structural equation models using the innovations from the FAVAR model. Once a local maximum is attained in the first step, the second step proceeds by single-edge deletions and compares the DAG scores in equivalence classes repeatedly. When the algorithm again reaches a local maximum, it obtains the optimal solution and DAG patterns (Chickering 2003). The results of the GES algorithm provide the Chi-square statistics and graphical patterns calculated by TETRAD IV.

3.2.5. Impulse Response Function

Since both VAR and FAVAR models have a large number of parameters, it is difficult to identify the dynamic interactions between the variables. Thus, we estimate the impulse response function (IRF), which graphically illustrates the dynamic effects of a structural shock on macroeconomic variables.

All stationary VAR (p) models can be illustrated as Moving Average (MA) process of infinite order (MA (∞)), where the current values of the variables are the weighted averages of all historical innovations. Thus, we compute the IRFs of the estimated factors and of the variables observed included in Y_t from equation (3.3) as:

$$\Phi^*(L)\begin{bmatrix}F_t\\Y_t\end{bmatrix}=e_t \tag{3.9}$$

where $\Phi^*(L) = 1 - \Phi(L)$, a matrix of conformable lag polynomial of finite order p in the lag operator L, in which $\Phi(L) = \Phi_1 L + \cdots + \Phi_p L^p$ and Φ_p is a $(K+M) \times (K+M)$ coefficient matrix. e_t is a $(K+M) \times 1$ vector of structural innovations within the diagonal covariance matrix.

The MA (∞) representation to estimate the dynamic effects is:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = [\Phi^*(L)]^{-1} e_t \tag{3.10}$$

We rewrite equation (3.10) as:

$$\begin{bmatrix} F_t \\ Y_t \end{bmatrix} = \sum_{i=0}^{\infty} \phi_i L^i \, e_t = \sum_{i=0}^{\infty} \phi_i \, e_{t-i} \tag{3.11}$$

where $\sum_{i=0}^{\infty} \phi_i L^i = [\Phi^*(L)]^{-1}$.

We can also express the estimator of X_t in equation (3.1) as:

$$\hat{X}_t = \hat{\Lambda}^f F_t + \hat{\Lambda}^y Y_t = [\hat{\Lambda}^f \quad \hat{\Lambda}^y] \begin{bmatrix} F_t \\ Y_t \end{bmatrix} \tag{3.12}$$

Using equations (3.11) and (3.12), we compute the IRF of each variable in X_t as:

$$X_t^{IRF} = [\hat{\Lambda}^f \quad \hat{\Lambda}^y] \sum_{i=0}^{\infty} \phi_i \, e_{t-i} \tag{3.13}$$

3.2.6. Forecast Error Variance Decomposition

Forecast error variance decomposition is often used as a complement to IRF when assessing the innovations of a VAR model. It determines the portion of the forecasting error of a variable, at any *t*, that is attributable to a given shock and it follows immediately from the coefficients in the moving average representation of the VAR framework and the variance of the structural shocks (Bernanke et al. 2005). This decomposition enables us to examine dynamically the relative influence of innovations in each endogenous variable out of the total variation of a variable.

We explain the proportion of the forecast error variance at horizon h of variable X_t due to the innovation e_t^{shock} by denoting that $\hat{X}_{t+h|t}$ is the optimal h-steps ahead forecast of X_{t+h} on time t information, and $X_{t+h} - \hat{X}_{t+h|t}$ is the forecast error. The fraction of the variance of the forecast error is due to the shocks, e_t^{shock}. The equation form of variance of the forecast error is:

$$\frac{Var\left(X_{t+h}-\hat{X}_{t+h|t} \middle| e_t^{shock}\right)}{Var\left(X_{t+h}-\hat{X}_{t+h|t}\right)} \tag{3.14}$$

3.3. Data Description

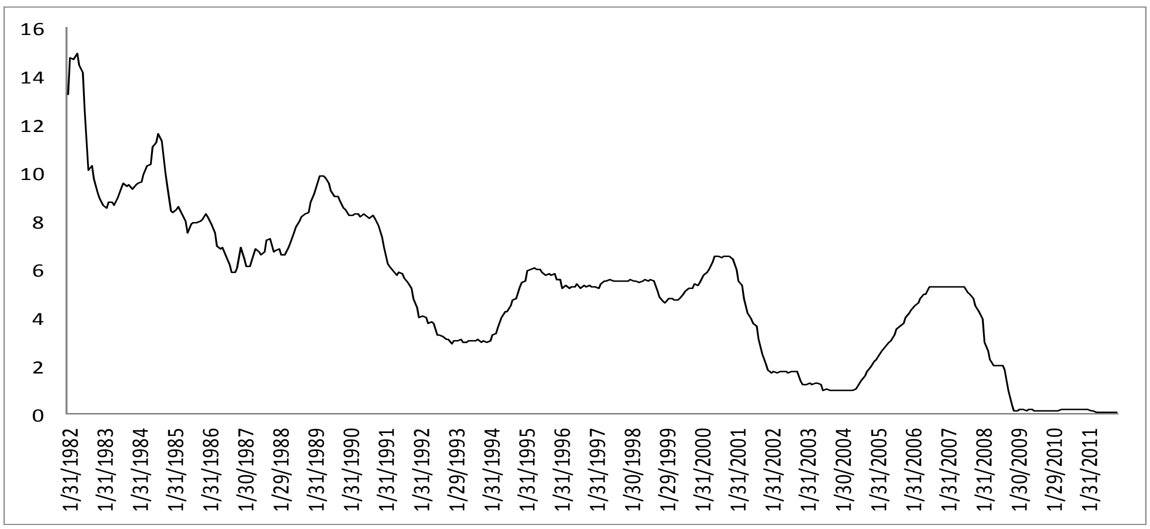
Our 127 monthly series comprises WTI crude oil price series traded on the New York Mercantile Exchange (NYMEX) and 126 macroeconomic indicator series included in the US federal fund rate. The data for WTI crude oil spot prices, federal fund rates and other macroeconomic and financial indicators derive from the DRI/McGraw Hill Basic Economics Database provided by IHS Global Insight[12]. We use observations of the data series between January 1982 and December 2008 for the in-sample estimation, and the observations of data series between January 2009 and November 2011 for the out-of-sample forecast for both models, i.e., the federal fund rate and WTI crude oil spot price, to check the robustness of estimation.

[12] We updated the Stock and Watson data from the IHS Global Insight Basic Economics Database by free trial access at Texas A&M University.

We select the 126 macroeconomic indicators based on the dataset used by Stock and Watson (2005). We choose November 2011, the most recent data available, as the endpoint of our sample. To provide different perspectives on the economy, we categorize our dataset by: real output and income; unemployment rate; employment and hours; housing starts and sales; orders and real inventories; money and credit quantity aggregates; stock prices; interest rates; spreads; exchange rates; price indexes; average hourly earnings; and miscellaneous.

The data is transformed in four ways (see Appendix A for details). First, many of the series are seasonally adjusted by the reporting agency. Second, the series are transformed by taking logarithms and/or differencing so that the transformed series are approximately stationary. In general, the first difference of logarithms (growth rates) is used for real variables, the second difference of logarithms (changes in growth rates) is used for price series, and the first differences are used for nominal interest rates. Third, outliers contained in some of the transformed series are identified as absolute median deviations larger than 6 times the inter quartile range and adjusted by replacing those observations with the one-sided median value of the preceding 5 observations. Fourth, the series are demeaned and standardized (Stock and Watson 2005). We use the DFGLS test Elliott et al. (1996) to assess the degree of integration of all series, and the Schwarz information criterion to select the optimal lag-length so that no serial correlation is left in the stochastic error term.

As mentioned, since some of the data is monthly, we use the monthly federal fund rate and WTI crude oil price and compute their price return as the first differences and first difference of logarithm, respectively. Figure 3.1 shows the federal fund rates and the first differences. Note the 1982 peak and gradually decreasing pattern. Figure 3.2 shows the WTI crude oil price and return. Note that the WTI crude oil price moved moderately until 2000, increased sharply until very recently, and fluctuated substantially in 2008~2009.



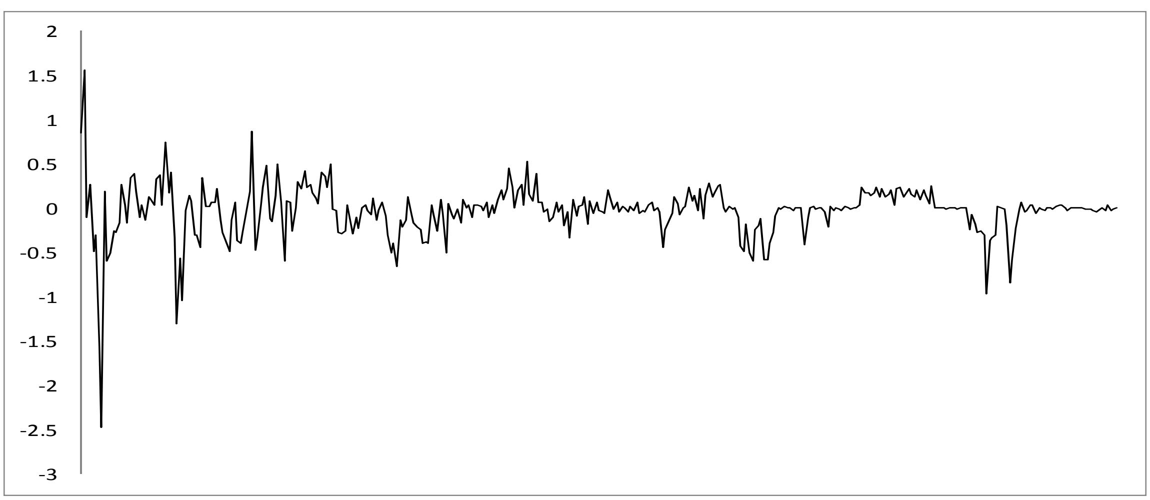
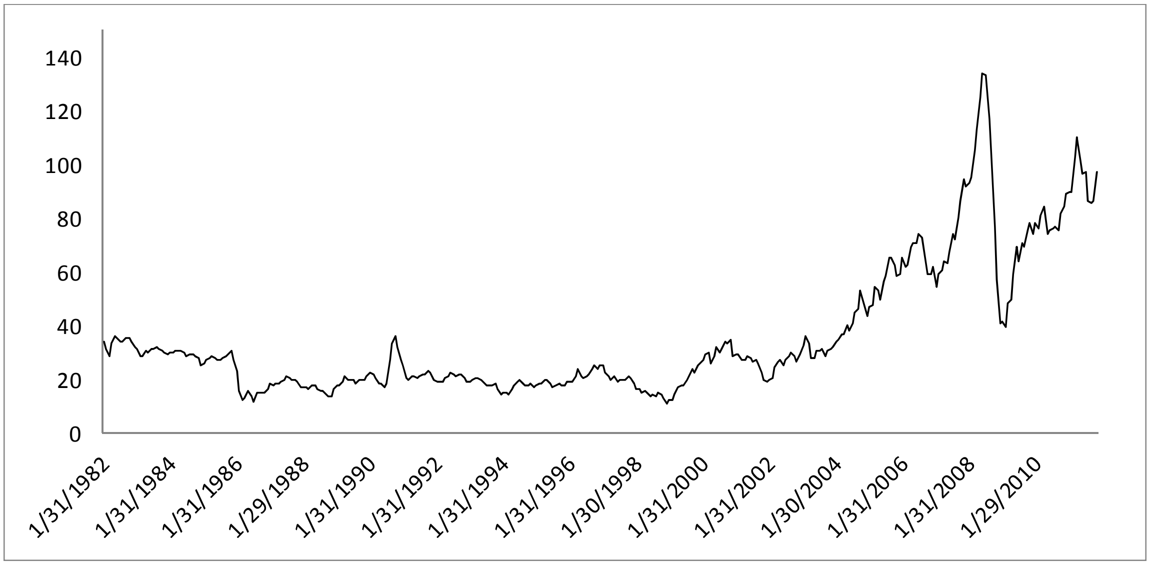


Figure 3.1. The US federal fund rates (upper graph) and first differences (lower graph) January 1982 through November 2011



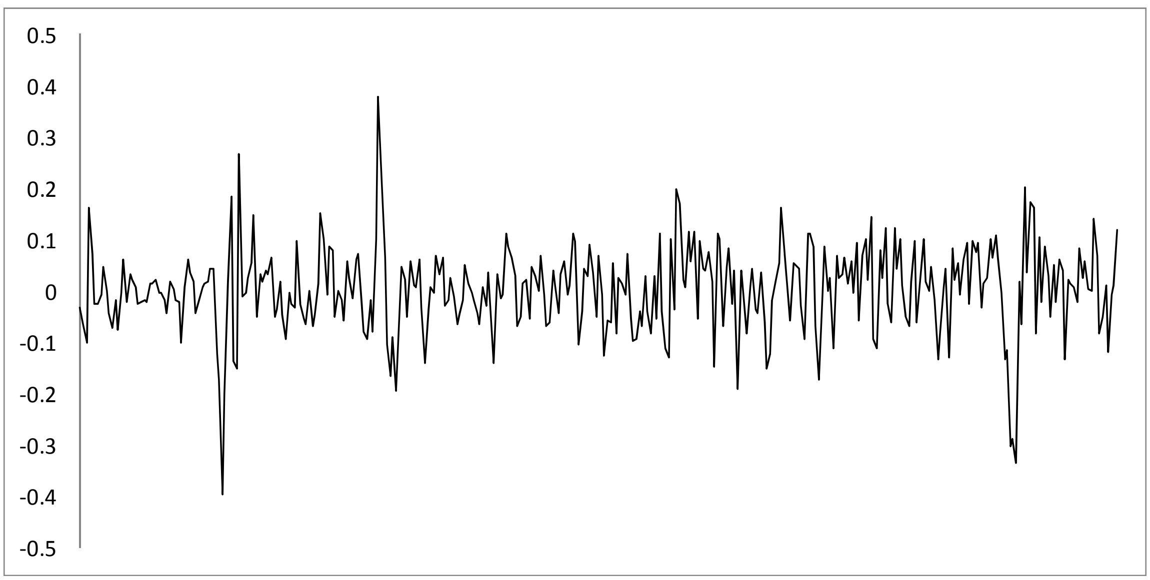


Figure 3.2. WTI crude oil monthly prices (upper graph) and returns (lower graph) January 1982 through November 2011

Table 3.1 reports the standard descriptive statistics for the first differences of the US federal fund rate and WTI crude oil price return. Both show evidence of excess kurtosis and negative skewness. Not surprisingly, the Jarque-Bera test cannot reject the hypothesis of a Gaussian distribution for WTI crude oil price return, whereas the hypothesis is rejected for the US federal fund rate. Some heteroscedasticity in the data may explain this non-normality as well as the presence of extremes (outliers). We do not explore the issue of heteroscedasticity because it is not our primary interest.

Table 3.1
Summary statistics for the monthly US federal fund rate and WTI crude oil price returns

	US Federal Fund Rate	WTI monthly Crude Oil Price Return
Mean	-0.0342	0.0028
Standard Deviation	0.3004	0.0843
Variance	0.0902	0.0071
Kurtosis	16.5184	3.2033
Skewness	-1.8225	-0.4026
Minimum	-2.4700	-0.3959
Maximum	1.5600	0.3768
JB (*p*-value)	816.81 (0.000)	2.8733 (0.238)
Number of Observation	359	359

3.4. Empirical Results

This section presents the empirical results from a FAVAR model fit to the data described earlier. Recall that we extracted the factors from 126 US macroeconomic and financial variables, i.e., the data set of Stock and Watson (2005). Technically, we construct two FAVAR models. The federal fund rate model is that the federal fund rate data series is only part of the observed variable, i.e., the federal fund rate (FFR_t), so that $Y_t = FFR_t$ and X_t is the 125 macroeconomic and financial time series. The WTI crude oil price return model uses only the state variables which include the dynamic latent factors, i.e., the WTI crude oil price return (WTI_t), so that $Y_t = WTI_t$ and X_t is the 126 macroeconomic and financial time series included in the federal fund rate.

We divide the analysis of the results into three stages. First, we infer the contemporaneous causal structure with innovations of the FAVARs by using DAG with the GES algorithm. Second, we describe the IRFs of each augmented factor and the two considered variables. Third, we look at the variance decomposition of the prediction errors and check our results for robustness by comparing the forecasting performance to the univariate AR model.

3.4.1. Determination of Estimating the Number of Factors

Table 3.2 shows the estimated numbers of factors for the two models. Clearly, there is no agreement on the optimal number of factors in both cases. We note that this result aligns with previous empirical studies which also find instability in determining the correct number of factors. According to Bai and Ng (2002)'s information criteria, the optimal number of factors is 10 to 18 for the federal fund rate model and 10 to 17 for the WTI crude oil price return model. Table 3.3 reports information on the autocorrelation and the explanatory power of estimated factors F_t for both models. We note that the first 3 factors for both models only explain 31.0% of the joint-variance of the 125 and 126 data in both cases, whereas 10 factors reach 55.9% and 55.8%, respectively. Therefore, we consider the set of the first 10 factors for both models as the potential set of regressors based on the information criteria IC_2 proposed in Bai and Ng (2002). The factors' autocorrelations up to 3 lags (table 3.3) show that most factors appear to be persistent.

Table 3.2
Static factors selection results

Method	Number of Factors	
	Model for the federal fund rate	Model for the WTI crude oil price return
IC_1	13	13
IC_2	10	10
IC_3	20	20
PCP_1	18	17
PCP_2	16	16
PCP_3	20	20

Note: IC_i and PCP_i respectively denote the number of factors given by the information criteria IC and PCP estimated with penalty function $g_i(N, T)$.

Table 3.3
Summary statistics of $f_{t,i}$ for $i = 1, \cdots, 10$

Factor i	Model for the federal fund rate				Model for the WTI crude oil price return			
	AR (1)	AR (2)	AR (3)	R_i^2	AR (1)	AR (2)	AR (3)	R_i^2
1	0.8341	0.8344	0.7802	0.1627	0.8356	0.8355	0.7803	0.1626
2	0.1746	0.0248	0.1988	0.2392	0.1579	0.0362	0.2078	0.2384
3	0.5277	0.4562	0.4520	0.3104	0.5201	0.4363	0.4344	0.3096
4	0.2668	0.2121	0.2407	0.3622	0.2705	0.2135	0.2410	0.3617
5	0.3035	0.3010	0.2850	0.4052	0.2976	0.3091	0.2903	0.4044
6	0.3807	0.2322	0.2900	0.4445	0.3454	0.1884	0.2521	0.4434
7	0.4356	0.3933	0.3628	0.4794	0.5255	0.4565	0.4182	0.4789
8	-0.1399	0.0678	0.1749	0.5098	-0.1799	0.0582	0.1739	0.5092
9	-0.0539	-0.0444	0.0645	0.5352	-0.0528	-0.0461	0.0675	0.5344
10	-0.1782	-0.0441	0.0686	0.5592	-0.1802	-0.0381	0.0662	0.5583

Note: For $i = 1, \cdots, 10$, $f_{t,i}$ for the federal fund rate and WTI crude oil price return models are estimated by PCA using a panel of data with 125 and 126 indicators of macroeconomic activity from January 1982 to December 2008 (324 time-series observations), respectively. The data is transformed (taking logs and differenced where appropriate) and standardized prior to estimation. AR (*p*) denotes the *p*-th autocorrelation. The relative importance of the common component, R_i^2, is calculated as the fraction of total variance in the data explained by factors 1 to *i*.

3.4.2. *Interpreting the Estimated Factors*

In attempting to assign interpretations to the factors, we are aware of the pitfalls. Since the factors are identified only up to an orthogonal transformation, a detailed interpretation of the individual factors is unwarranted. However, looking at the fit of the regression of the individual series in our dataset against each of the factors still gives us an idea of the underlying economic concepts.

Following Stock and Watson (2002b), we use simple linear regressions to estimate the relationship between every variable against each factor to determine if a particular factor strongly relates to a specific group of macroeconomic variables. We graph R^2, or explanatory powers, of the first 10 factors for individual series as bar charts in figures 3.3 through 3.12 for the federal fund rate model, and figures 3.13 through 3.22 for the WTI crude oil price return model. The numbers on the horizontal axis refer to variables (see Appendix A for details).

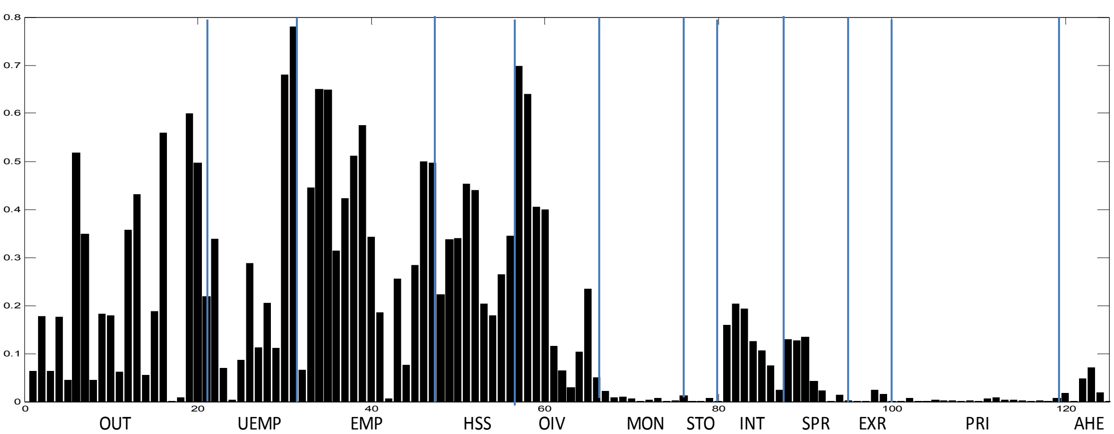


Figure 3.3. R^2 values for factor 1 of the federal fund rate model

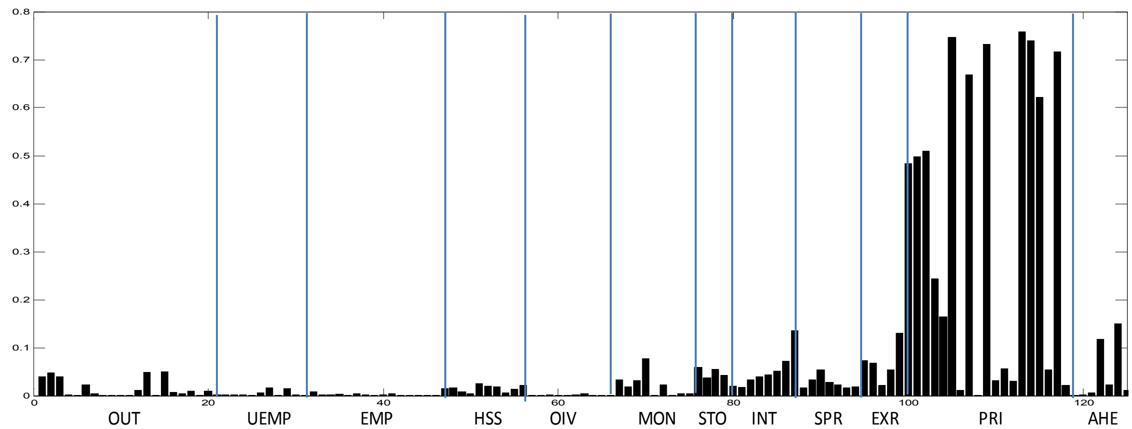


Figure 3.4. R^2 values for factor 2 of the federal fund rate model

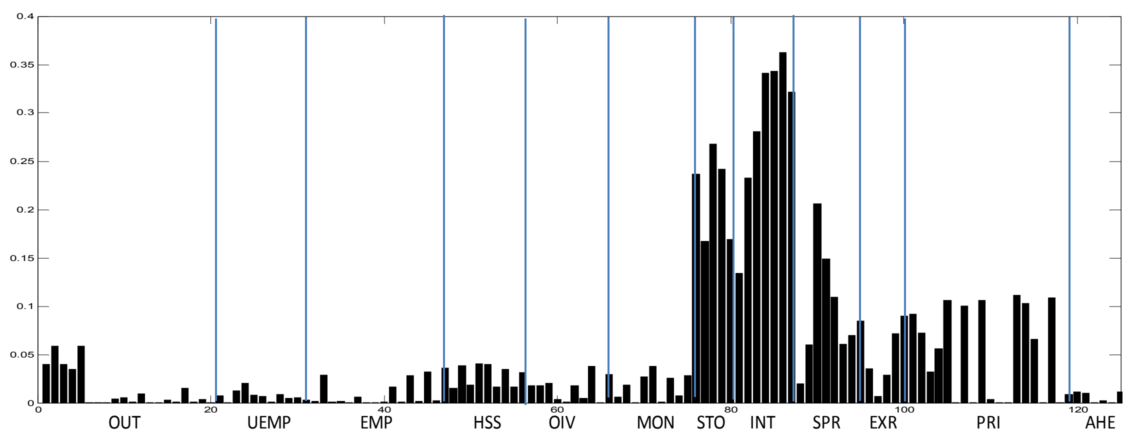


Figure 3.5. R^2 values for factor 3 of the federal fund rate model

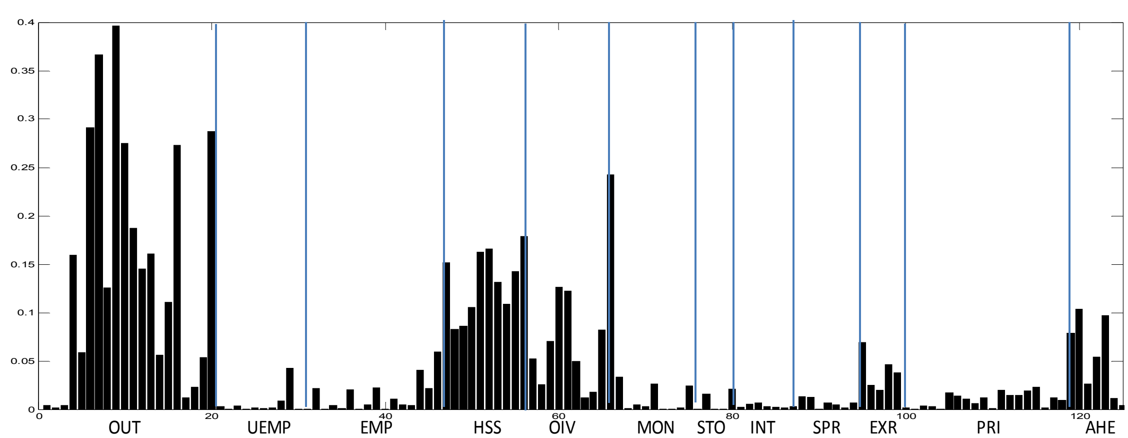


Figure 3.6. R^2 values for factor 4 of the federal fund rate model

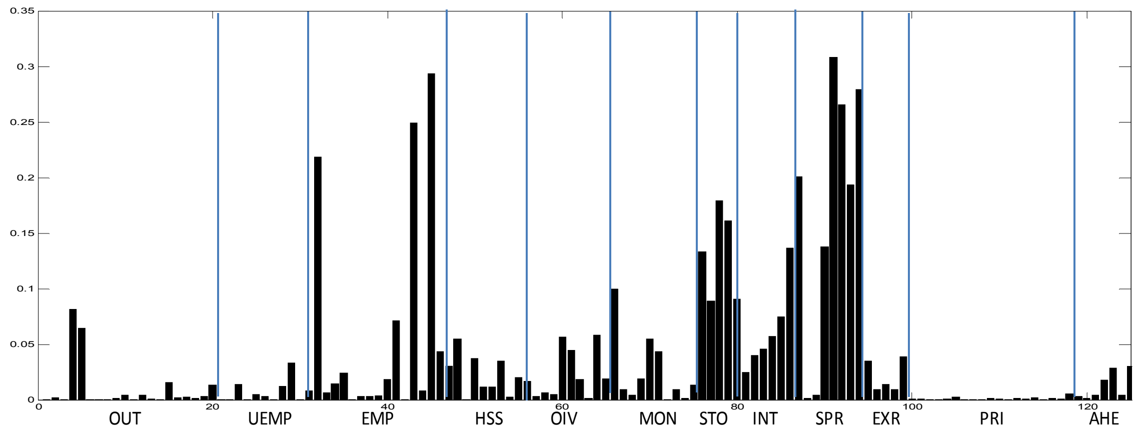


Figure 3.7. R^2 values for factor 5 of the federal fund rate model

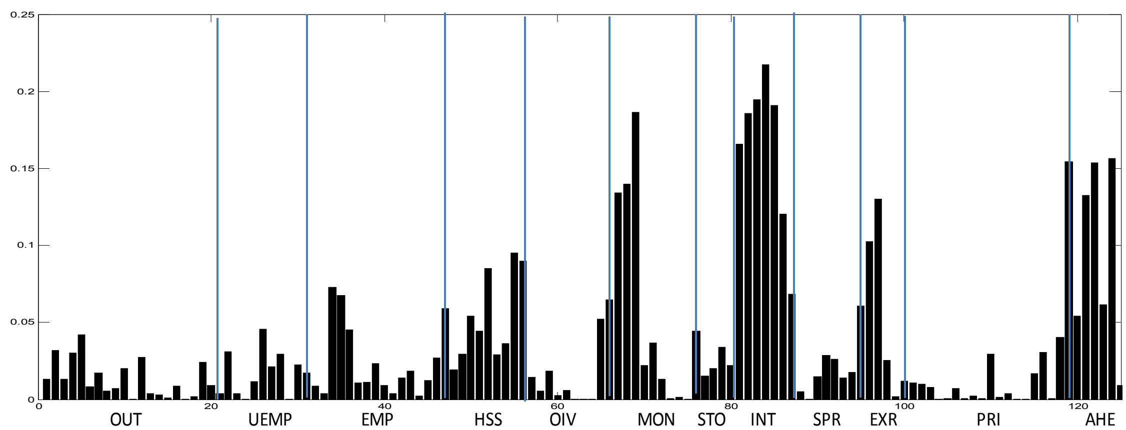


Figure 3.8. R^2 values for the factor 6 of the federal fund rate model

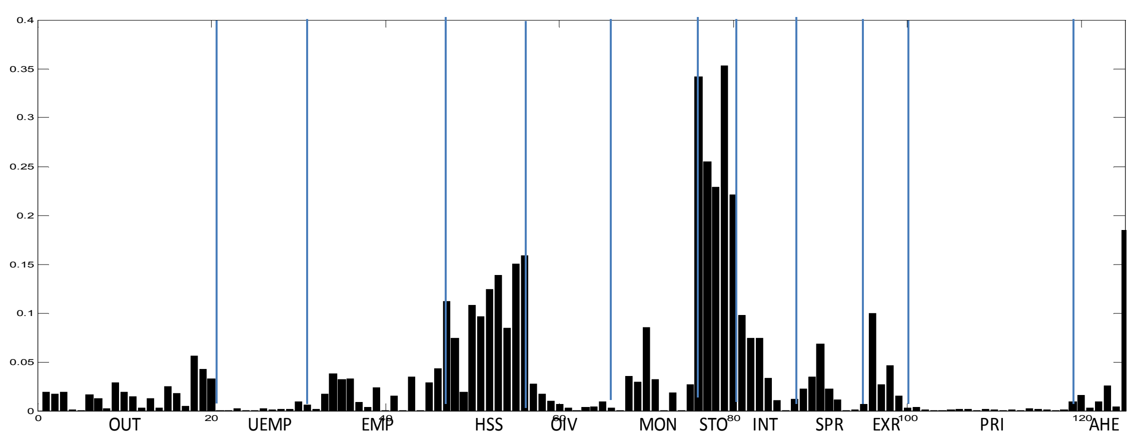


Figure 3.9. R^2 values for factor 7 of the federal fund rate model

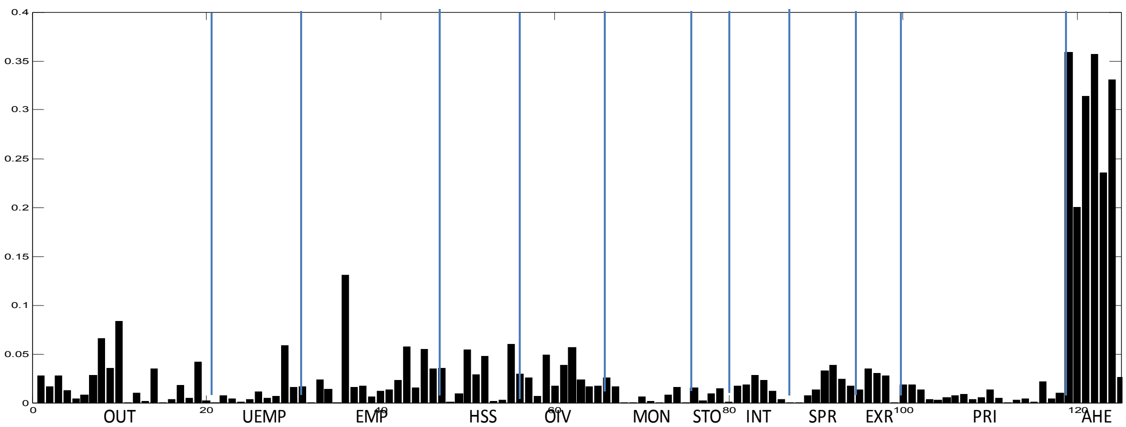


Figure 3.10. R^2 values for factor 8 of the federal fund rate model

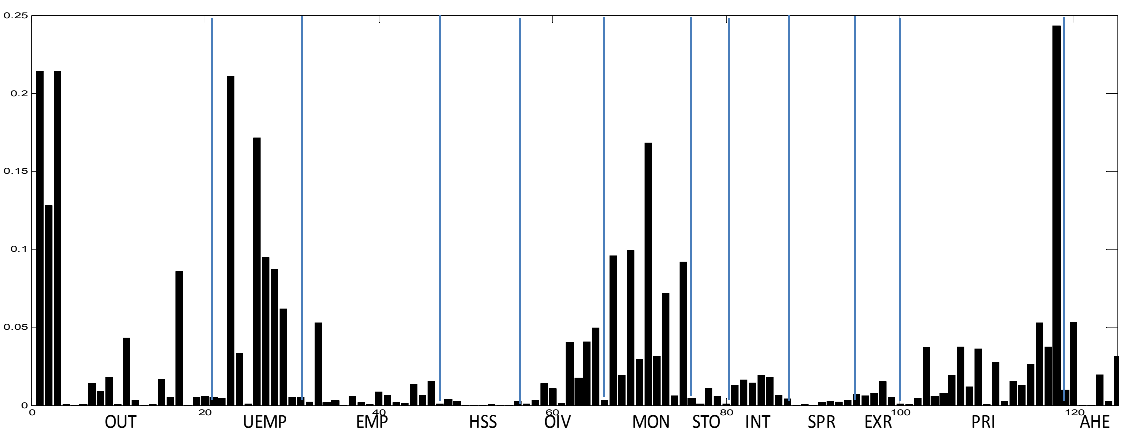


Figure 3.11. R^2 values for factor 9 of the federal fund rate model

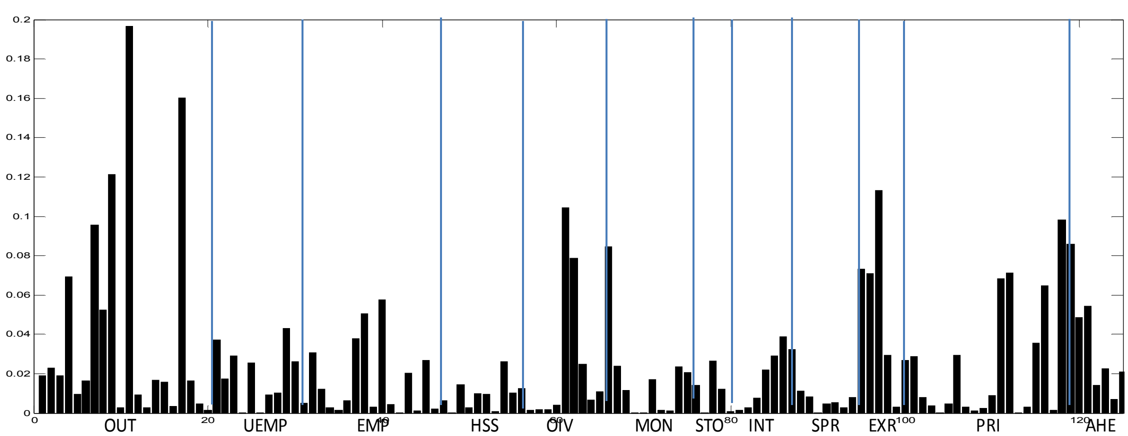


Figure 3.12. R^2 values for factor 10 of the federal fund rate model

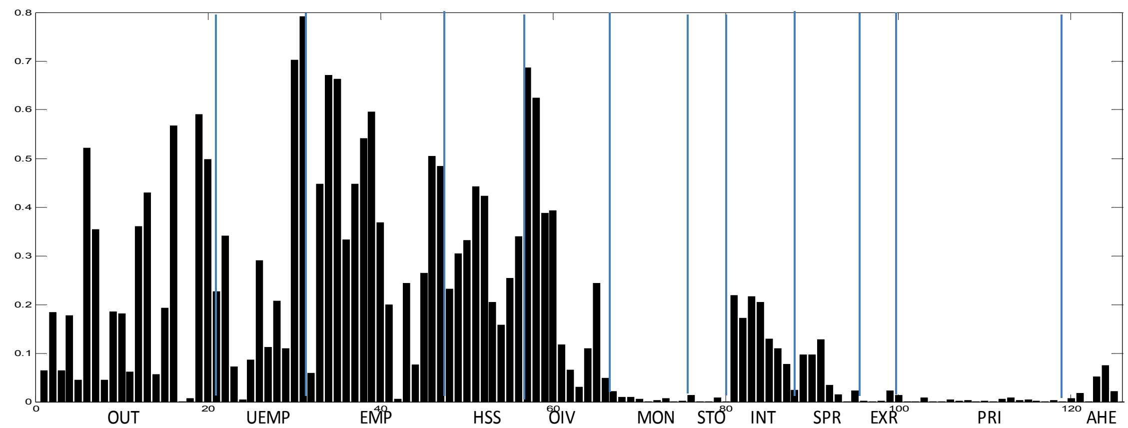


Figure 3.13. R^2 values for factor 1 of the WTI crude oil price return model

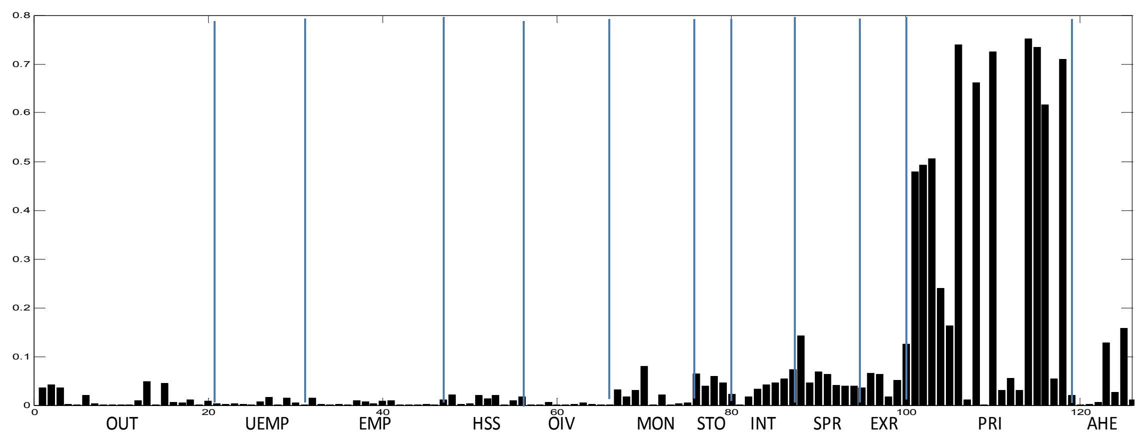


Figure 3.14. R^2 values for factor 2 of the WTI crude oil price return model

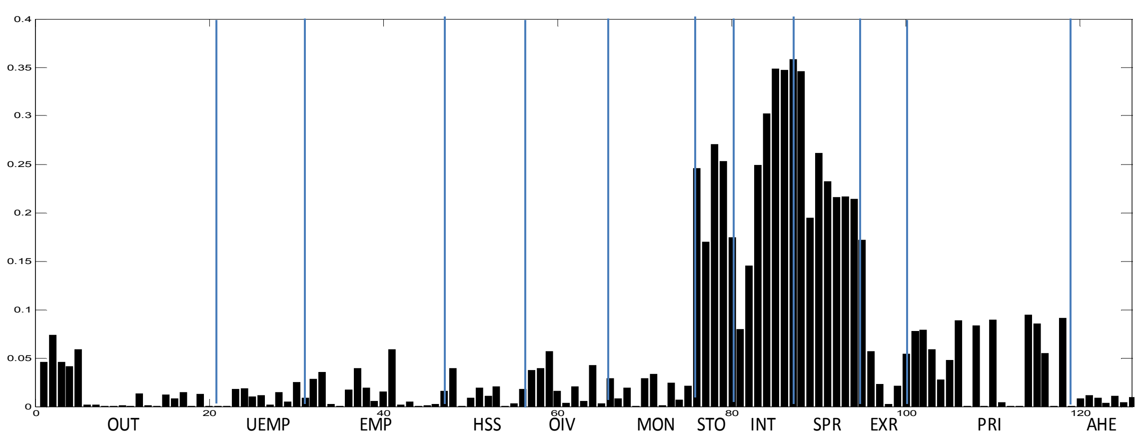


Figure 3.15. R^2 values for factor 3 of the WTI crude oil price return model

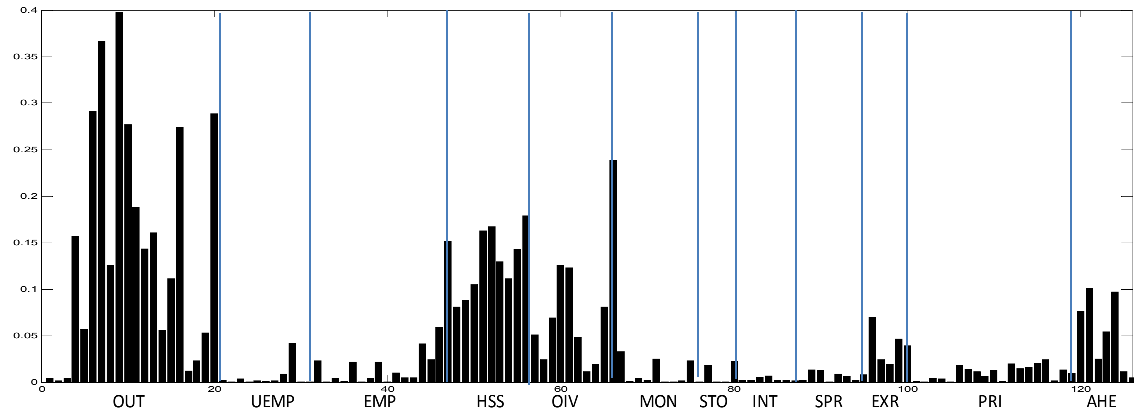


Figure 3.16. R^2 values for factor 4 of the WTI crude oil price return model

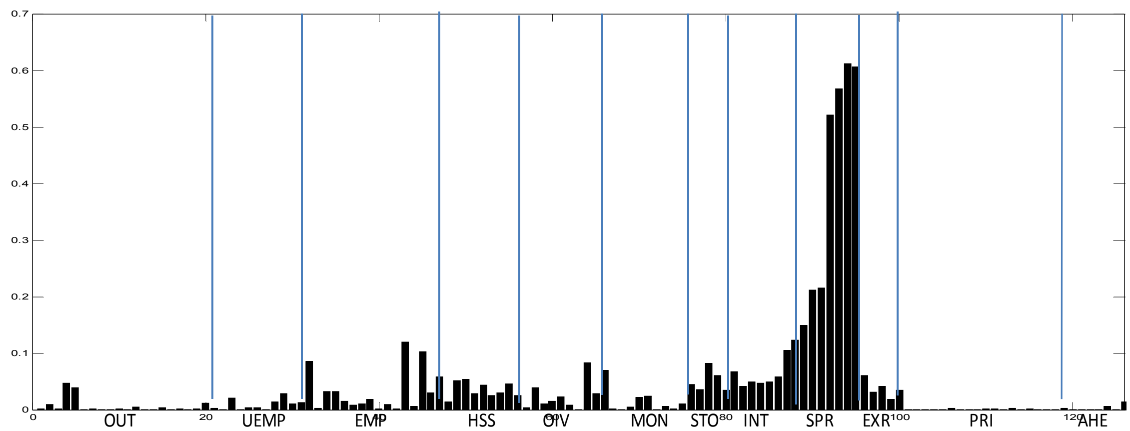


Figure 3.17. R^2 values for factor 5 of the WTI crude oil price return model

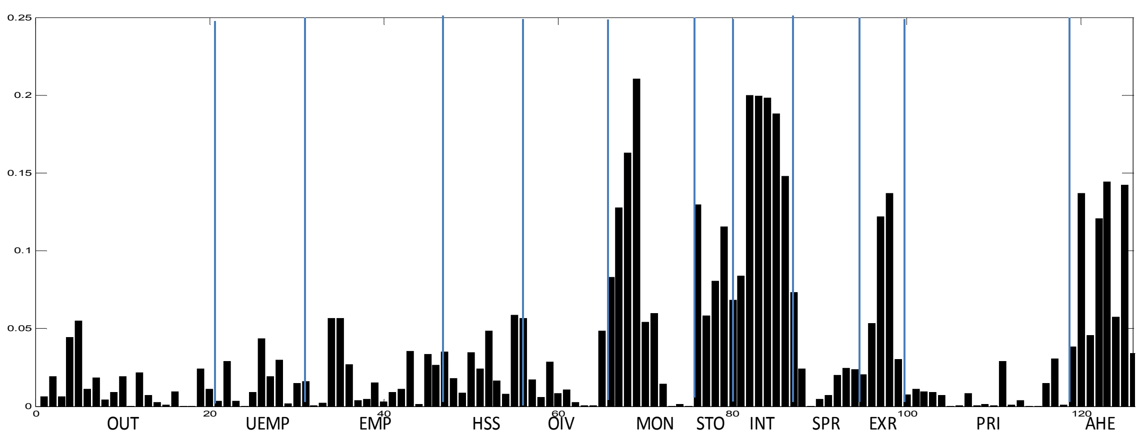


Figure 3.18. R^2 values for factor 6 of the WTI crude oil price return model

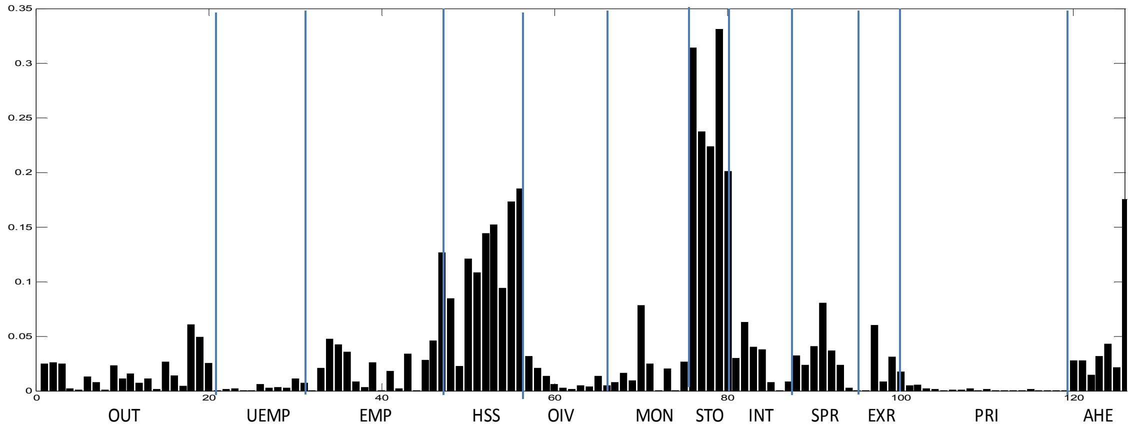


Figure 3.19. R^2 values for factor 7 of the WTI crude oil price return model

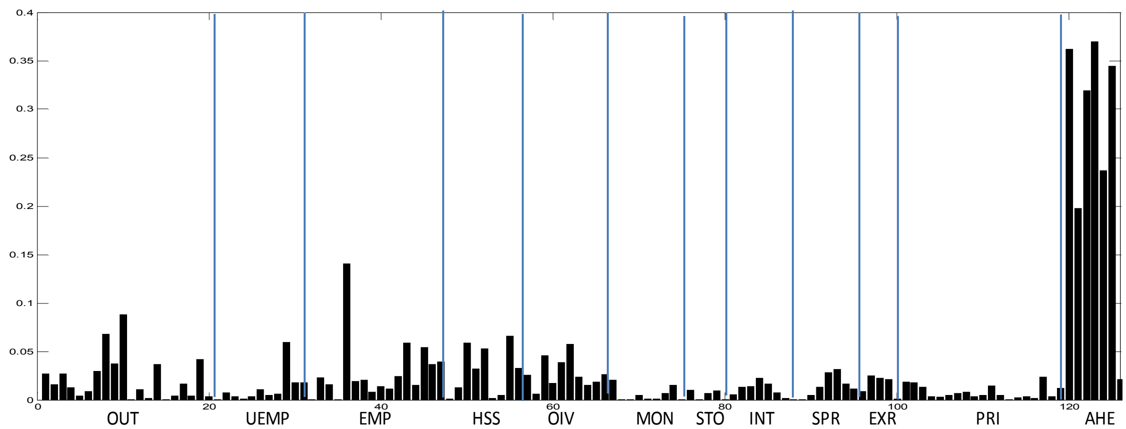


Figure 3.20. R^2 values for factor 8 of the WTI crude oil price return model

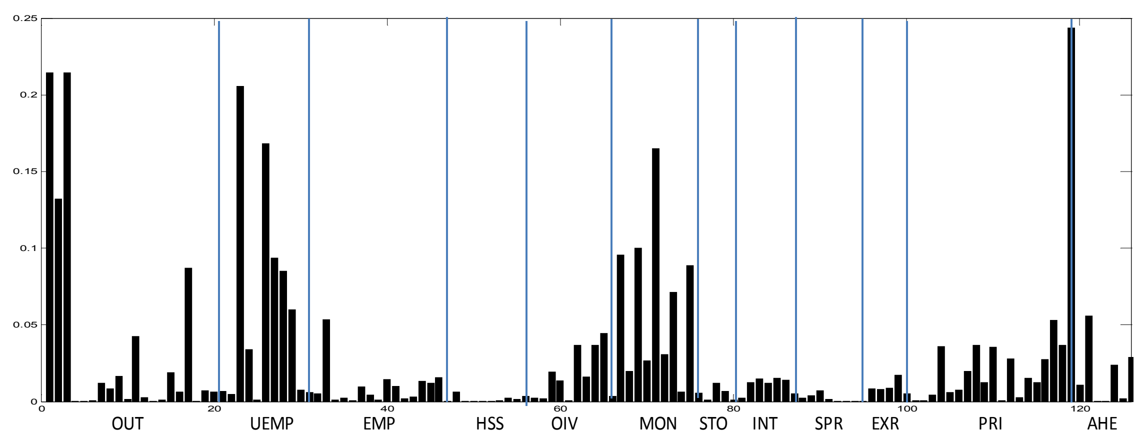


Figure 3.21. R^2 values for factor 9 of the WTI crude oil price return model

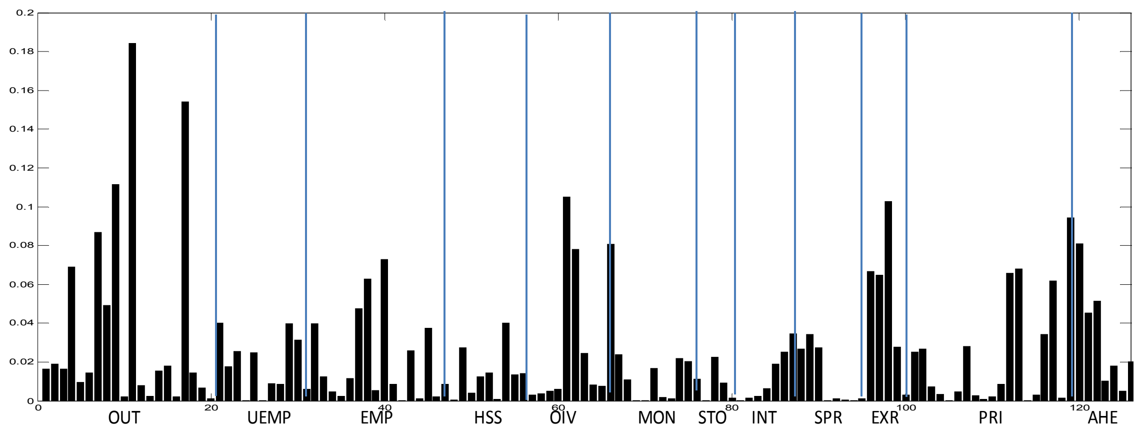


Figure 3.22. R^2 values for factor 10 of the WTI crude oil price return model

The results for both models are similar except for factor 5. Nevertheless, factor 5 represents a specific group of macroeconomic variables. Specifically, factor 1 is clearly an employment and hour factor with R^2 values with employment and hour variables as well as housing starts variables often exceeding 50%. Factor 2 is primarily related to price index variables with R^2 values often exceeding 70%. Factor 3 is primarily related to interest rate variables. Factor 4 is a real output factor with higher R^2 values. Factor 5 reveals the highest explanatory power of spread variables. Factor 6 is primarily related to interest rate as well as money and credit quantity aggregates variables. Factor 7 shows a somewhat high explanatory power of stock market price variables. Factor 8 is primarily related to average hourly earnings. Factor 9 is primarily related to real income and unemployment. Factor 10 is primarily related to the exchange rate.

3.4.3. Determinations of the Lag Order

We use the conventional information criteria for standard VAR in the FAVAR. Table 3.4 reports the AIC, SIC and HQIC. For the FFR rate model, AIC and HQIC indicate two as the optimal number of lag ($p = 2$) and SIC indicates one lag ($p = 1$). For the WTI model, SIC indicates $p = 1$, HQIC indicates $p = 2$ and AIC indicates $p = 3$. One reason for the differences is that the magnitude of penalties to determine the optimal number of lag in the FAVAR model in equation (3.3) varies substantially for the optimal lag selection criteria. Thus, we apply the modeling philosophy of parsimony (Box et al. 1976), and follow the suggestion of SIC (*p=1*) for both models.

Table 3.4
Determination result of optimal lag length for the model of the federal fund rate (FFR) and the WTI crude oil price return (WTI)

Lag Order	Akaike Information Criterion (AIC)		Schwarz Information Criterion (SIC)		Hannan and Quinn Information Criterion (HQIC)	
	FFR	WTI	FFR	WTI	FFR	WTI
0	-1.1361	-0.6351	-1.0060	-0.5136	-1.0841	-0.5868
1	-13.362	-8.5008	-11.801*	-7.0427*	-12.739	-7.9204
2	-14.282*	-9.3566	-11.289	-6.5620	-13.087*	-8.2441*
3	-14.199	-9.4043*	-9.7747	-5.2731	-12.432	-7.7598
4	-14.012	-9.3541	-8.1563	-3.8863	-11.673	-7.1775

Note: * indicates the most appropriate lag order for the considered model; the information criteria used to identify the optimal lag-length (*p*) of a VAR process are $AIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{2n}{T}\right)$, $SIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{nlnT}{T}\right)$, and $HQIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{2nln(lnT)}{T}\right)$, where $\widehat{\Omega}_p$ is the maximum likelihood estimate of variance-covariance matrix of Ω, *p* is the proposed lag-length, *n* is the number of variables, and *T* is the sample size.

3.4.4. Contemporaneous Causal Structure

Figures 3.23 and 3.24 show the inferred contemporaneous information flows, and tables 3.5 and 3.6 show the correlation-covariance matrices among innovations for both models. For ease of explanation, we split the entire contemporaneous causal structure into three sectors. The first part is the real economy sector, which consists of employment and housing starts (factor 1), output (factor 4), average hourly earnings (factor 8), and real income and unemployment (factor 9). The second part is the money and interest sector, which consists of the federal fund rate (recall that we include it in the FFR model only), interest rate (factor 3), spread (factor 5), and interest rate and money aggregate (factor 6). The third part consists of the WTI crude oil price return (recall that we include it in the WTI model only), price (factor 2), stock (factor 7), and exchange rate (factor 10).

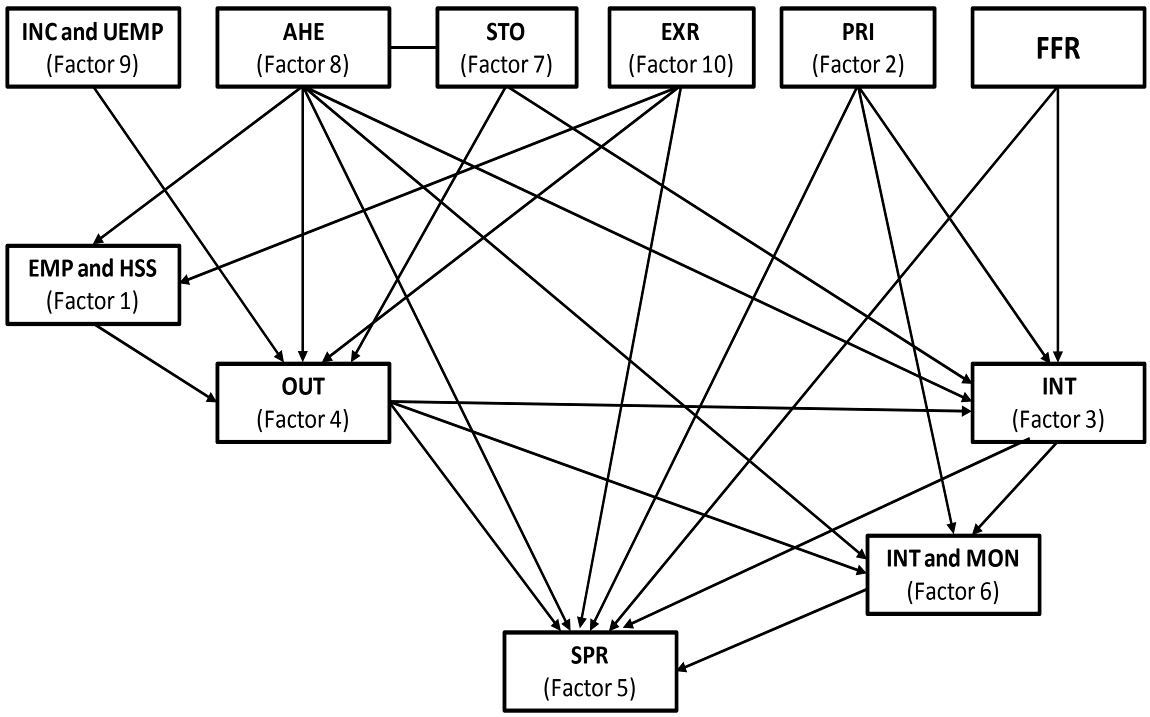


Figure 3.23. The contemporaneous causal patterns inferred by the GES algorithm for the federal fund rate model

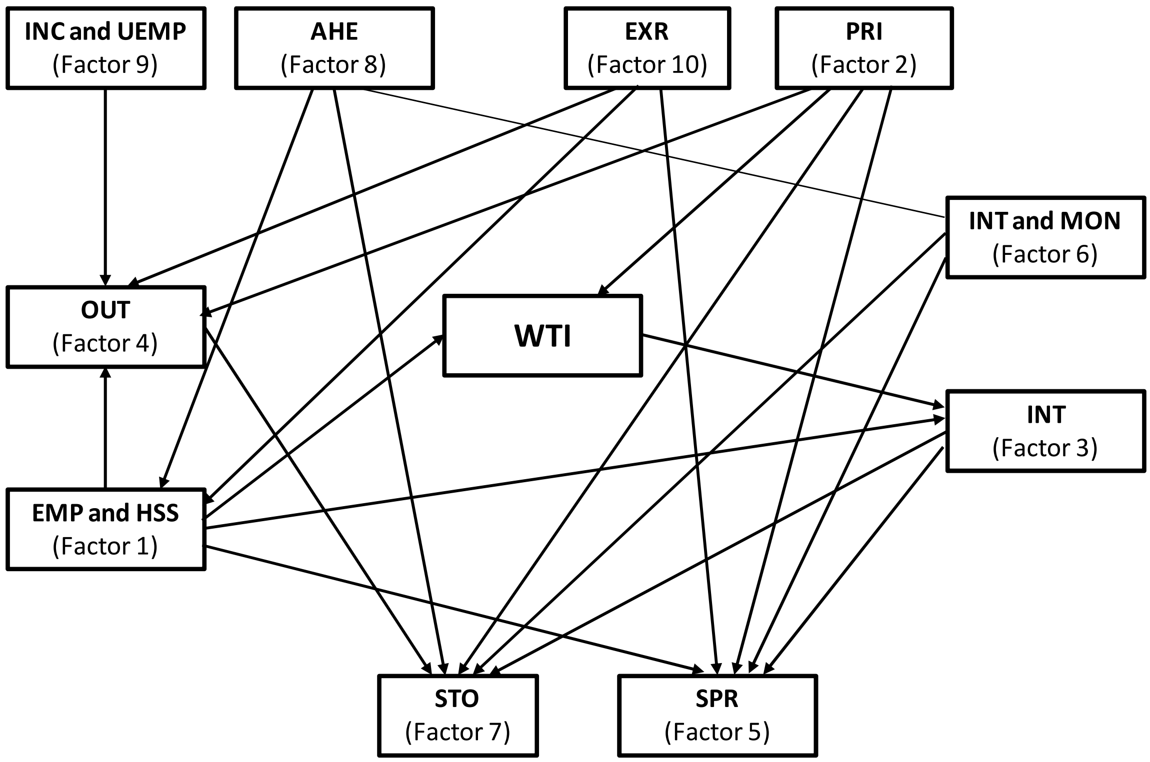


Figure 3.24. The contemporaneous causal patterns inferred by the GES algorithm for the WTI crude oil price return model

Table 3.5.
Correlation-covariance matrix among innovations of FAVAR model for the federal fund rate

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	Factor10	FFR
Factor1	0.032	-0.002	0.005	0.022	-0.001	-0.001	0.002	0.006	-0.002	0.005	-0.003
Factor2	-0.034	0.060	-0.008	0.004	-0.002	-0.005	-0.001	-0.003	0.004	0.003	0.001
Factor3	0.143	-0.161	0.038	0.004	-0.011	-0.003	-0.005	-0.001	0.002	0.002	0.007
Factor4	0.692	0.089	0.114	0.032	-0.001	0.000	0.005	0.006	0.003	0.000	0.001
Factor5	-0.052	-0.090	-0.543	-0.041	0.010	-0.006	0.006	-0.001	-0.001	-0.003	0.001
Factor6	-0.050	-0.129	-0.101	-0.018	-0.420	0.022	-0.007	-0.005	0.002	0.000	0.001
Factor7	0.075	-0.027	-0.181	0.182	0.392	-0.356	0.020	-0.003	-0.001	0.002	0.002
Factor8	0.235	-0.099	-0.028	0.251	-0.072	-0.249	-0.154	0.019	0.001	-0.001	0.001
Factor9	-0.091	0.121	0.061	0.134	-0.046	0.091	-0.070	0.028	0.020	-0.002	0.000
Factor10	0.193	0.080	0.060	0.009	-0.181	-0.024	0.081	-0.062	-0.098	0.019	0.000
FFR	-0.057	0.015	0.153	0.033	0.041	0.026	0.048	0.026	0.012	-0.013	0.061

Note: The lower triangular is for correlation values and the upper triangular is for covariance values. See Appendix A for a description of the factors, where variables are in the same order.

Table 3.6.
Correlation-covariance matrix among innovations of FAVAR model for the WTI crude oil price return

	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8	Factor9	Factor10	WTI
Factor1	0.033	-0.002	0.005	0.022	0.000	0.001	0.001	0.007	-0.002	0.005	0.022
Factor2	-0.038	0.058	-0.006	0.003	-0.003	-0.004	-0.002	-0.003	0.004	0.003	0.093
Factor3	0.143	-0.135	0.038	0.004	-0.010	-0.002	-0.006	-0.001	0.002	0.002	-0.050
Factor4	0.690	0.078	0.125	0.032	0.000	-0.001	0.004	0.007	0.003	0.000	0.021
Factor5	0.026	-0.124	-0.522	-0.027	0.010	-0.002	0.003	0.000	-0.001	-0.002	0.010
Factor6	0.018	-0.102	-0.057	-0.022	-0.134	0.030	-0.008	-0.003	0.002	0.000	0.002
Factor7	0.030	-0.053	-0.215	0.171	0.192	-0.351	0.018	-0.004	-0.001	0.001	0.005
Factor8	0.272	-0.094	-0.044	0.262	0.015	-0.136	-0.227	0.020	0.000	-0.001	0.004
Factor9	-0.096	0.109	0.065	0.128	-0.067	0.067	-0.061	0.013	0.020	-0.002	0.010
Factor10	0.194	0.078	0.068	0.010	-0.175	-0.010	0.068	-0.047	-0.101	0.019	-0.004
WTI	0.135	0.423	-0.281	0.126	0.109	0.014	0.037	0.029	0.075	-0.031	0.834

Note: The lower triangular is for correlation values and the upper triangular is for covariance values. See Appendix A for a description of factors, where variables are in the same order.

For the FFR model, we interpret the results as follows: (a) there is observational equivalence between stock market (factor 7) innovations and average hourly earnings (factor 8) innovations, i.e., the causal direction cannot be decided based on statistical observations only, or either direction between them is statistically equivalent; (b) there are several first causes (causal roots) and last effects (causal sinks), i.e., the federal fund rate variable, price (factor 2), real income and unemployment (factor 9), and exchange rate (factor 10) are indicated to be causal root, whereas the spread (factor 5) is a causal sink in terms of innovations discovery.

In the real economy sector case, we note that average hourly earnings (factor 8) affect output (factor 4) either directly or through employment and housing starts (factor 1). Factor 4 also causes the real income and unemployment (factor 9) in contemporaneous time. Stock (factor 7) and exchange rate (factor 10) affect factor 4 at the same time. In the money and interest sector case, the federal fund rate directly affects interest rate (factor 3) and spread (factor 5). Interest rate (factor 3) affects spread (factor 5) either directly or through interest rate and money aggregates (factor 6). Price (factor 2), output (factor 4), average hourly earnings (factor 8), and exchange rate (factor 10) affect spread (factor 5) in contemporaneous time. In the third sector case, there is no direct causal link among price (factor 2), stock (factor 7), and exchange rate (factor 10). However, figure 3.23 and table 3.5 show that the fulfillments of those factors are a causal root of the entire contemporaneous causal structure.

Interestingly, the effects of the monetary policy integrate into spread (factor 5), but do not transmit to other sectors of the overall economy in contemporaneous time.

Moreover, we note that the federal fund rate shock is exogenous in contemporaneous time. Explicitly, our results show that the US macroeconomic and financial indicators do not respond contemporaneously to realization of the monetary policy shock. In this respect, the monetary transmission mechanism identified in this inferred causal structure is consistent with the identifying assumption of the FAVAR model in Bernanke et al. (2005).

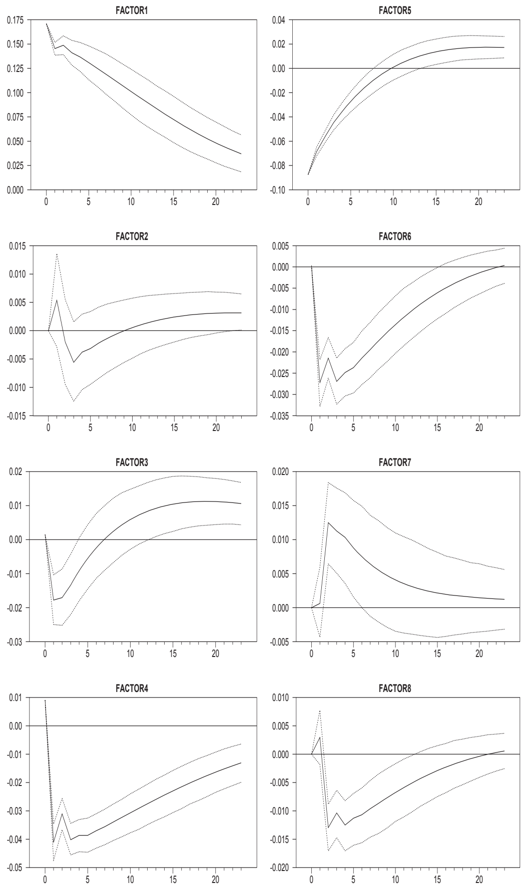
Similarly, we interpret the results for the WTI model as follows: (a) there is observational equivalence between interest rate and money aggregate (factor 6) innovations and average hourly earnings (factor 8) innovations; and (b) there are several first causes (causal roots) and last effects (causal sinks). Price (factor 2), average hourly earnings (factor 8), real income and unemployment (factor 9), and exchange rate (factor 10) are indicated as causal root innovations. Spread (factor 5) and stock (factor 7) are indicated as causal sink in terms of innovations discovery.

In the real economy sector case, we find that the contemporaneous causal order is average hourly earnings (factor 8), employment and housing starts (factor 1), and output (factor 4). Real income and unemployment (factor 9) directly affects output (factor 4). Moreover, the effect of output (factor 4) influences the real economy sector as well as the immediate cause of price and exchange rate transfer to stock market (factor 7) in contemporaneous time. For the money and interest sector case, spread (factor 5) directly causes interest rate (factor 3) as well as interest rate and money aggregates (factor 6). Spread (factor 5) is also directly affected by employment and housing starts (factor 1), price (factor 2), and exchange rate (factor 10). In the third sector case, price (factor 2) is an immediate cause of WTI crude oil price return and stock (factor 7), whereas exchange rate (factor 10) shows no direct causal link with other factors of the intra-sector. However, exchange rate (factor 10) affects employment and housing starts (factor 1) and output (factor 4) in the real economy sector as well as spread (factor 5) in the money and interest sector. Surprisingly, WTI crude oil price return is influenced by real economy (factor 1) and price (factor 2), whereas it has an effect on interest rate (factor 3) in the money and interest sector in contemporaneous time. In general the typical view of the oil price shocks transmission mechanism is that the oil price shocks are not affected by the systematic response to variation in the overall economy since the shocks are exogenous in contemporaneous time. However, figure 3.24 and table 3.6 show that the WTI crude oil price return is clearly not exogenous in contemporaneous time. In this respect, we argue that the oil price shocks transmission mechanisms identified in this causal information is inferred from the data. Moreover, WTI crude oil price returns are a bridge to transmit the causal influences from the overall economy into the money and interest sector in contemporaneous time.

3.4.5. Impulse Response Function

Based on the identified contemporaneous causal relationships in figures 3.23 and 3.24, we compute the IRFs of both FAVAR models. This section presents the results of each of the augmented 10 factors and 1 variable, i.e., the federal fund rate or WTI crude oil price return. From this, we verify information on the direction and significance of dynamic responses for a 24 month-horizon following an initial shock of each augmented factor and variable. The solid line indicates the estimated response, and the upper and lower dashed lines plotted in each graph represent 90% bootstrap confidence intervals based on 1500 bootstrap samples. Figures 3.25 through 3.35 and figures 3.36 through 3.46 show the IRFs of both FAVAR models.

In the FFR model, for the real economy sector, employment and housing starts (factor 1) and real income and unemployment (factor 8) move opposite to output (factor 4) and average hourly earnings (factor 8). For the money and interest sector, the federal fund rate moves opposite to interest rate (factor 3), factor 5 (spread), and interest rate and money aggregate (factor 6). For the third sector, stock (factor 7) moves opposite to price (factor 2) and exchange rate (factor 10).



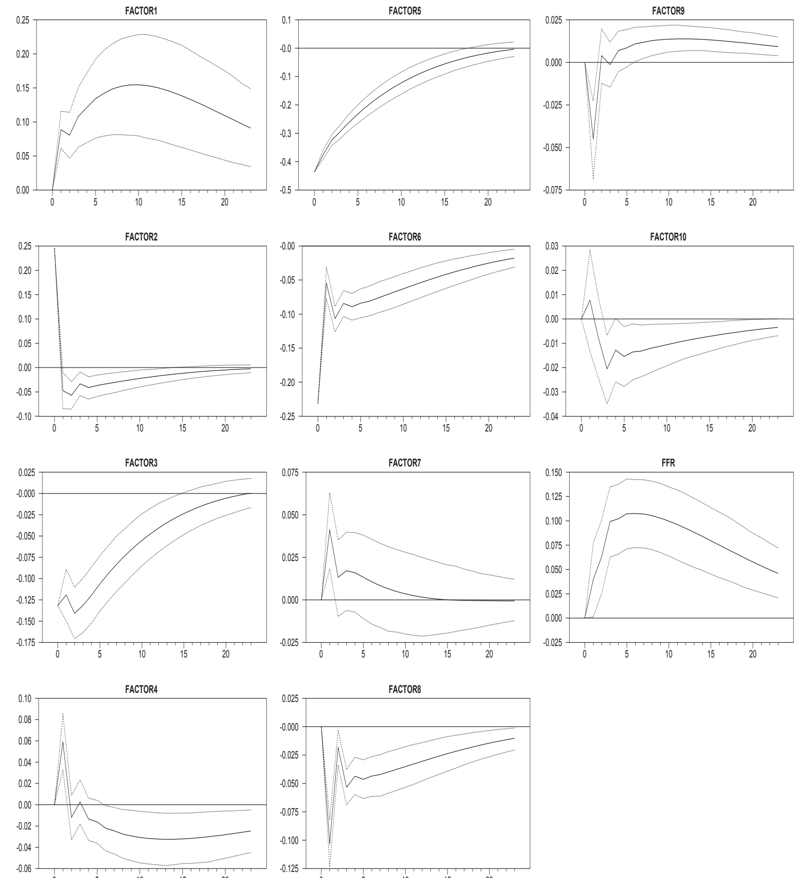
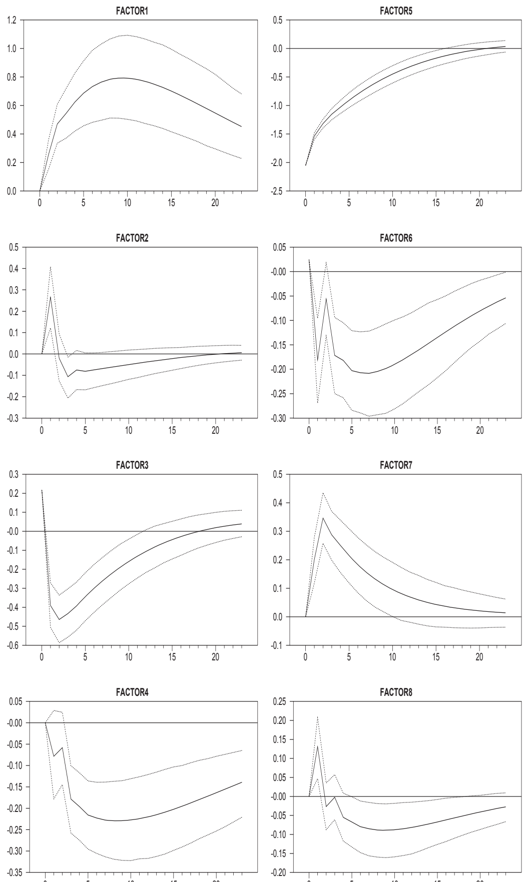


Figure 3.25. Impulse responses to Factor 1 (Employment and Housing starts) shocks

Figure 3.26. Impulse responses to Factor 2 (Price) shocks



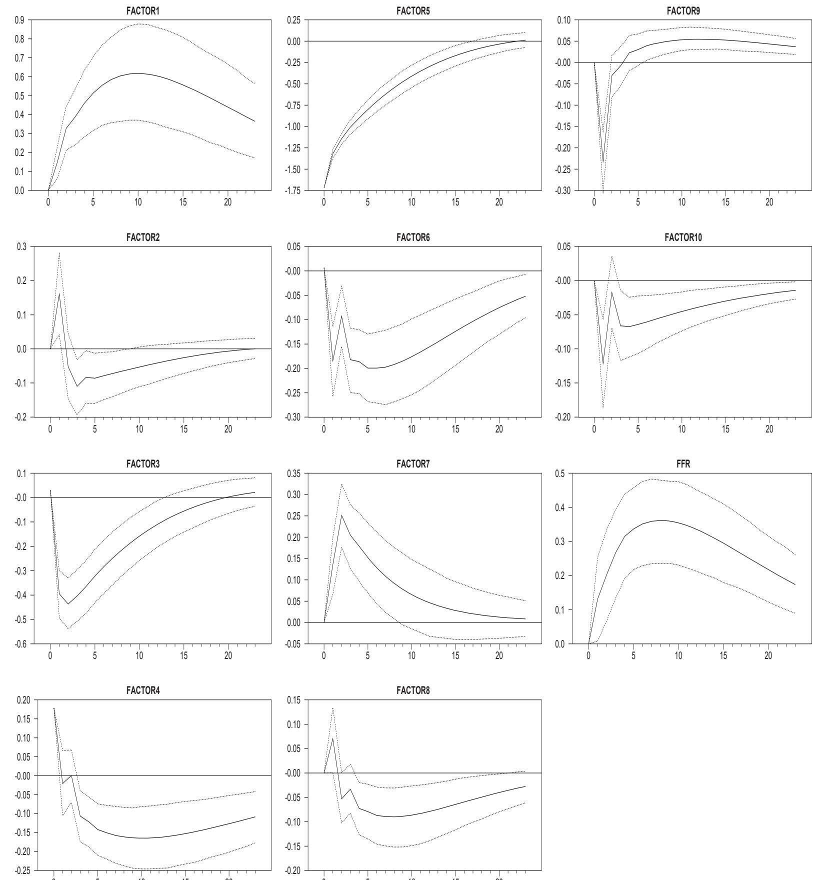


Figure 3.27. Impulse responses to Factor 3 (Interest rate) shocks

Figure 3.28. Impulse responses to Factor 4 (Output) shocks

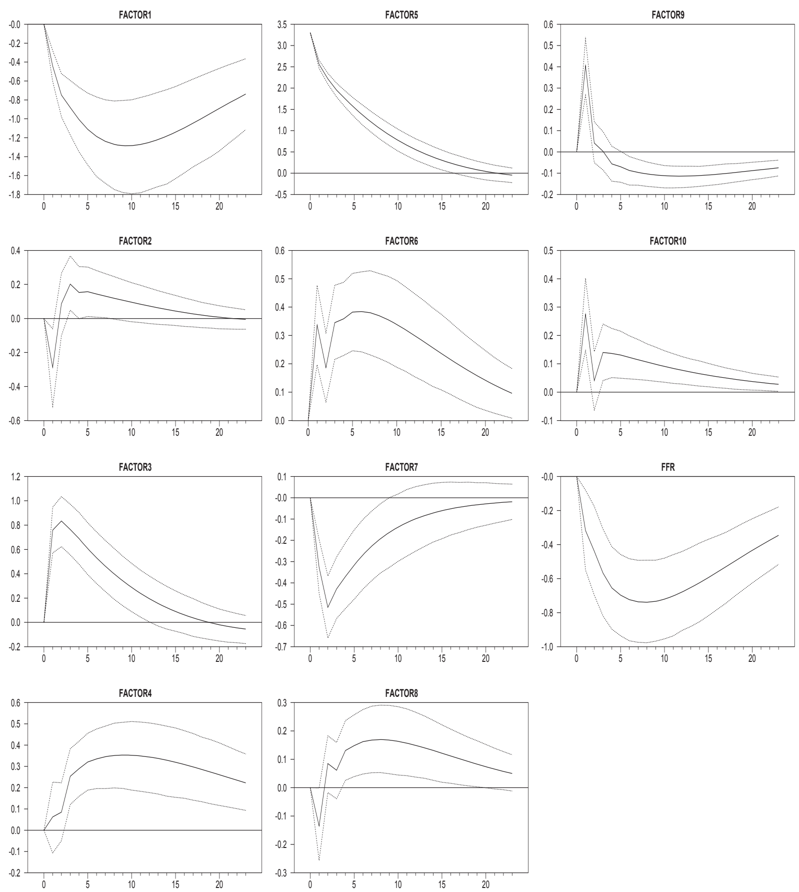


Figure 3.29. Impulse responses to Factor 5 (Spread) shocks

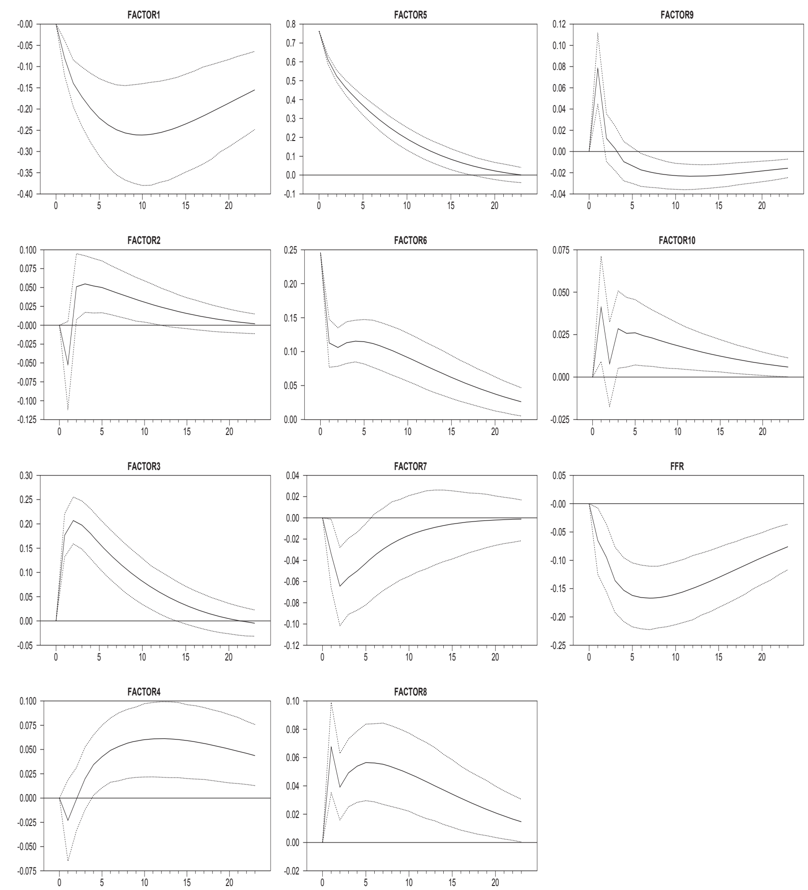


Figure 3.30. Impulse responses to Factor 6 (Interest rate and Money aggregate) shocks

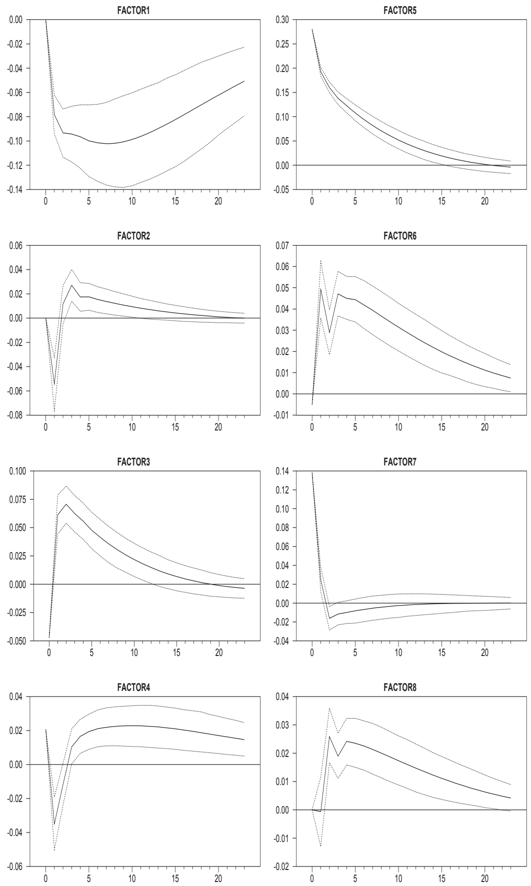


Figure 3.31. Impulse responses to Factor 7 (Stocks) shocks

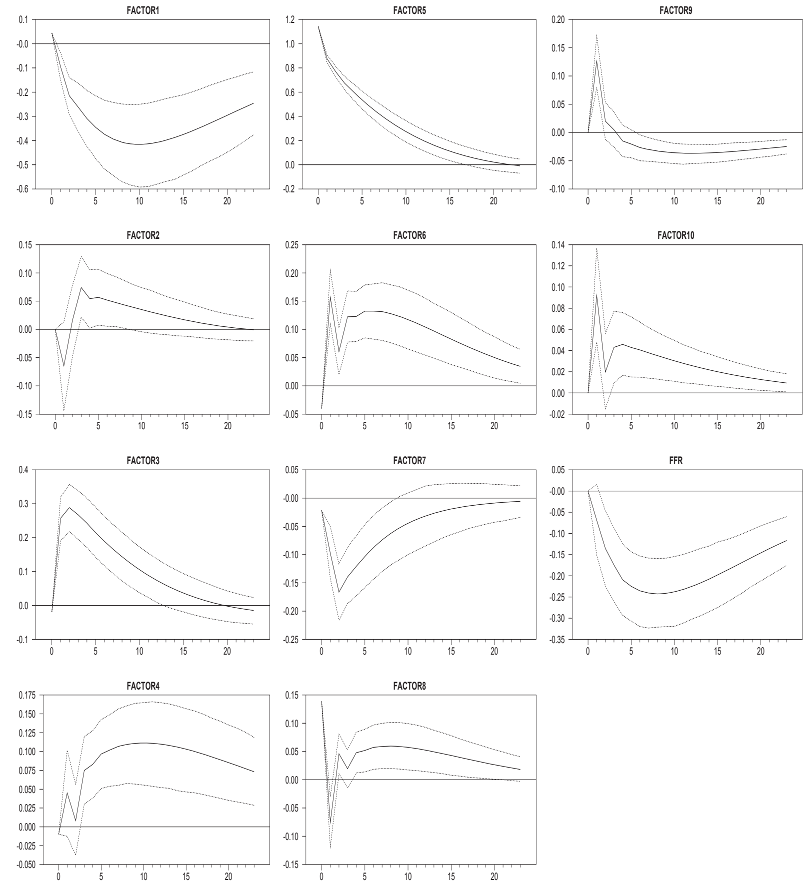
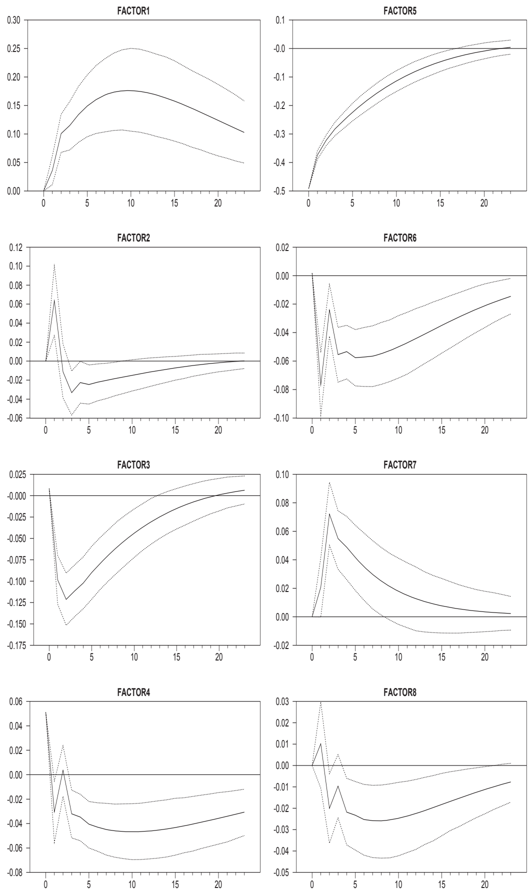


Figure 3.32. Impulse responses to Factor 8 (Average hourly earnings) shocks



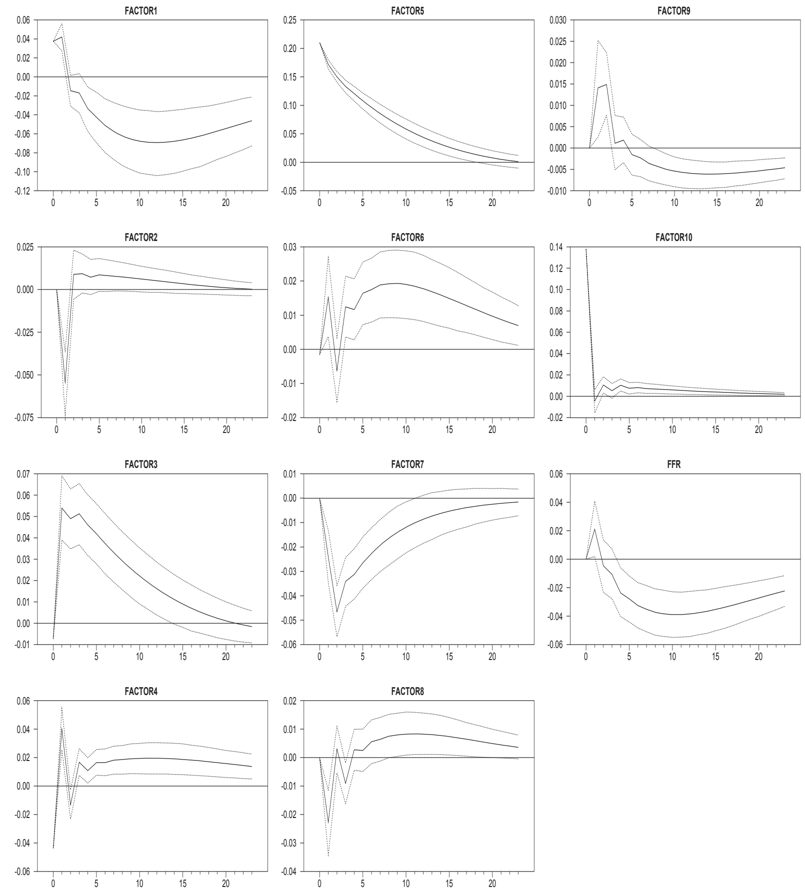


Figure 3.33. Impulse responses to Factor 9 (Real income and Unemployment) shocks

Figure 3.34. Impulse responses to Factor 10 (Exchange rate) shocks

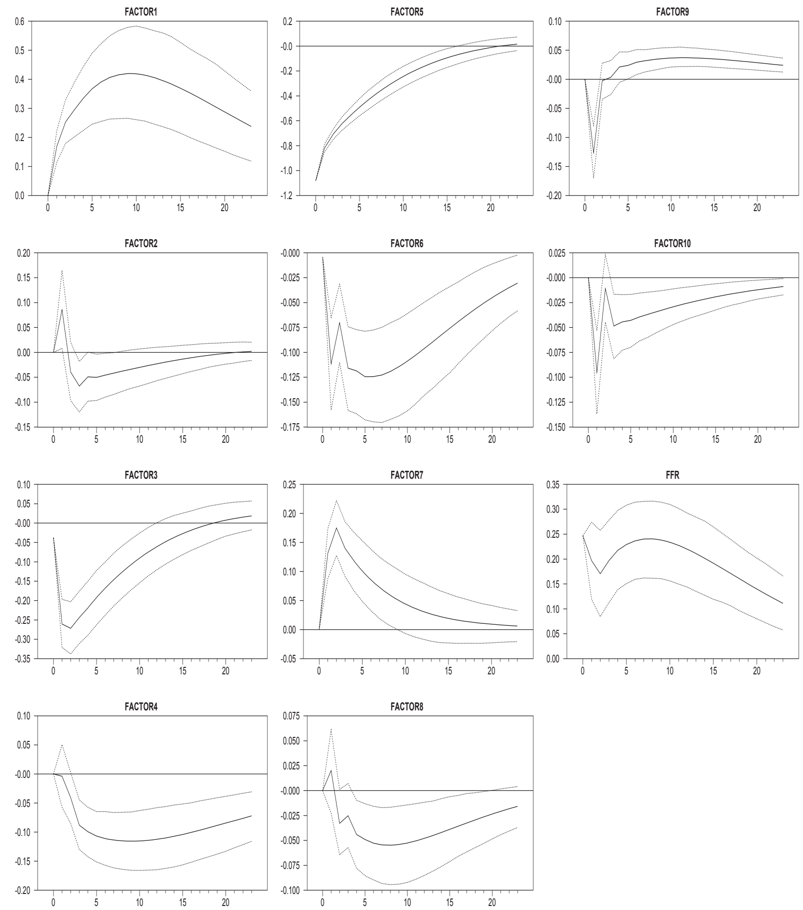
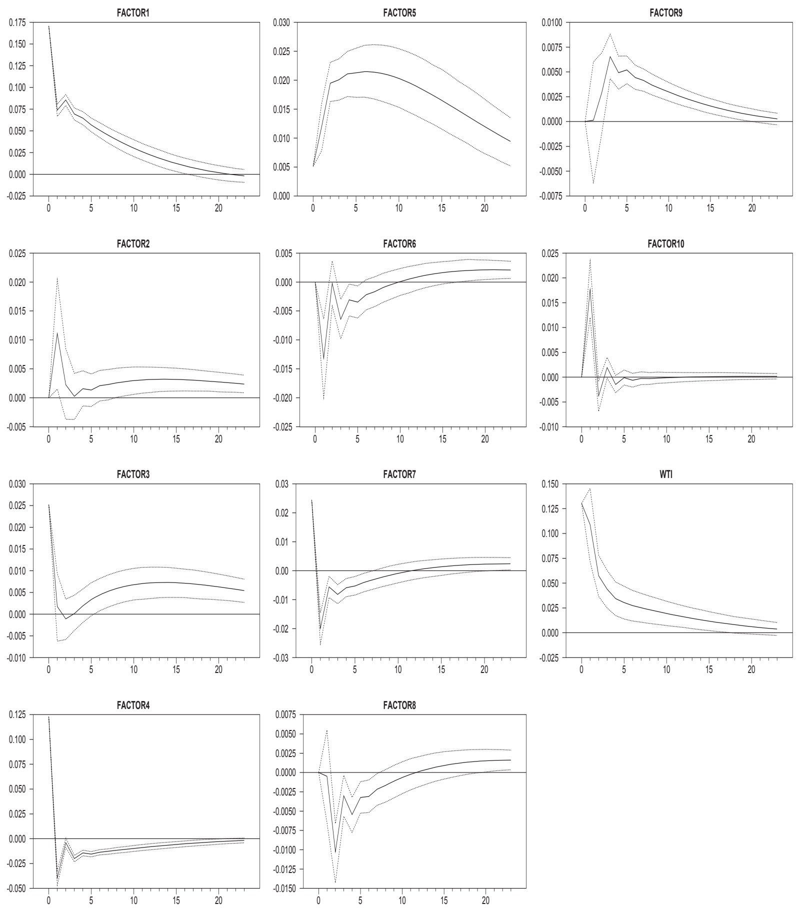


Figure 3.35. Impulse responses to FFR (Federal fund rate) shocks



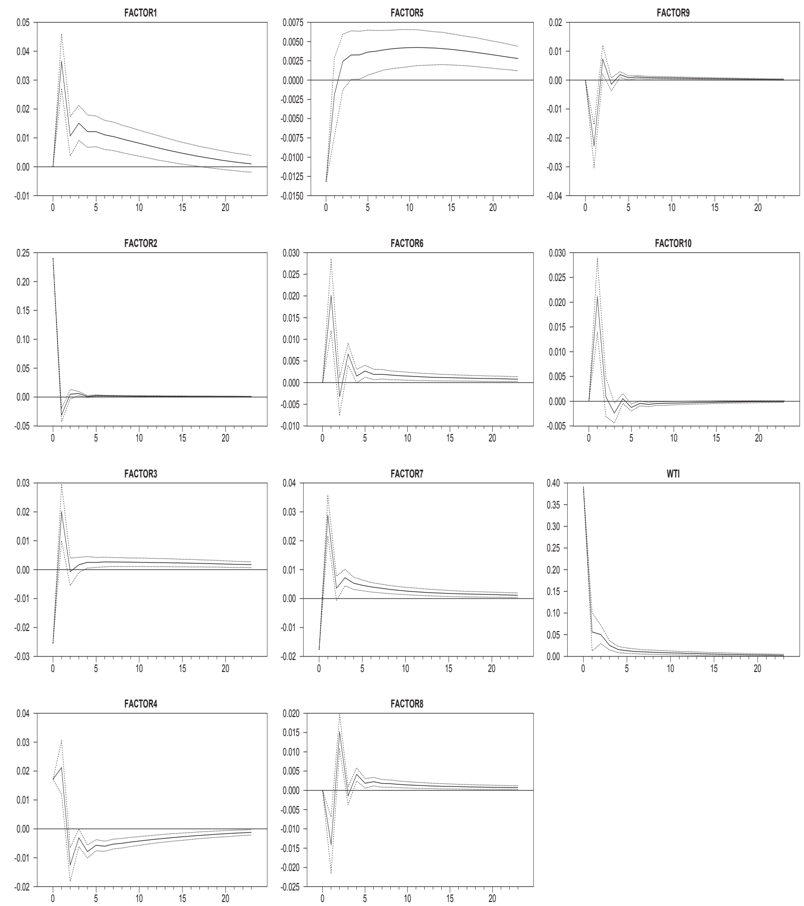


Figure 3.36. Impulse responses to Factor 1 (Employment and Housing starts) shocks

Figure 3.37. Impulse responses to Factor 2 (Price) shocks

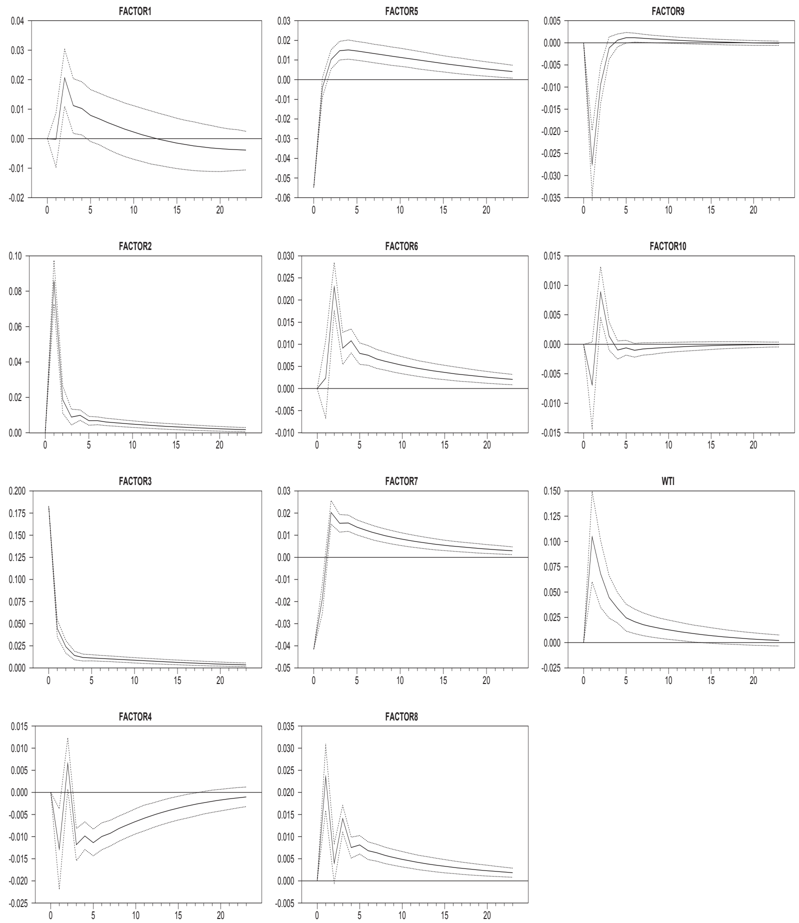


Figure 3.38. Impulse responses to Factor 3 (Interest rate) shocks

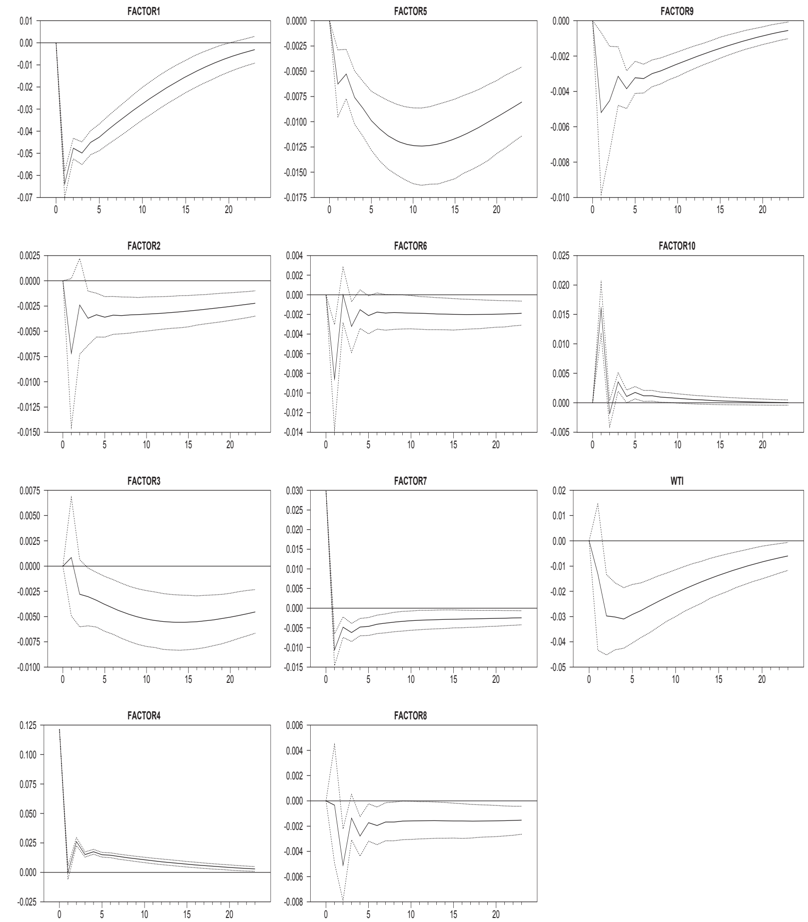


Figure 3.39. Impulse responses to Factor 4 (Output) shocks

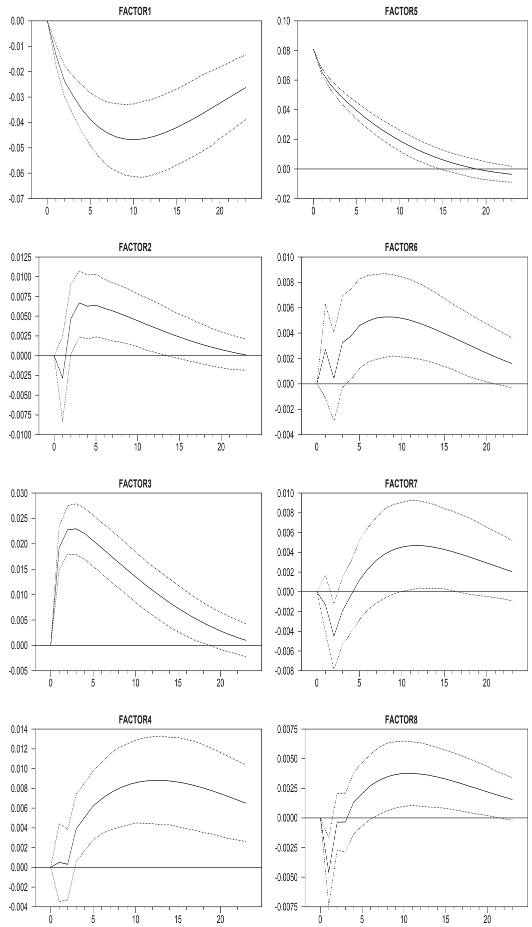


Figure 3.40. Impulse responses to Factor 5 (Spread) shocks

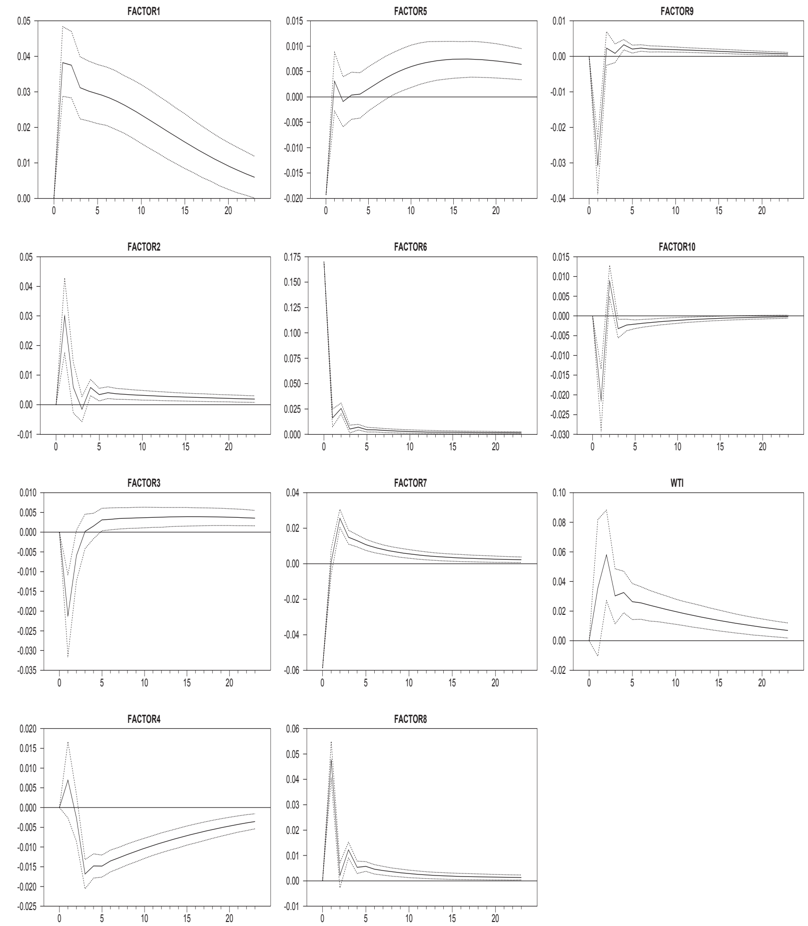


Figure 3.41. Impulse responses to Factor 6 (Interest rate and Money aggregate) shocks

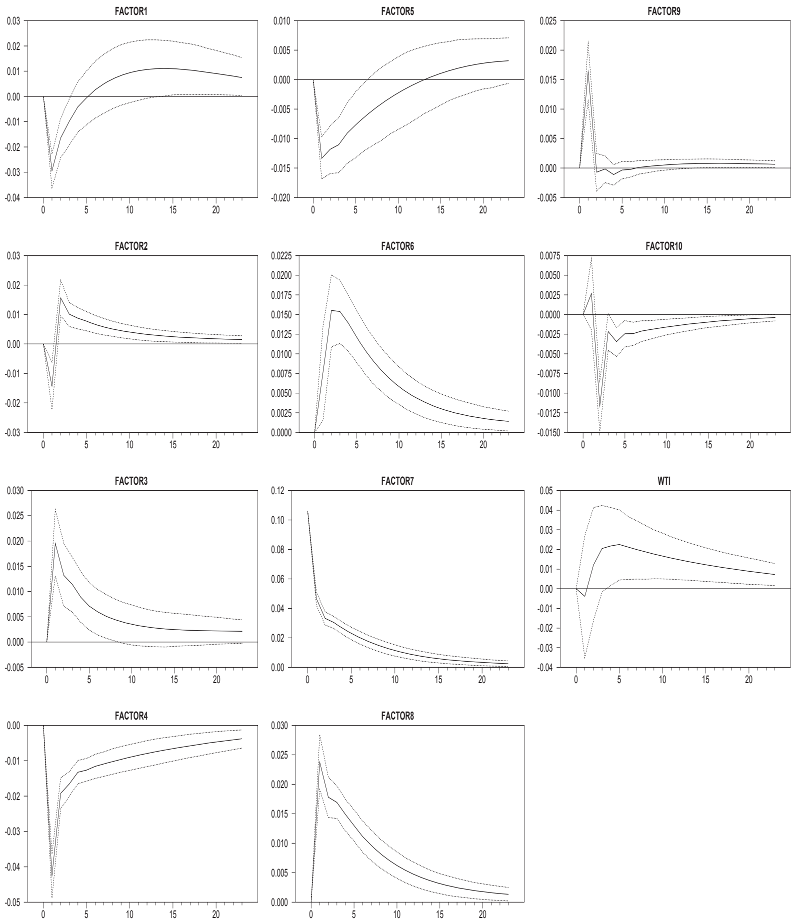


Figure 3.42. Impulse responses to Factor 7 (Stocks) shocks

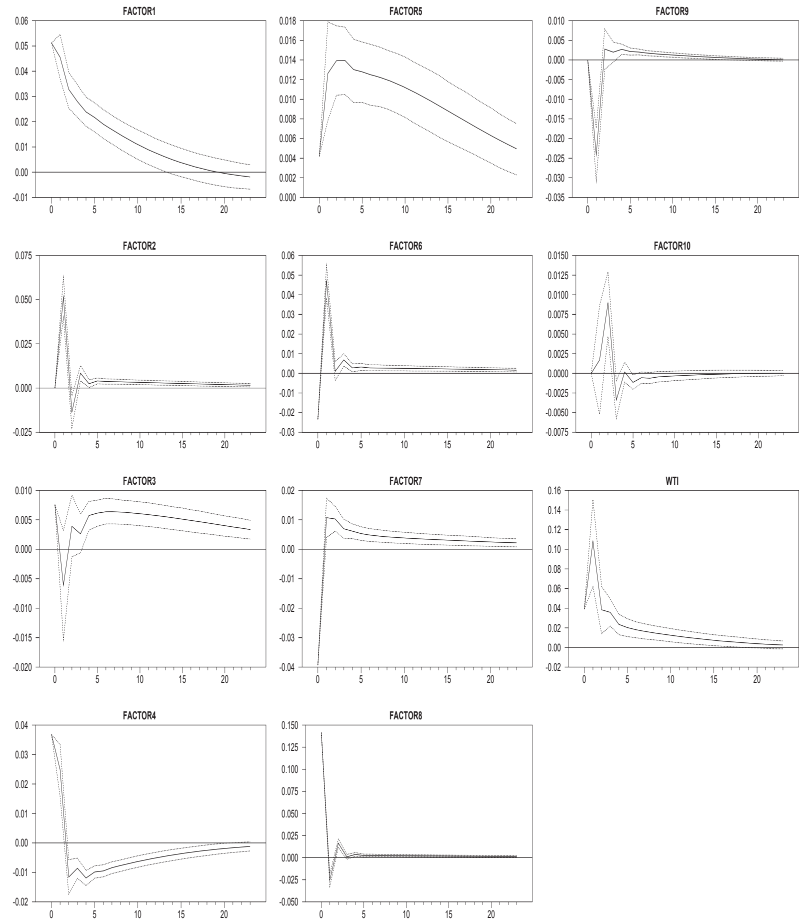


Figure 3.43. Impulse responses to Factor 8 (Average hourly earnings) shocks

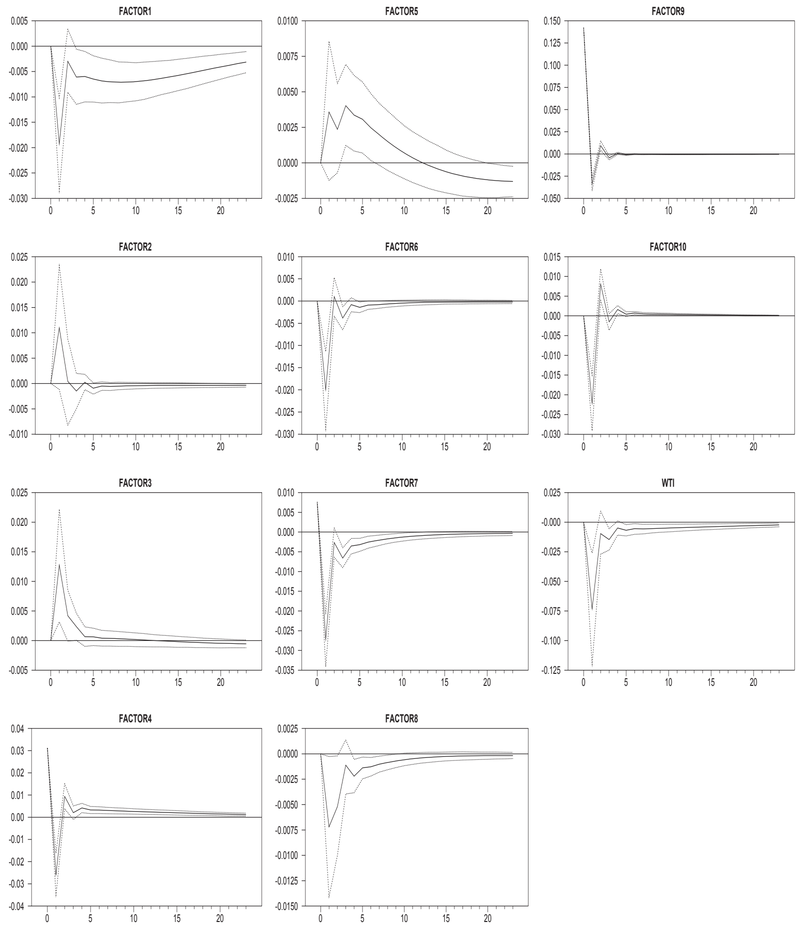


Figure 3.44. Impulse responses to Factor 9 (Real income and Unemployment) shocks

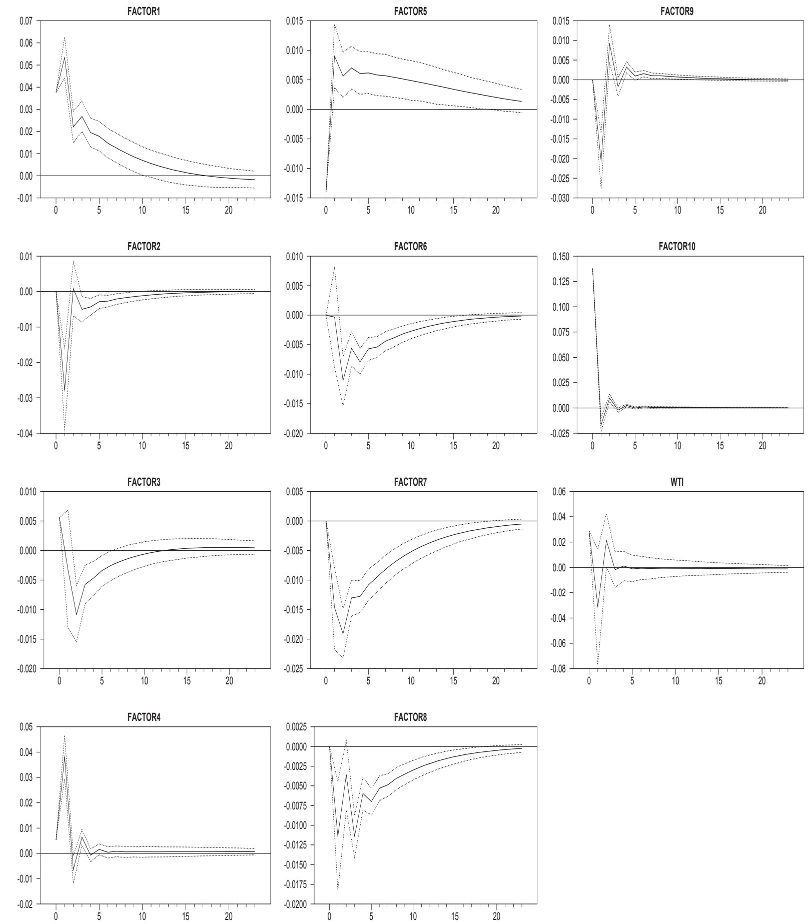


Figure 3.45. Impulse responses to Factor 10 (Exchange rate) shocks

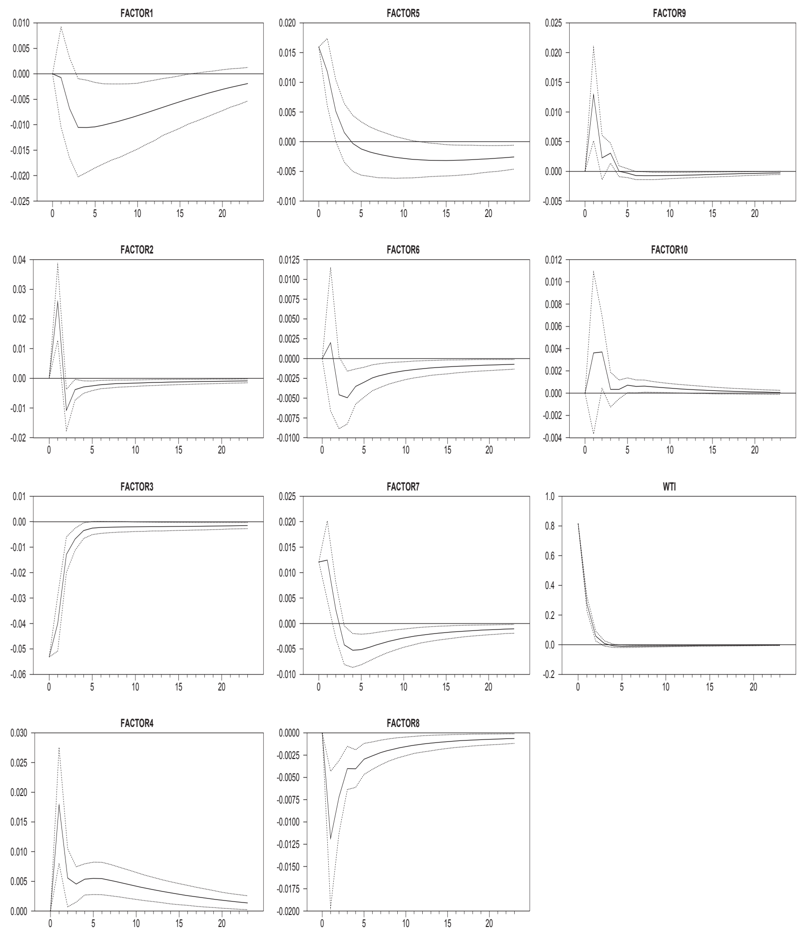


Figure 3.46. Impulse responses to WTI (WTI crude oil price return) shocks

The specific descriptions for the IRFs are: (a) a shock to employment and housing starts (factor 1) (figure 3.25) does not contemporaneously affect the augmented factors and the federal fund rate with the exception of output (factor 4), but it positively affects stock (factor 7), real income and unemployment (factor 9), and the federal fund rate dynamically; (b) a shock to price (factor 2) (figure 3.26) positively affects employment and housing starts (factor 1), stock (factor 7), real income and unemployment (factor 9), and the federal fund rate dynamically, whereas output (factor 4) increases initially but drops after two months, and real income and unemployment (factor 9), which is negatively influenced in the short-run then increases; (c) shocks to interest rate (factor 3) (figure 3.27) and output (factor 4) (figure 3.28) positively affect employment and housing starts (factor 1), stock (factor 7), real income and unemployment (factor 9), and the federal fund rate dynamically, whereas price (factor 2) and average hourly earnings (factor 8) slightly decrease after a big jump in the short-run, and real income and unemployment (factor 9) gradually increases after a drop; (d) a shock to spread (factor 5) (figure 3.29) positively affects price (factor 2), interest rate (factor 3), (output) factor 4, interest rate and money aggregate (factor 6), average hourly earnings (factor 8), and exchange rate (factor 10) dynamically, whereas price (factor 2) and average hourly earnings (factor 8) drop initially, but increase in the long-run; (e) a shock to interest rate and money aggregate (factor 6) (figure 3.30) positively affects spread (factor 5) contemporaneously, whereas employment and housing starts (factor 1), stock (factor 7), real income and unemployment (factor 9), and the federal fund rate decrease dynamically, and the other factors increase; (f) for a shock to stock (factor 7) (figure 3.31), employment and housing starts (factor 1), real income and unemployment (factor 9), and the federal fund rate decrease dynamically, while the other factors increase, whereas price (factor 2) and output (factor 4) have negatively strong influences on stock (factor 7) in the short-run, but then increase; (g) a shock to average hourly earnings (factor 8) (figure 3.32) positively or negatively affects employment and housing starts (factor 1), interest rate (factor 3), output (factor 4), spread (factor 5), interest rate and money aggregate (factor 6), and stock (factor 7) in contemporaneous time. Similar to (e) and (f) cases, for a shock to average hourly earnings (factor 8), employment and housing starts (factor 1), stock (factor 7), real income and unemployment (factor 9), and the federal fund rate decrease dynamically, whereas the other factors increase; (h) for a shock to real income and unemployment (factor 9) (figure 3.33), employment and housing starts (factor 1), stock (factor 7), and the federal fund rate are positively influenced dynamically, whereas these factors are not influenced in contemporaneous time, and the other factors decrease dynamically; (i) a shock to exchange rate (factor 10) (figure 3.34) positively affects employment and housing starts (factor 1) and spread (factor 5) in contemporaneous time, whereas it negatively affects output (factor 4), whereas price (factor 2), interest rate (factor 3), output (factor 4), interest rate and money aggregate (factor 6), and average hourly earnings (factor 8) decrease dynamically; (j) a shock to the federal fund rate (figure 3.35) negatively affects interest rate (factor 3) and spread (factor 5) in contemporaneous time, and employment and housing starts (factor 1), stock (factor 7), and real income and unemployment (factor 9) increase dynamically, whereas the other factors decrease. In particular, price (factor 2) eventually decreases dynamically after a slight jump. From this last result, we cannot observe a strong example of the so-called price puzzle.

For the WTI model, general interpretations of results are that the movements of all augmented factors and the WTI crude oil price return show high volatilities in the short-run but then stabilize. However, we cannot provide conclusive interpretations of the movements of all augmented factors and WTI crude oil price return except the opposite movements of employment and housing starts (factor 1) and output (factor 4) in the real economy sector. The specific descriptions for the IRFs are: (a) a shock to employment and housing starts (factor 1) (figure 3.36) affects interest rate (factor 3), output (factor 4), and spread (factor 5) in contemporaneous time, whereas a shock to employment and housing starts (factor 1) positively affects price (factor 2), interest rate (factor 3), spread (factor 5), real income and unemployment (factor 9), and the WTI crude oil price return dynamically. And, output (factor 4), interest rate and money aggregate (factor 6), stock (factor 7), average hourly earnings (factor 8), and exchange rate (factor 10) reveal higher volatilities in the short-run, but shortly converge to zero; (b) a shock to price (factor 2) (figure 3.37) strongly influences all augmented factors in the short-run and then converges to zero, whereas employment and housing starts (factor 1) slightly increase after a big jump, and spread (factor 5) slightly increases after an initial drop; (c) a shock to interest rate (factor 3) (figure 3.38) negatively affects spread (factor 5) and stock (factor 7) in contemporaneous time, whereas it strongly influences all augmented factors in the short-run, but stabilizes in the long-run; (d) a shock to output (factor 4) (figure 3.39) only positively influences stock (factor 7) in contemporaneous time, and all augmented factors decrease dynamically with the exception of exchange rate (factor 10); (e) a shock to spread (factor 5) (figure 3.40) does not contemporaneously affect the augmented factors, whereas it negatively affects average hourly earnings (factor 9) and the WTI crude oil price return dynamically; (f) a shock to interest rate and money aggregate (factor 6) (figure 3.41) positively affects employment and housing starts (factor 1), spread (factor 5), and WTI crude oil price return dynamically, negatively affects output (factor 4), and slightly influences the other factors which then stabilize; (g) a shock to stock (factor 7) (figure 3.42) strongly affects the augmented factors in the short-run except the WTI crude oil price return, whereas interest rate (factor 3), interest rate and money aggregate (factor 6), average hourly earnings (factor 8), and the WTI crude oil price return increase, but employment and housing starts (factor 1), output (factor 4), spread (factor 5), and exchange rate (factor 10) slightly decrease; (h) a shock to average hourly earnings (factor 8) (figure 3.43) strongly influences the augmented factors in the short-run, but they shortly stabilize, whereas it positively affects employment and housing starts (factor 1) and spread (factor 5) dynamically; (i) a shock to real income and unemployment (factor 9) (figure 3.44) and a shock to exchange rate (factor 10) (figure 3.45) strongly influence the augmented factors in the short-run, but they shortly stabilize; (j) a shock to the WTI crude oil price return (figure 3.46) negatively affects employment and housing starts (factor 1), interest rate (factor 3), interest rate and money aggregate (factor 6), and average hourly earnings (factor 8) dynamically, whereas the other augmented factors reveal higher volatilities in the short-run, but shortly stabilize. These results demonstrate that oil price shock for fluctuations in the WTI crude oil price return are an important source for fluctuations in US macroeconomic and financial indicators in the short-run.

Overall, we consider our results to be satisfactory. However, we note that the above descriptions only provide one possible interpretation to check the models' empirical plausibility, since we lack sufficient information on the complete causal structures among variables of the overall economy over the full dynamic interactions beyond contemporaneous time. Despite this drawback, the IRFs obtained from our FAVAR models generally align with the literature and make economic sense. First, spread (factor 5) shock has a positive effect on output (factor 4). Cuaresma et al. (2004) and Estrella (2005) explain that since spread accounts for future output growth, theoretically, the relationship can be positively or negatively correlated. Second, price (factor 2) shock has a negative effect on output (factor 4) which is supported by Christiano et al. (1999). Third, the positive relationship between interest rate (factor 3) and stock (factor 7) has been advocated by Tufte and Wohar (1999). Fourth, our finding that the federal fund rate shock has a negative effect on interest rate (factor 3), spread (factor 5), interest rate and money aggregate (factor 6), and exchange rate (factor 10) is supported by the FAVAR models in Bernanke et al. (2005)[13] and Kwon (2007)[14]. Although we find that the federal fund rate shock has a positive effect on stock (factor 7), Lagana and Mountford (2005) provided the same results in their FAVAR model applied to a UK dataset. Fifth, the price puzzle is considerably reduced and prices (factor 2)

[13] The FAVAR model in Bernanke et al. (2005) represents the method based on the estimated factors from the entire dataset and the assumed full recursive restrictions. This study is generally accepted as the benchmark of the monetary policy effect (Stock and Watson, 2005).
[14] The FAVAR model in Kwon (2007) represents the method based on the estimated factors from the inductively classified groups of variables and inferred causal structures.

eventually decrease. From the results of our FAVAR model, Bernanke et al. (2005), and Kwon (2007), we conclude that the inclusion of the information captured by the factors into the VAR framework succeeds in mitigating the price puzzle.

3.4.6. Forecast Error Variance Decompositions

Tables 3.7 and 3.8 report the forecast error variance decompositions based on a structural factorization using the contemporaneous information flows from figures 3.23 and 3.24. The table entries show the variation of the 11 selected variables, i.e., the 10 unobserved factors and the federal fund rate/WTI crude oil price return, due to innovations from those variables at the time horizons of contemporaneous time, short-horizon (1 and 2 months) and long-horizon (12 months).

From the results, we note that the percentage of the forecast error variance explained by each innovation does not differ much in both models. Thus, we conclude that there is no dominant innovation for explaining each of the forecast error variances of all of the factors and observed variables. Although the relative importance of each innovation is not interpreted easily in this condition, we offer the following interpretations.

For the FFR model, interest rate (factor 3), output (factor 4), and spread (factor 5) explain the variations of each augmented factor and federal fund rate except in contemporaneous time, thus factor 3, factor 4, and factor 5 innovations appear to be important at long-horizon. The specific descriptions for the forecast error variance decompositions are: (a) innovations of own contribution (federal fund rate), which account for 100%, explain the uncertainty associated with the federal fund rate in contemporaneous time, but own contributions (46.3% and 25.1% at short-horizon, and 6.7% at long-horizon) appear to be less important over time; (b) the innovation of average hourly earnings (factor 8, 6.1%), exchange rate (factor 10, 4.3%), and own contribution (89.6%) in contemporaneous time explain the variation in employment and housing starts (factor 1), but interest rate (factor 3, 20.0%), output (factor 4, 11.4%), and spread (factor 5, 52.5%) innovations appear to be important at long-horizon; (c) for variations in price (factor 2), innovations of own contribution (100% in contemporaneous time, 25.2 and 24.3% at short-horizon and 10.6% at long-horizon), and interest rate (factor 3, 26.6% and 24.4% at short-horizon and 17.5% at long-horizon), output (factor 4, 9.5% and 10.0% at short-horizon and 13.2% at long-horizon) and spread (factor 5, 30.0% and 31.1% at short-horizon and 43.1% at long-horizon) appear to be important in the whole time horizon; (d) innovations of price (factor 2, 24.9%) and own contribution (67.8%) explain the variance of interest rate (factor 3) in contemporaneous time, whereas output (factor 4, 14.3%), spread (factor 5, 50.9%), and own contribution (16.3%) explain it at long-horizon; thus, own contribution (interest rate), factor 4 (output), and factor 5 (spread) innovations appear to be important for explaining the interest rate at long-horizon, with the exception of factor 2 (1.8%); (e) innovations of real income and unemployment (factor 9, 7.1%), exchange rate (factor 10, 5.1%), and own contribution (86.2%) show high explanatory power for the variance of output (factor 4) in contemporaneous time, whereas own contribution (11.8%) still remains dominant, although interest rate (factor 3, 22.7%) and spread (factor 5, 51.2%)

appear to be important factors at long-horizon; (f) innovations of interest rate (factor 3), output (factor 4), and own contribution explain the variation in spread (factor 5) at any time-horizon; (g) for the variation in interest rate and money aggregate (factor 6), price (factor 2, 46.0%) and own contribution (52.0%) appear to be important in contemporaneous time, whereas interest rate (factor 3), output (factor 4), and spread (factor 5) innovations show high ranked explanatory power as time passes, and price (factor 2) innovation and own contribution show diminished explanatory power as time passes; (h) for the variation in stock (factor 7), innovations of average hourly earnings (factor 8, 2.4%) and own contribution (97.6%) appear to be important in contemporaneous time, whereas interest rate (factor 3, 23.2%), output (factor 4, 11.5%), and spread (factor 5, 51.1%) innovations are important at long-horizon; (i) own contribution only explains the variance of average hourly earnings (factor 8) in contemporaneous time, whereas interest rate (factor 3, 13.3%), output (factor 4, 13.0%), and spread (factor 5, 44.9%) innovations as well as own contribution (10.1%) appear to be important at long-horizon; (j) innovations of own contribution (100%) explain the uncertainty associated with real income and unemployment (factor 9) in contemporaneous time, whereas interest rate (factor 3, 20.7%), output (factor 4, 13.7%), and spread (factor 5, 46.7%) innovations as well as own contribution (6.0%) show strong explanatory power at long-horizon; (k) innovations of own contribution (100%) explain the variance of exchange rate (factor 10) in contemporaneous time, whereas interest rate (factor 3, 19.5%), output (factor 4, 11.0%), and spread (factor 5, 49.7%) innovations appear to be more important at long-horizon.

For the WTI model, innovations of own contribution appear to be important at any time-horizon for explaining the uncertainties of the augmented factors and the WTI crude oil price return. The specific descriptions are: (a) innovations of price (factor 2, 18.3%) and own contribution (79.4%) in contemporaneous time, and 15.7% and 73.1%, respectively, at long-horizon explain the uncertainty associated with the WTI crude oil price return for whole time-horizon; (b) average hourly earnings (factor 8, 7.9%), exchange rate (factor 10, 4.3%), and own contribution (87.8%) innovations explain the variance of employment and housing starts (factor 1) in contemporaneous time, whereas innovations of output (factor 4, 15.1%), spread (factor 5, 11.8%), interest rate and money aggregate (factor 6, 7.8%), and own contribution (48.2%) appear to be important at long-horizon; (c) own contribution (100%) explains the variance of price (factor 2) in contemporaneous time, whereas own contribution (82.1% and 80.9% at short-horizon and 78.7% at long-horizon) and interest rate (factor 3, 10.3% and 10.8% at short-horizon and 11.1% at long-horizon) innovations appear to be important; (d) for the variation in interest rate (factor 3), own contribution (88.8% in contemporaneous time, 82.2% and 80.4% at short-horizon and 72.7% at long-horizon) and WTI crude oil price return (7.5% in contemporaneous time, 10.4% at short-horizon and 9.3% at long-horizon) innovations show high explanatory power for whole time-horizon; (e) for the variation in output (factor 4), innovations of employment and housing starts (factor 1, 46.4% in contemporaneous time, 41.8% and 40.3% at short-horizon and 36.1% at long-horizon) and own contribution (45.4% in contemporaneous time, 37.3% and 37.5% at short-horizon and 34.1% at long-horizon) appear to be important for whole time-horizon; (f) innovations of interest rate (factor 3, 28.4% in contemporaneous time, 19.3% and 15.6% at short-horizon and 11.9% at long-horizon) and own contribution (61.7% in contemporaneous time, 69.2% and 71.3% at short-horizon and 63.5% at long-horizon) explain the variance of spread (factor 5); (g) for the variation of interest rate and money aggregate (factor 6), innovations of average hourly earnings (factor 8, 1.9% in contemporaneous time, 8.3% and 8.0% at short-horizon and 7.7% at long-horizon) and own contribution (98.1% in contemporaneous time, 88.3% and 86.2% at short-horizon and 80.0% at long-horizon) show high ranked explanatory power; (h) for the variation of stock (factor 7), interest rate (factor 3, 8.6% in contemporaneous time, 8.3% and 8.8% at short-horizon and 10.7% at long-horizon), interest rate and money aggregate (factor 6, 17.2% in contemporaneous time, 13.7% and 14.7% at short-horizon and 13.8% at long-horizon), average hourly earnings (factor 8, 7.7% in contemporaneous time, 6.6% and 6.3% at short-horizon, and 5.5% at long-horizon), and own contribution (56.4% in contemporaneous time, 53.8% and 52.6% at short-horizon and 51.6% at long-horizon) innovations show high ranked explanatory power; (i) innovations of own contribution (100%) explains variance of average hourly earnings (factor 8) in contemporaneous time, whereas interest rate and money aggregate (factor 6, 9.4% and 9.0% at short-horizon and 9.3% at long-horizon), stock (factor 7, 2.3% and 3.6% at short-horizon and 7.8% at long-horizon), and own contribution (83.9% and 81.3% at short-horizon and 72.9% at long-horizon) innovations appear to be important; (j) own contribution (100% at contemporaneous time, 85.2% and 84.3% at short-horizon and 83.0% at long-horizon) explain the variations in real income and unemployment (factor 9) for whole time-horizon; (k) innovations of own contribution (100% in contemporaneous time, 90.3% and 88.5% at short-horizon and 87.6% at long-horizon) explain uncertainty associated with exchange rate (factor 10).

We interpret the overall results as follows: (a) for the FFR model, interest rate (factor 3), output (factor 4), and spread (factor 5) innovations appear to be dominant for explaining each of the forecast error variances, and the federal fund rate innovation is important for explaining each of the forecast error variances of almost all augmented factors except in contemporaneous time; (b) for the WTI model, no dominant innovation explains each of the forecast error variances, yet the innovations of own contribution appear to be important for the variations in each augmented factor and the WTI crude oil price return at any time-horizon; (c) innovations of own contribution only explain the variances of federal fund rate and WTI crude oil price return in contemporaneous time. These results suggest that the federal fund rate and the WTI crude oil price return appear to be exogenous in contemporaneous time. We note that they are consistent with our previous DAG patterns which inductively inferred the contemporaneous causal structure by the GES algorithm without any deductive information in figures 3.23 and 3.24.

Table 3.7
Forecast error variance decomposition for the FFR model

Period	FFR	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Forecast error variance decomposition for the federal fund rate shock											
0 month	100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 month	46.271	1.264	0.417	7.295	5.249	34.186	1.178	1.739	1.256	0.847	0.299
2 month	25.092	1.612	0.753	11.433	8.484	45.159	1.793	1.415	3.296	0.832	0.13
12 month	6.706	0.492	1.104	18.046	12.086	51.94	2.627	0.505	5.382	1.002	0.11
Forecast error variance decomposition for Factor 1											
0 month	0.000	89.575	0.000	0.000	0.000	0.000	0.000	0.000	6.105	0.000	4.320
1 month	7.235	13.934	1.988	16.185	5.156	48.413	1.56	1.655	2.648	0.293	0.932
2 month	6.382	5.47	0.986	19.492	8.583	51.45	1.721	1.088	3.807	0.76	0.261
12 month	5.715	0.845	0.758	20.045	11.417	52.502	2.066	0.418	5.172	0.949	0.114
Forecast error variance decomposition for Factor 2											
0 month	0.000	0.000	100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 month	2.585	0.014	25.157	26.555	9.524	29.998	0.925	1.133	1.424	1.545	1.141
2 month	3.094	0.014	24.272	24.366	10.001	31.103	2.022	1.084	1.498	1.475	1.070
12 month	4.437	0.016	10.599	17.521	13.152	43.138	3.826	0.795	4.706	1.316	0.494

Table 3.7
(Continued)

Period	FFR	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Forecast error variance decomposition for Factor 3											
0 month	2.042	0.003	24.884	67.835	1.278	0.000	0.000	3.217	0.559	0.105	0.076
1 month	6.054	0.028	2.727	17.075	13.664	50.196	2.801	0.503	5.823	0.871	0.258
2 month	5.763	0.025	2.056	16.456	13.945	51.044	3.045	0.431	6.026	0.995	0.214
12 month	5.335	0.014	1.756	16.288	14.277	50.941	3.436	0.359	6.278	1.083	0.232
Forecast error variance decomposition for Factor 4											
0 month	0.000	0.223	0.000	0.000	86.162	0.000	0.000	1.160	0.251	7.059	5.145
1 month	0.104	2.770	4.973	11.750	52.156	9.321	0.613	2.586	4.110	5.860	5.757
2 month	2.626	3.347	4.031	14.840	40.334	19.654	0.477	2.162	3.400	4.526	4.602
12 month	5.625	0.717	0.387	22.686	11.826	51.175	1.11	0.269	4.869	1.012	0.324
Forecast error variance decomposition for Factor 5											
0 month	5.386	0.035	0.876	19.423	13.596	50.289	2.684	0.363	6.031	1.114	0.204
1 month	5.331	0.036	0.955	19.056	13.585	50.601	2.768	0.335	6.016	1.103	0.214
2 month	5.304	0.036	0.985	18.851	13.590	50.767	2.787	0.320	6.044	1.098	0.218
12 month	5.257	0.027	1.065	18.381	13.694	50.989	2.86	0.289	6.114	1.090	0.234

Table 3.7
(Continued)

Period	FFR	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Forecast error variance decomposition for Factor 6											
0 month	0.016	0.000	46.027	0.523	0.035	0.000	51.997	0.022	1.375	0.003	0.002
1 month	3.240	0.205	17.188	8.668	9.120	29.869	22.057	0.685	7.279	1.634	0.056
2 month	3.710	0.276	17.068	7.685	9.319	31.815	21.014	0.769	6.806	1.477	0.060
12 month	5.149	0.178	4.551	13.553	13.23	48.192	6.939	0.659	6.244	1.201	0.103
Forecast error variance decomposition for Factor 7											
0 month	0.000	0.000	0.000	0.000	0.000	0.000	0.000	97.617	2.383	0.000	0.000
1 month	7.906	0.000	0.761	18.843	8.539	49.311	0.560	8.934	4.669	0.209	0.269
2 month	6.494	0.019	0.247	21.874	11.082	50.635	0.720	2.637	5.152	0.764	0.376
12 month	5.731	0.031	0.129	23.163	11.528	51.145	0.799	0.935	5.338	0.832	0.368
Forecast error variance decomposition for Factor 8											
0 month	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100.000	0.000	0.000
1 month	0.529	0.012	13.025	21.253	6.118	22.443	5.687	0.000	30.193	0.120	0.620
2 month	1.301	0.171	11.610	18.676	7.554	24.744	6.424	0.660	27.865	0.458	0.539
12 month	4.552	0.194	5.479	13.272	12.987	44.916	6.320	0.922	10.119	1.060	0.179

Table 3.7
(Continued)

Period	FFR	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Forecast error variance decomposition for Factor 9											
0 month	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100	0.000
1 month	4.367	0.001	0.551	20.915	14.476	44.819	1.663	0.887	4.381	7.887	0.052
2 month	4.299	0.017	0.545	21.008	14.551	44.720	1.681	0.876	4.437	7.752	0.113
12 month	4.692	0.248	0.613	20.682	13.728	46.748	1.798	0.769	4.616	6.000	0.107
Forecast error variance decomposition for Factor 10											
0 month	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100
1 month	5.475	0.010	0.039	19.049	8.881	45.911	1.011	0.632	5.107	2.273	11.612
2 month	5.438	0.037	0.068	18.877	8.893	46.072	1.024	0.692	5.266	2.225	11.408
12 month	5.518	0.089	0.443	19.525	11.047	49.706	1.629	0.396	5.481	1.492	4.672

Table 3.8
Forecast error variance decomposition for the WTI model

Period	WTI	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Forecast error variance decomposition for WTI crude oil price shock											
0 month	79.429	2.023	18.267	0.000	0.000	0.000	0.000	0.000	0.182	0.000	0.099
1 month	77.441	3.032	16.241	1.103	0.015	0.007	0.127	0.001	1.289	0.547	0.197
2 month	76.135	3.325	16.157	1.554	0.106	0.020	0.476	0.017	1.427	0.546	0.236
12 month	73.132	3.993	15.681	2.091	0.803	0.165	1.120	0.442	1.774	0.572	0.228
Forecast error variance decomposition for Factor 1											
0 month	0.000	87.820	0.000	0.000	0.000	0.000	0.000	0.000	7.894	0.000	4.286
1 month	0.002	66.732	2.470	0.000	7.800	0.295	2.850	1.681	9.062	0.701	8.408
2 month	0.072	63.588	2.122	0.631	9.682	0.976	4.456	1.767	8.759	0.562	7.384
12 month	0.715	48.173	1.871	0.696	15.144	11.765	7.790	1.235	6.820	0.577	5.215
Forecast error variance decomposition for Factor 2											
0 month	0.000	0.000	100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
1 month	0.994	0.181	82.056	10.337	0.068	0.013	1.235	0.257	3.610	0.175	1.074
2 month	1.128	0.187	80.912	10.758	0.076	0.044	1.280	0.616	3.768	0.173	1.059
12 month	1.177	0.248	78.690	11.149	0.244	0.493	1.446	1.309	3.916	0.175	1.153

Table 3.8
(Continued)

Period	WTI	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Forecast error variance decomposition for Factor 3											
0 month	7.508	1.695	1.727	88.835	0.000	0.000	0.000	0.000	0.152	0.000	0.083
1 month	10.403	1.478	2.442	82.227	0.001	0.890	0.986	0.922	0.197	0.363	0.091
2 month	10.407	1.422	2.343	80.365	0.019	2.088	1.027	1.342	0.230	0.388	0.369
12 month	9.315	1.676	2.179	72.720	0.374	8.658	1.101	2.185	0.903	0.353	0.535
Forecast error variance decomposition for Factor 4											
0 month	0.000	46.429	0.912	0.000	45.398	0.000	0.000	0.000	4.173	2.999	0.089
1 month	0.777	41.825	1.882	0.394	37.321	0.001	0.125	4.582	5.052	4.108	3.933
2 month	0.836	40.268	2.153	0.470	37.535	0.002	0.134	5.422	5.176	4.143	3.861
12 month	1.164	36.091	2.297	2.091	34.124	0.879	3.581	7.319	5.675	3.546	3.233
Forecast error variance decomposition for Factor 5											
0 month	2.404	0.246	1.655	28.444	0.000	61.701	3.535	0.000	0.166	0.000	1.849
1 month	2.531	1.051	1.145	19.299	0.213	69.237	2.415	1.161	1.122	0.080	1.744
2 month	2.13	2.667	0.926	15.645	0.277	71.309	1.904	1.673	1.836	0.094	1.538
12 month	1.14	10.903	0.785	11.916	2.580	63.488	1.249	1.844	4.410	0.191	1.493

Table 3.8
(Continued)

Period	WTI	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Forecast error variance decomposition for Factor 6											
0 month	0.000	0.000	0.000	0.000	0.000	0.000	98.146	0.000	1.854	0.000	0.000
1 month	0.012	0.493	1.203	0.011	0.242	0.022	88.333	0.196	8.342	1.144	0.001
2 month	0.070	0.470	1.168	1.522	0.232	0.022	86.155	0.954	7.962	1.090	0.355
12 month	0.285	0.598	1.302	3.047	0.376	0.600	80.033	3.996	7.711	1.066	0.987
Forecast error variance decomposition for Factor 7											
0 month	0.729	2.990	1.568	8.630	4.456	0.000	17.227	56.377	7.729	0.294	0.000
1 month	1.238	3.979	4.404	8.279	3.983	0.010	13.680	53.811	6.568	3.199	0.849
2 month	1.149	3.705	4.016	8.824	3.668	0.099	14.655	52.587	6.258	2.907	2.134
12 month	1.343	3.326	3.588	10.672	3.344	0.335	13.843	51.561	5.496	2.495	3.997
Forecast error variance decomposition for Factor 8											
0 month	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100	0.000	0.000
1 month	0.603	0.000	0.774	2.239	0.000	0.077	9.387	2.337	83.851	0.215	0.516
2 month	0.792	0.415	1.637	2.220	0.126	0.074	9.025	3.561	81.286	0.318	0.544
12 month	0.973	0.593	1.635	4.137	0.276	0.359	9.271	7.809	72.897	0.342	1.709

Table 3.8
(Continued)

Period	WTI	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Forecast error variance decomposition for Factor 9											
0 month	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100	0.000
1 month	0.651	0.000	2.013	2.875	0.104	0.268	3.769	1.054	2.306	85.190	1.770
2 month	0.668	0.024	2.161	3.224	0.184	0.284	3.726	1.040	2.295	84.340	2.054
12 month	0.707	0.707	2.165	3.203	0.495	0.371	3.821	1.034	2.375	83.021	2.099
Forecast error variance decomposition for Factor 10											
0 month	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	100
1 month	0.071	1.540	2.055	0.218	1.204	0.130	2.107	0.022	0.011	2.311	90.331
2 month	0.142	1.550	2.010	0.551	1.185	0.137	2.362	0.684	0.385	2.531	88.465
12 month	0.155	1.562	2.028	0.581	1.278	0.375	2.497	0.954	0.442	2.531	87.597

3.4.7. Robustness Check

We employ a common framework for generating pseudo out-of-sample-forecasts from the FAVAR and benchmark models. Initially, we use observations from January 1982 to December 2008 and calculate the one-step-ahead predictions to estimate each forecasting model described below. Thereafter, we augment the sample by one month, re-compute the unobserved factors, re-specify the models, and re-estimate their parameters and the corresponding one-step-ahead predictions generated by moving the forecast window forward. We repeat this procedure until we reach the November 2011 end-date, at which point we make our final set of forecasts. The result is a combined total of 35 out-of-sample predictions for the WTI oil price return.

3.4.7.1. Specifications of Forecasting Models

To ensure the robustness of our analysis, first we estimate FAVAR models for both the federal fund rate and the WTI crude oil price return by extracting the factors from a large set of predictors (see Section 3.2). Second, we use simple univariate AR models for each observed variable. Third, we compare the forecasting accuracies of both FAVAR models to the benchmark AR models.

Following Stock and Watson (2002a), the forecasting equation to predict y_t is:

$$y_{t+h}^{FAVAR} = \alpha + \sum_{j=1}^{m} \beta_i \hat{F}_{t+h-j} + \sum_{i=1}^{p} \gamma_i y_{t+h-i}^{FAVAR} \tag{3.15}$$

where y_{t+h}^{FAVAR} denotes the h-steps ahead forecasts of y_t, $\hat{F}_{t+h}$ is the h-steps ahead prediction for the r factors, β_i is $1 \times r$ vector of coefficients, and the m, p, and r are either fixed or selected with information criteria.

Next, we construct our benchmark model. For variable y (WTI crude oil price return), we estimate the following AR (p) model:

$$y_t^{AR} = \alpha^{AR} + \sum_{i=1}^{p} \beta_i^{AR} y_{t-i}^{AR} + v_t \tag{3.16}$$

where α^{AR} and β_i^{AR} are the coefficients to be estimated, and v_t is the residual.

For the h-steps ahead horizon, we determine the forecasts as:

$$y_{t+h}^{AR} = \alpha^{AR} + \sum_{i=1}^{p} \beta_i^{AR} y_{t+h-i}^{AR}$$

where y_{t+h}^{AR} denotes the forecast value of y for horizon $t + h$, and equation (3.16) estimates the coefficients $\alpha^{AR}$and β_i^{AR}.

3.4.7.2. Forecasting Performances

To compare the models' forecast performances, we calculate the out-of-sample-forecast residuals which are generated from the difference between the forecasted and the actual returns prices of each observed variable. Based upon these exercise results, we use two forecast performance measure statistics: Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). Table 3.9 summarizes the forecast performances.

For the two FAVAR models, we note that the MSEs (0.00 and 0.93) and the MAPEs (2.33% and 73.23% for each observed variable) are less than the AR benchmark models. These results confirm our expectation that the MSEs of the FAVAR models are superior.

Table 3.9
Summary of out-of-sample forecast performance of FAVAR and AR models

	MSE		MAPE	
	FAVAR	AR	FAVAR	AR
R_{FFR}	0.000983	0.002859	2.33097	3.090269
R_{WTI}	0.927359	1.073899	73.23035	79.93896

Note: $\text{MSE} = \frac{1}{T}\sum_{t=1}^{T}(Y_t - \hat{Y}_t)^2$ and $\text{MAPE} = \frac{100}{T}\sum_{t=1}^{T}|Y_t - \hat{Y}_t|$; a lower loss measure indicates a higher forecasting power.

However, these two performance measures do not provide a statistical significance of the similarities/differences between the models' forecasts. Therefore, we apply the DM test (Diebold and Mariano 1995) and the forecast encompassing test (Harvey et al. 1997) to examine the out-of-sample predictability. Table 3.10 summarizes the test statistics.

From the DM test statistics, the hypothesis of equality of forecast errors between the models cannot be rejected in all variables at the 5% significance level. We conclude that the forecasting errors of the FAVAR and AR models statistically perform similarly.

Table 3.10
Summary of test statistics of DM and out-of-sample forecast encompassing test for FAVAR and AR models

	DM		Forecast Encompassing			
			Dependent variable			
			e_t^{FAVAR}		e_t^{AR}	
	test statistics	p-value	$\hat{\lambda}$	p-value	$\hat{\lambda}$	p-value
R_{FFR}	-1.4038	0.1604	-0.3225	0.038	1.3225	0.000
R_{WTI}	-1.2963	0.1949	-0.0613	0.894	1.0613	0.026

Note: The DM test and forecast encompassing test are based on the null hypothesis of no difference in the accuracy (equal predictive ability) between the forecasting errors of the FAVAR and AR models.

In the DM test, the null hypothesis of equal forecast accuracy is tested based on $E(d_t) = 0$, where E is expectation operator and $d_t = e^2_{FAVAR,t} - e^2_{AR,t}$. The variables $e_{FAVAR,t}$ and $e_{AR,t}$ are the forecast errors generated by the FAVAR and AR model,s respectively. The DM test statistic is $DM = [\hat{V}(\bar{d})]^{-1/2}\bar{d}$, where $\bar{d}$ is sample mean of d_t, and $\hat{V}(\bar{d})$ is sample variance of $\bar{d}$ which is asymptotically estimated by $T^{-1}\left[\gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k\right]$, where γ_k is k[th] autocovariance of d_t which can be estimated from $T^{-1}\sum_{t=k+1}^{T}\left(d_t - \bar{d}\right)\left(d_{t-k} - \bar{d}\right)$. Under the null hypothesis, this statistic follows an asymptotic standard normal distribution.

In the forecast encompassing test, the test determines weights based on the covariance between the errors from FAVAR $e_{FAVAR,t}$, and the difference between the errors of the FAVAR and AR models, $e_{FAVAR,t} - e_{AR,t}$. If the covariance is not equal to zero, then information can be gained and a composite forecast can be built. This is tested based on $e_{FAVAR,t} = \lambda\left(e_{FAVAR,t} - e_{AR,t}\right) + \varepsilon_t$, where ε_t is a composite forecast error. The null hypothesis is $\lambda = 0$. If the null is true, then the FAVAR model encompasses the AR model. The actual test involves an OLS regression of $e_{FAVAR,t}$ on $\left(e_{FAVAR,t} - e_{AR,t}\right)$. In this study, we use t-test of $\hat{\lambda}$ for forecast encompassing.

For the forecast encompassing test, the hypothesis is that if $\hat{\lambda}$ is significantly different from zero when the dependent variable is the residuals of the FAVAR/AR models, then the AR/FAVAR models encompass the FAVAR/AR models. For the FFR model, the results show that the forecasting residuals of the FAVAR model encompass the AR model (*p*-value is 0.000), whereas the residuals of the AR model do not encompass the FAVAR model (*p*-value is 0.038) at the 1% significance level. For the WTI model, the forecasting residuals of the FAVAR model encompass the AR model (*p*-value is 0.026), whereas the residuals of the AR model do not encompass the FAVAR model (*p*-value is 0.894) at the 5% significance level. Thus, we conclude that both FAVAR models are superior in forecasting ability to the benchmark univariate AR models. Moreover, these forecast encompassing test results are consistent with the findings of the equality tests.

3.5. Conclusion

In Chapter III, we constructed econometric models for the federal fund rate and the WTI crude oil price return using a large panel of macroeconomic time series. We summarized the information with a few estimated factors by using PCA, which allowed us to interpret and identify the underlying factors. We augmented these factors as regressors in a VAR framework to assess the effects of the federal fund rate and WTI crude oil price shocks upon the US economy. In sequence, the contemporaneous causal relationships among innovations of both FAVAR models were inductively inferred by using the GES algorithm. Based on the casual structures identified by the graphical model, we estimated the impulse response functions and forecast error variance decompositions with respect to a shock in each of the augmented factors and two considered variables.

Based on DAG analysis, we found that the federal fund rate shock is exogenous in contemporaneous time as the identifying assumption in the VAR framework of the monetary shock transmission mechanism. This result is consistent with the identification assumption of Bernanke et al. (2005) and the stylized fact of Kwon (2007). However, we found that the WTI crude oil price return is not exogenous in contemporaneous time. Thus, we argue that the oil price shocks transmission mechanism identified from information flows is inferred from the data.

From the innovation accounting analysis based on our FAVAR models, we found that our results generally align with previous literature (Bernanke et al. 2005; Christiano et al. 1999; Cuaresma et al. 2004; Estrella 2005; Fair 2002; Kwon 2007; Tufte and Wohar 1999) and appear to make economic sense. In particular, we find that the price puzzle (Sims 1992) is considerably reduced and the price responses on shocks to the federal fund rate eventually decrease as noted by Bernanke et al. (2005) and Kwon (2007). Therefore, we conclude that inclusion of the information captured by the factors into the VAR framework succeeds in mitigating the price puzzle.

We conclude that using the larger macroeconomic information set for analyzing monetary policy and oil price shock transfer mechanisms is advantageous. Moreover, the results from out-of-sample-forecasts of the federal fund rate and the WTI crude oil price return reinforce this conclusion. The forecasting performance of the FAVAR models based on common factors clearly outperforms the model based on individual variables. More importantly, the FAVAR model is superior with respect to univariate AR models in out-of-sample-forecasts.

CHAPTER IV

PRICE DYNAMIC CAUSATIONS AMONG ENERGY, AGRICULTURAL, AND FINANCIAL MARKETS UNDER STRUCTURAL BREAKS

4.1. Introduction

The findings of earlier studies generally agree that energy commodity markets play an important role in macroeconomic and financial activities. Numerous works have discussed the interrelationships of the prices and volatilities among energy, macroeconomic, and financial markets. In fact, studying the dynamics and statistical properties of energy commodities has become an important part of price dynamics and commodity market analysis as well as macroeconomic and financial analysis.

In addition to this stylized fact, the potential impact of energy markets on agricultural market and their co-movements have recently attracted a lot of attention since rapidly increasing the usage of biofuels (i.e., corn-based ethanol) in 2005. Historically, US energy markets have always linked to agricultural markets. For instance, gasoline and electricity are directly used as input sources for producing and transporting agricultural products. According to the US Department of Agriculture (USDA), total energy costs were approximately 15% of annual agricultural production expenses in 2011. The potential impacts – and relationship – of energy markets on US agricultural markets gained more attention after Congress passed the Energy Policy Act of 2005, which promotes adding corn-based ethanol to gasoline, and the Chicago Board of Trade (CBOT) launched an ethanol futures contract. US corn-based ethanol production rose from 3.9 billion gallons in 2005 to 13.2 billion gallons in 2011 (RFA 2011), while wholesale corn prices rose from 1.96 $/bushel to 6.01 $/bushel in the same timeframe (NASS 2011).

Economic theory based on market fundamentals and arbitrage activities suggests that energy, agricultural, and financial markets are plausibly interrelated. Although there are many separate studies of integration between energy and agricultural markets, energy and stock markets, and/or agricultural and stock markets (Brown and Yücel 2002; Hamilton 2003; Hanson et al. 1993; Harri et al. 2009; Jiménez-Rodríguez and Sánchez 2005; Kaltalioglu and Soytas 2009; Mutuc et al. 2010; Sadorsky 1999; Soytas et al. 2009; Yu et al. 2006; Zhang et al. 2010), to our knowledge no study has yet empirically examined the dynamic causations of the prices among the energy, agricultural, and financial markets simultaneously. Therefore, Chapter IV focuses on the dynamics of contemporaneous causations among the crude oil, gasoline, corn and stock markets, allowing for structural changes or regime shifts.

We note, too, that previous studies assume that the relationship between two considered markets is unchanging or remains stable over time. However, the failure to account for structural changes or regime shifts may incur biased and unreliable results. In other words, when we employ only partial time series before or after a structural change, we provide incomplete or misleading information on potential market linkages.

In this respect, finding structural changes allows us to produce one sample before and one sample after the identified change point, and to carefully investigate the dynamics of contemporaneous causations among the three markets. However, it is widely known that the accurate directions and magnitude of these linkages are difficult to capture because their dynamic relationships vary by time, strengthening and weakening during periods of crisis. Therefore, we build an econometric model to quantify the dynamic relationships among the prices of crude oil, gasoline, corn, and the S&P 500 by using a VAR framework. Simultaneously, we apply the Bai-Perron test (Bai and Perron 2003) to investigate the possible existence of multiple structural breaks with unknown points. Based on the results of the structural break test, we divide the entire sample period into sub-periods, and impose the dynamic relationships among four prices in econometric models. In addition, we investigate the empirical contemporaneous causal relationships using the DAG approach following Bessler and Lee (2002), Bessler and Yang (2003), Demiralp and Hoover (2003), Moneta (2004), Moneta (2008), Swanson and Granger (1997), and Kim and Bessler (2007). Finally, we apply forecast error variance decomposition and an impulse response function (IRF), to analyze the information transmission among the prices of crude oil, gasoline, corn, and the S&P 500.

The reminder of this chapter is organized as follows. Section 2 describes the structural break test. Section 3 presents the data, summary statistics, and basic non-stationary test results. Section 4 discusses the analytical results of the daily price returns of crude oil, gasoline, corn, and the S&P 500. Section 5 concludes.

4.2. Empirical Methodology

This section focuses on assessing price level dynamic causations based on the VAR modelSims (1980) widely used for empirical analysis of time series. The VAR model allows us to infer the contemporaneous causal structures by using statistical properties without too much *a priori* theory and/or information from the data and to easily perform innovation accounting analysis. After estimating the dynamic relationships among the prices of crude oil, gasoline, corn, and the S&P 500, we infer the contemporaneous causal structures from innovations by using DAG. However, since the contemporaneous causal structures may change if there are structural changes or regime shifts, we first need to test for structural breaks.

Economic time series most likely contain structural breaks due to shifts in market fundamentals, such as depressions, financial crises, oil shocks, production technology, government policies, etc. The earliest tests for structural breaks were developed by Chow (1960), who proposed an analysis of variance test and a predictive test. Since the 1970s, the Chow tests have been used extensively in empirical studies. However, they have several limitations, one of which is that the tests are generally valid only under the strong assumptions that the regression error term does not suffer from autocorrelation or heteroscedasticity, and the break point is known *a priori*.

To overcome the limitations, the Chow tests were extended by the alternative CUSUM and CUSUMSQ tests of Brown et al. (1975), which depend on the basis of inference in Quandt (1960), i.e., we must infer the break point because we do not know the actual break point with certainty. Later, Krämer et al. (1988), Ploberger et al. (1989), and Ploberger and Kramer (1992) extended Brown et al. to show how the CUSUM test can be accomplished using OLS residuals. One noticeable drawback of the CUSUM test is its asymptotically low power against instability in the intercept but not in the entire coefficient vector. Therefore, Ploberger et al. (1989) proposed a fluctuation test based on comparisons between parameter estimates from the partial samples and the complete sample, assuming stationary regressors of the model. Andrews (1993) derived the asymptotic null distribution of the sequential likelihood ratio test (Quandt 1960) of parameter constancy. He also showed that this test has nontrivial local asymptotic power against all alternatives of non-constant parameters. Andrews and Ploberger (1994) developed tests with stronger optimality properties than Andrews (1993) in the context of Maximum Likelihood Estimators (MLEs).

Another effort to overcome the limitations of Chow tests were against the alternative to constancy, i.e., parameters are stochastic and fluctuate according to some time series model. By using the assumption that if the null is not true, the parameters follow a random walk, LaMotte and McWhorter Jr (1978) provided an exact F test for testing against the alternative to constancy. Extensions have since been made by Nyblom and Makelainen (1983) and Nyblom (1989). They developed the locally most powerful test against a parameter variation in the form of a martingale. Recently, Bai and Perron (1998) and Bai and Perron (2003) considered issues related to multiple structural changes occurring at unknown dates in their linear regression model estimated by OLS. In particular, they examined several aspects of the structural break models including the consistency of the break fraction estimators, the rate of convergence, and the construction of tests that allow inferences for deciding the presence of structural change and the number of breaks. They considered a simulation study and empirical application, and presented an efficient algorithm to obtain global minimizing of the sum of squared residuals (Bai and Perron 2003; Bai and Perron 1998).

Since one of our objectives is to identify possible multiple structural breaks absent prior information of break dates, we select the method suggested by Bai and Perron (1998) and Bai and Perron (2003) to determine whether the considered series contain unknown structural breaks. The remainder of this section provides the details.

4.2.1. Test for Structural Changes

The multiple break testing and estimation methodology of Bai and Perron (1998) and Bai and Perron (2003) requires no *a priori* information regarding the number and timing of potential breaks, and allows for serial correlation and heteroscedasticity in the errors across structural regimes.

First, we consider a multiple linear regression model with *m* breaks:

$$y_t = \beta_j z_t + u_t \tag{4.1}$$

where y_t is the observed dependent variable, z_t is $(q \times 1)$ vector of covariates, β_j is the mean of volatility in $(j+1)^{th}$ regime, and u_t is the error term at time *t*. Also, $t = T_{j-1} + 1, \cdots, T_j, j = 1, \cdots, m+1, T_0 = 0$ and $T_m = T$.

We treat the break points $(T_1, \cdots, T_m)$ as unknown. Note that this is a pure structural change model and that all coefficients are subject to change. In addition, the model permits correlation and heterogeneity in the residuals (Bai and Perron 1998).

We express equation (4.1) as the matrix form of the multi-variables linear model:

$$Y = BZ + U \tag{4.2}$$

where $Y = (y_1, \cdots, y_T)'$, $B = (\beta_1, \beta_2, \cdots, \beta_{m+1})'$, $U = (u_1, \cdots, u_T)'$, and Z is the diagonal matrix with $Z = diag(z_1, \cdots, z_{m+1})$.

We estimate the unknown regression coefficients and the break points $(\beta_1, \beta_2, \cdots, \beta_{m+1}, T_1, \cdots, T_m)$ when T observations on (y_t, z_t) are available based on the least-squares principle proposed by Bai and Perron (1998). For each m-partitions $(T_1, \cdots, T_m)$, denoted T_j, we obtain the associated least-squares estimates of $\beta_j(T_1, \cdots, T_m)$ by minimizing the sum of squared residuals which we express as:

$$SSR_T(T_1, \cdots, T_m) = \sum_{i=1}^{m+1} \sum_{t=T_{i-1}+1}^{T_i} \left(y_t - \beta_j z_t\right)^2 \tag{4.3}$$

The estimated parameters are the mean estimates of regimes based on the partitions; we put them back into the objective function and denote the sum of squared residuals as $SSR_T(T_1, \cdots, T_m)$. By considering all the possible m-partitions $(T_1, \cdots, T_m)$, we can obtain the estimated break points from:

$$\left(\hat{T}_1, \cdots, \hat{T}_m\right) = \operatorname{argmin}_{T_1, \cdots, T_m} SSR_T(T_1, \cdots, T_m) \tag{4.4}$$

where the minimization is taken over all possible partitions $(T_1, \cdots, T_m)$ such that $T_i - T_{i-1} > q$. Thus the break point estimators are global minimizers of the objective function. Finally, we can identify that the regression parameter estimates are the associated least-squares estimates at the estimated m-partition (T_m), i.e., $\hat{\beta}_{m+1} = \hat{\beta}_{m+1}(\{\hat{T}_m\})$.

The Bai and Perron test uses an efficient dynamic programming algorithm (Bai and Perron 2003) to determine the number of breaks and their break dates, which begins by testing for a single break, proceeds to two breaks, etc. The optimal number of breaks (*m-1*) is evaluated based on the optimal break that gives the lowest sum of squared residuals. To check for structural breaks in the series, we implement the procedure to determine the existence of structural change and to select the number of breaks suggested by Bai and Perron (1998) and Bai and Perron (2003). Our procedure consists of three tests.

Test 1: We find the *F*-statistics for testing null of no structural breaks against alternative m breaks where the breaks are selected according to equation (4.4). We call this the $SupF_T(m)$ test. To test the null hypothesis of no structural breaks in beta against the alternative of m breaks, let $(T_1, \cdots, T_m)$ be a partition such that $T_i = [T\lambda_i]$, where $i = 1, \cdots, m$. Also define R such that $(R\beta)' = (\beta_1' - \beta_2', \cdots, \beta_m' - \beta_{m+1}')$. Calculate the statistic as:

$$F_T(\lambda_1, \cdots, \lambda_m) = \frac{1}{T}\left(\frac{T-m-2}{m}\right) \hat{\beta}' R' \left[R\hat{V}'\left(\hat{\beta}\right)R'\right]^{-1} R\hat{\beta} \tag{4.5}$$

where $\hat{V}\left(\hat{\beta}\right)$ is an estimate of the variance-covariance matrix for $\hat{\beta}$ that is robust to heteroscedasticity and serial correlation.

Next, consider the maximum *F* statistics corresponding to the following equations:

$$SupF_T(m) = F_T\left(\hat{\lambda}_1, \cdots, \hat{\lambda}_m\right) \tag{4.6}$$

where $(\hat{\lambda}_1, \cdots, \hat{\lambda}_m)$ minimize the global sum of squared residuals, $S_T(T\hat{\lambda}_1, \cdots, T\hat{\lambda}_m)$, under the restriction that $(\hat{\lambda}_1, \cdots, \hat{\lambda}_m) \in \Theta_\pi$, where $\Theta_\pi = \{(\lambda_1, \cdots, \lambda_m); |\lambda_{i+1} - \lambda_i| \geq \pi, \lambda_1 \geq \pi, \lambda_m \leq 1 - \pi\}$ for some arbitrary small positive number π (the trimming parameter).

Test 2: We examine the test for null of no structural breaks against $1 \leq m \leq M$ breaks where M is an upper bound on the number of possible breaks. Given no specification of the number of breaks, Bai and Perron (1998) introduce a new class of tests of no structural break against an unknown number of breaks, given some upper bound *M*. We call these the double maximum tests which define for some fixed weight $(a_1, \cdots a_M)$:

$$DmaxF_T(M, q, a_1, \cdots a_M) = \max_{1 \leq m \leq M} a_m \, supF_T(\hat{\lambda}_1, \cdots, \hat{\lambda}_m; q) \tag{4.7}$$

where q is degrees of freedom.

The first version of the double maximum tests suggested by Bai and Perron (1998) sets all weights equal to unity, employing the statistic

$$UDmaxF_T(M, q) = \max_{1 \leq m \leq M} supF_T(\hat{\lambda}_1, \cdots, \hat{\lambda}_m; q) \tag{4.8}$$

where $\hat{\lambda}_j = \frac{\hat{T}_j}{T}$, $j = 1, \cdots, m$ are the estimates of the break points obtained using the global minimization of the sum of squared residuals. However, a problem associated with the $UDmaxF_T(M, q)$ test concerns a fixed m. Since $F_T(\hat{\lambda}_1, \cdots, \hat{\lambda}_m; q)$ is the sum of m dependent Chi-square random variables with q degrees of freedom and each is divided by m, the critical values of the individual tests decrease as m increases. Thus, the marginal *p*-values decrease with *m,* which may lead to a test with low power if there is a large number of breaks (Bai and Perron 1998). As an alternative, Bai and Perron (1998) suggested the $WDmaxF_T(M,q)$ test:

$$WDmaxF_T(M,q) = \max_{1\le m\le M} \frac{c(q,\alpha,1)}{c(q,\alpha,m)} \times supF_T(\hat{\lambda}_1,\cdots,\hat{\lambda}_m;q) \tag{4.9}$$

where $\frac{c(q,\alpha,1)}{c(q,\alpha,m)}$ is the weight for $m = 1$ as $\frac{c(q,\alpha,1)}{c(q,\alpha,m)} = 1 = a_1$ and for $m > 1$ as $\frac{c(q,\alpha,1)}{c(q,\alpha,m)} = a_m$. This test assumes weights such that the marginal p-values are equal across values of m[15].

Test 3: Bai and Perron (1998) proposed a sequential test of m versus $m+1$ breaks, which we call the $Sup\ F_T(m+1|m)$ test. This test is based on the difference between the sums of squared residuals obtained with m breaks and with $m+1$ breaks. For each segment containing the observations $\hat{T}_{i-1}+1$ to T_i, where $i = 1,\cdots,m+1$, we test the null hypothesis of no structural break versus the alternative of a single change. If the overall minimal value of the sum of squared residuals (over all segments where an additional break is included) is sufficiently smaller than the sum of squared residuals obtained from the m break model, we reject the model with m breaks. The break date selection is the one with the overall minimum.

[15] The asymptotic distributions of these statistics are derived by Bai and Perron (1998); the critical values appear in Bai and Perron (2003).

4.3. Data Description

We use daily prices series for US crude oil, gasoline, corn, and the S&P 500 from January 2, 2001 through December 30, 2011 for a total of 2870 observations, excluding all public holidays for all markets. The data series derive from Thomson DataStream. The crude oil price is West Texas Intermediate (WTI) crude oil FOB spot prices, gasoline price is New York Harbor conventional gasoline regular FOB spot prices, corn price is No.2 yellow corn FOB Gulf prices, and the S&P 500 index price is based upon the daily closing prices of the S&P 500. We use the first difference of log transformed price series as a measure of returns and denote the price returns as R_{WTI}, $R_{GASOLINE}$, R_{CORN}, and $R_{S\&P500}$. Table 4.1 and figure 4.1 report the summary statistics and time series plots of the daily returns.

We start by analyzing the dynamic behavior of each univariate series, which serves to facilitate the multivariate modeling and the understanding of multivariate dynamics. Table 4.1 reports the summary statistics. We note that the S&P 500 return only shows negative mean returns, whereas the other market price returns show positive mean returns. Based on the magnitude of the unconditional standard deviations, the gasoline market is slightly volatile and shows positive skewness and moderate kurtosis, whereas the other markets are not volatile and show negative skewness. The S&P 500 return shows the highest magnitude in kurtosis, whereas the corn market shows the smallest.

Table 4.1
Summary statistics for each series of price returns

	WTI Crude Oil	Gasoline	Corn	S&P 500
Mean	0.00020	0.00020	0.00017	-0.00001
Standard Deviation	0.01096	0.01248	0.00884	0.00591
Variance	0.00012	0.00016	0.00008	0.00003
Kurtosis	4.56852	4.52089	2.71403	7.97309
Skewness	-0.18764	0.07839	-0.10490	-0.16864
Minimum	-0.07423	-0.07769	-0.05260	-0.04113
Maximum	0.07128	0.10219	0.04729	0.04759
JB (*p*-value)	10.8378 (0.004)	9.74038 (0.008)	0.52416 (0.769)	103.523 (0.000)
Number of Observation	2870	2870	2870	2870

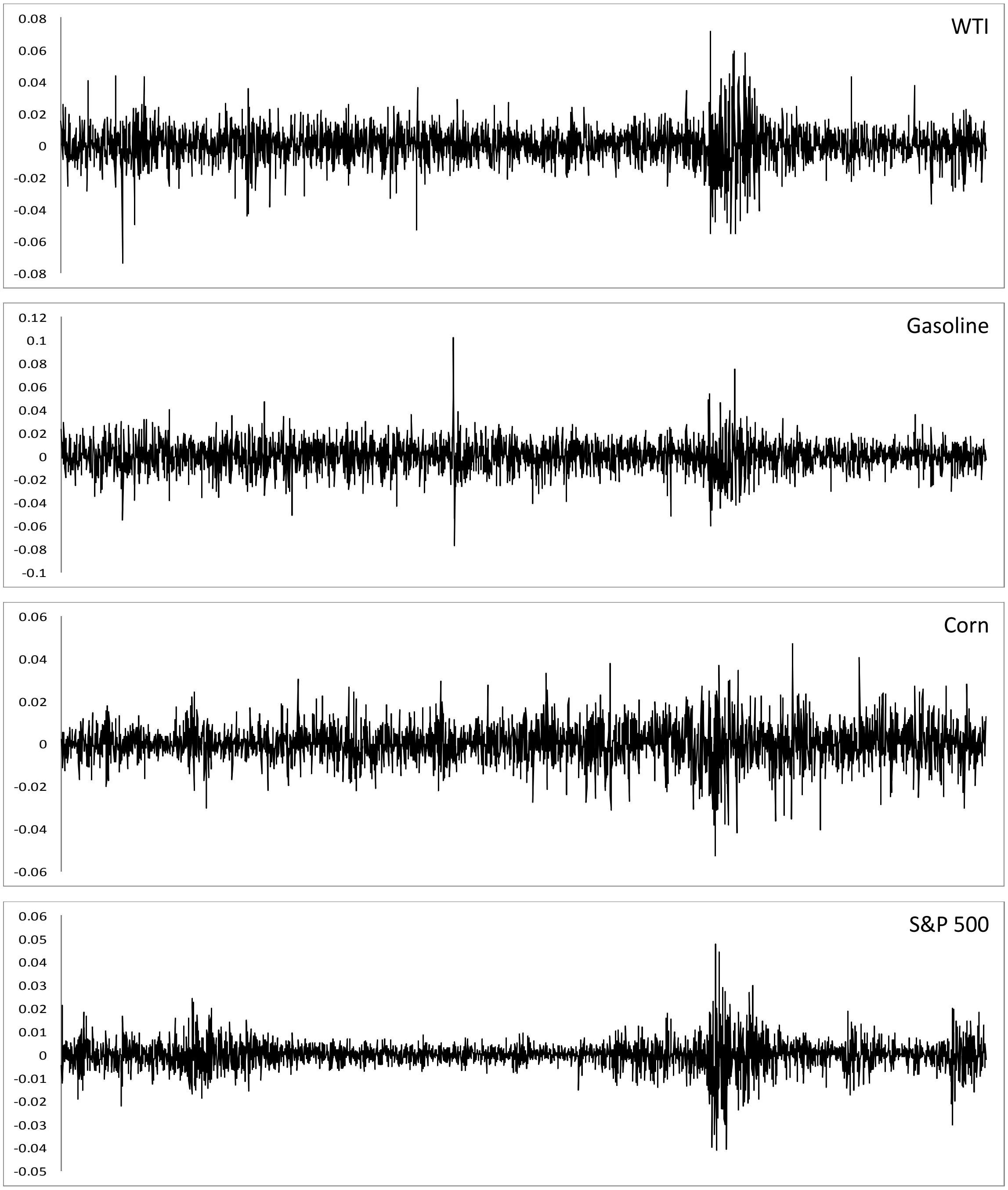


Figure 4.1. Plots of the daily return for each series

We test for the presence of a unit root for the log transformed prices and price returns of each market. A series with a unit root is non-stationary with an infinite unconditional variance, making it impossible to generalize it to other time periods. Table 4.2, which reports the Dickey-Fuller test and augmented Dickey-Fuller test statistics, shows that the log level prices of all variables fail to reject the null hypothesis of a unit root at the 10% significance level. However, all price returns, i.e., first differencing of the logarithm of the price series, result in rejecting the null hypothesis at the 1% significance level, indicating stationary.

Table 4.2
Tests for non-stationary of energy price returns

Price return Series	DF Test		ADF Test (k)[a]	
	Log Level	First Difference	Log Level	First Difference
WTI Crude Oil	-1.43	-55.04*	-1.38	-38.90*
Gasoline	-1.63	-52.46*	-1.66	-38.05*
Corn	-0.76	-53.63*	-0.77	-37.63*
S&P 500	-2.38	-59.12*	-2.13	-41.83*

Note: * indicates 1% significance level; the critical value is -3.51 at the 1% significance level;
[a] indicates the number of lag determined by the optimal lag order selection criteria.

4.4. Empirical Results

This section reports the empirical results of the tests for multiple structural breaks and VAR models fitted to the data. Based on the results of the break test, we divide the entire sample period into several sub-periods, and impose the dynamic relationships among four prices in econometric models. We also report the estimated results of the contemporaneous causal relationships for each sub-period using the PC algorithm. Finally, we implement innovation accounting analysis based on structural innovations.

4.4.1. Test Results for Structural Breaks

Table 4.3 summarizes the Bai and Perron test results for the four price returns. For the WTI crude oil and gasoline cases, the $Sup\ F_T(1)$ test results indicate insignificant at all three levels of significance. However, when we test the $Sup\ F_T(m)$ up to 5 breaks, the test rejects the null hypothesis of no structural breaks and accepts the existence of 2, 3, and 4 potential structural breaks in the WTI crude oil mean price returns at the 1%, 5%, and 10% significance levels, respectively. For the gasoline case, all test results show the existence of 2, 3, 4, and 5 breaks at the 5% significance level. However, the validation of only two breaks from the $Sup\ F_T(m+1|m)$ test, sequential procedure, and BIC results indicates that the $Sup\ F_T(2|1)$ test statistics are significant at 1%; similarly, the sequential procedure and BIC suggest the existence of two breaks in both cases. In contrast, when referring to the LWZ[16], this test unfailingly selects no

[16] LWZ is a modified criterion of Schwarz which is proposed by Lui, Wu and Zidek (1997).

structural break for both cases. In a sense, this is inevitable since the $Sup\ F_T(1)$ tests are statistically insignificant. Thus, the LWZ shows results similar to the $Sup\ F_T(1)$ test. In short, the existence of two breaks are strongly supported by the significant results of the $Sup\ F_T(2)$, $Sup\ F_T(2|1)$, $UDmax$ and $WDmax$ tests as well as the sequential procedure and BIC. Thus, we find two significant structural breaks in the WTI crude oil and gasoline mean between January 1, 2001 and December 30, 2011, i.e., August 25, 2008 and December 15, 2008 for WTI crude oil, and September 5, 2005 and November 12, 2008 for gasoline.

In the corn and S&P 500 cases, all results of the $Sup\ F_T(m)$, $UDmax$ and $WDmax$ tests indicate insignificance at all significance levels. However, the $Sup\ F_T(2|1)$ test only rejects the null hypothesis of no structural breaks at the 10% significance level for both cases. The sequential procedure, LWZ, and BIC reach a similar conclusion. The LWZ and BIC report no breaks, whereas the sequential procedures suggest two structural breaks for both cases. Similar to the WTI and gasoline cases, the non-existence of structural breaks is due to the results of the $Sup\ F_T(1)$ tests. Thus, we find two significant structural break points in the corn and S&P 500 mean returns between January 1, 2001 and December 30, 2011, i.e., October 20, 2005 and September 18, 2008 for corn, and September 23, 2008 and December 16, 2008 for the S&P 500.

Table 4.3
Test results for structural breaks

	WTI	Gasoline	Corn	S&P 500
$Sup\ F_T(1)$	2.92	3.35	3.01	3.35
$Sup\ F_T(2)$	11.11***	10.09**	5.77	6.08
$Sup\ F_T(3)$	8.48**	8.21**	5.36	5.89
$Sup\ F_T(4)$	6.98*	8.29**	4.91	5.34
$Sup\ F_T(5)$	6.04	7.47**	4.56	5.16
$Sup\ F_T(2\|1)$	19.21***	16.73***	8.49*	8.77*
$Sup\ F_T(3\|2)$	3.13	4.35	4.47	5.43
$Sup\ F_T(4\|3)$	2.41	8.21	3.50	3.58
$Sup\ F_T(5\|4)$	2.24	4.01	3.05	4.32
$UDmax$	11.11**	10.09*	5.77	5.52
$WDmax$	11.32**	10.28*	5.88	5.63
Number of Breaks Selected				
Sequential	2	2	2	2
LWZ	0	0	0	0
BIC	2	2	0	0
Break Points				
$\hat{T}_1$	08/25/2008	09/05/2005	10/20/2005	09/23/2008
$\hat{T}_2$	12/15/2008	11/12/2008	09/18/2008	12/16/2008

Note: *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.

We note that the majority of the break points occur between September and October 2005, and between September and November 2008. The first break point coincides with Hurricane Katrina and the ethanol boom. Since Katrina caused heavy damage to US refinery and domestic oil production capacity in the Gulf of Mexico, the effect on gasoline prices was significant, whereas the effect on crude oil prices was moderate. The ethanol boom was caused by oil price increases, passage of the US energy legislation, and CBOT's introduction of an ethanol futures market. The effect mainly influenced the US corn market. The second break point (between September and November 2008) coincides with the collapse of Lehman Brothers, i.e., the S&P 500 index fell 4.71%, the Dow Jones industrial average fell 4.42%, and the NASDAQ index fell 3.60%. After this, hedge funds gradually withdrew positions from the oil futures markets and the price of oil price sharply decreased. Based on this analysis, we divide the full sample period into 01/01/2001 ~ 09/05/2005 for sub-period 1, 09/06/2005 ~ 12/16/2008 for sub-period 2, and 12/17/2008 ~ 12/30/2011 for sub-period 3.

4.4.2. Vector Autoregression Results

This section describes our preliminary data analysis conducted by applying a VAR (p) process to the daily price returns of each series split over the three sub-periods. We apply the MLE procedure of Johansen (1991) to construct a VAR (p) process and determine the optimal lag length based on loss information criteria, i.e., Akaike, Schwarz, and Hannan and Quinn losses. Table 4.4 reports the results. The criteria show somewhat ambiguous results: AIC, HQIC and SIC indicate $p = 1$ as an optimal lag

Table 4.4
VAR optimal lag-length determination

Period	Lag Order	Akaike Information Criterion (AIC)	Hannan and Quinn Information Criterion (HQIC)	Schwarz Information Criterion (SIC)
Entire period	0	-26.8039	-26.8009	-26.7836
	1	-26.8252*	-26.8102*	-26.7956*
	2	-26.8219	-26.7949	-26.7470
Sub-period 1	0	-27.4609	-27.4294	-27. 3771
	1	-27.4691*	-27.4628*	-27.4524*
	2	-27.4474	-27.3906	-27.2964
Sub-period 2	0	-26.4530	-26.4444	-26.4307
	1	-26.5663	-26.5235*	-26.4546*
	2	-26.5813*	-26.5043	-26.3803
Sub-period 3	0	-27.2344	-27.2038	-27.1382
	1	-27.2443	-27.2263*	-27. 2132*
	2	-27.2492*	-27.1763	-27.0582

Note: * indicates the most appropriate lag order for the considered model; the information criteria used to identify the optimal lag length (p) of a VAR process are $AIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{2n}{T}\right)$, $SIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{nlnT}{T}\right)$, and $HQIC = \ln\left(det\widehat{\Omega}_p\right) + p\left(\frac{2nln(lnT)}{T}\right)$, where $\widehat{\Omega}_p$ is the maximum likelihood estimate of variance-covariance matrix of Ω, p is the proposed lag length, n is the number of variables, and T is the sample size.

order for the entire period and sub-period 1, HQIC and SIC indicate $p = 1$, and AIC indicates $p = 2$ for sub-periods 2 and 3. Following SIC, we select $p = 1$ as an optimal lag order for the sub-periods.

Thus, we choose the most parsimonious specification of a VAR (1) model and proceed to fit the three VAR models to the four-variate daily price returns of the time series. Tables 4.5 through 4.7 report the estimated parameters and robust standard errors for each sub-period. We note some dramatic changes. For example, in sub-period 1, two coefficients in the gasoline equation are only statistically significant, whereas the others in the remaining equations are not statistically significant like the random walk model. However, in sub-period 2, six parameters of the WTI crude oil, corn, and S&P 500 equations are statistically significant, whereas the gasoline equation is similar to the random walk model. Again, only two parameters in the S&P 500 equation are statistically significant in sub-period 3.

Table 4.5
VAR (1) model estimation results for sub-period 1

Parameters	$R_{WTI,}$ $(i = 1)$		$R_{GASOLINE,}$ $(i = 2)$		$R_{CORN,}$ $(i = 3)$		$R_{S\&P500,}$ $(i = 4)$	
	Coefficient	Std. Err	Coefficient	Std. Err	Coefficient	Std. Err	Coefficient	Std. Err
d_i	0.00034	0.00030	0.00042	0.00038	-0.00003	0.00020	-0.00003	0.00014
$\gamma_{1,i1}$	-0.07264*	0.03753	-0.08958*	0.04686	-0.00206	0.02455	0.00151	0.01768
$\gamma_{1,i2}$	0.02964	0.02999	0.11265***	0.03744	-0.01676	0.01962	0.00573	0.01413
$\gamma_{1,i3}$	-0.03568	0.04409	0.01652	0.05506	-0.01573	0.02884	-0.00200	0.02077
$\gamma_{1,i4}$	0.05538	0.06093	0.05561	0.07608	-0.01514	0.03986	-0.03864	0.02871

Diagnostics tests	R_{WTI}	$R_{GASOLINE}$	R_{CORN}	$R_{S\&P500}$
R^2	0.0046	0.0079	0.0016	0.0018
RMSE	0.01062	0.01326	0.00694	0.00501
χ^2	5.6968	9.6655**	1.8988	2.2322
Log-likelihood	16762.99			
# of observations	1220			

Note: *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.

Table 4.6
VAR (1) model estimation results for sub-period 2

Parameters	$R_{WTI,}$ $(i = 1)$		$R_{GASOLINE,}$ $(i = 2)$		$R_{CORN,}$ $(i = 3)$		$R_{S\&P500,}$ $(i = 4)$	
	Coefficient	Std. Err	Coefficient	Std. Err	Coefficient	Std. Err	Coefficient	Std. Err
d_i	-0.00015	0.00036	-0.00042	0.00043	0.00034	0.00034	-0.00020	0.00023
$\gamma_{1,i1}$	-0.15414***	0.04695	-0.07547	0.05667	-0.12110***	0.04521	-0.08053***	0.02966
$\gamma_{1,i2}$	0.00686	0.03708	-0.03161	0.04476	0.02628	0.03571	0.00651	0.02342
$\gamma_{1,i3}$	0.01919	0.03813	0.09158*	0.04602	0.03270	0.03671	-0.02048	0.02408
$\gamma_{1,i4}$	0.32907***	0.05510	0.02076	0.06650	0.27410***	0.05306	-0.11266***	0.03481

Diagnostics tests	R_{WTI}	$R_{GASOLINE}$	R_{CORN}	$R_{S\&P500}$
R^2	0.051	0.0083	0.0361	0.0380
RMSE	0.01441	0.012602	0.010054	0.006595
χ^2	55.0236***	7.17085	31.9727***	33.6652***
Log-likelihood	11349.71			
# of observations	854			

Note: *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.

Table 4.7
VAR (1) model estimation results for sub-period 3

Parameters	R_{WTI}, $(i = 1)$		$R_{GASOLINE}$, $(i = 2)$		R_{CORN}, $(i = 3)$		$R_{S\&P500}$, $(i = 4)$	
	Coefficient	Std. Err	Coefficient	Std. Err	Coefficient	Std. Err	Coefficient	Std. Err
d_i	0.00042	0.00042	0.00054	0.00038	0.00037	0.00035	0.00021	0.00022
$\gamma_{1,i1}$	0.07240	0.04723	-0.03042	0.04259	-0.01942	0.03962	0.01466	0.02513
$\gamma_{1,i2}$	-0.03692	0.05178	0.07491	0.04669	-0.02858	0.04344	-0.01597	0.02755
$\gamma_{1,i3}$	0.01204	0.04461	-0.06464	0.04022	-0.00362	0.03742	0.05199**	0.02374
$\gamma_{1,i4}$	0.01862	0.07609	-0.05386	0.06861	0.02724	0.06383	-0.12244***	0.04049

Diagnostics tests	R_{WTI}	$R_{GASOLINE}$	R_{CORN}	$R_{S\&P500}$
R^2	0.0039	0.0065	0.0020	0.0159
RMSE	0.01114	0.01010	0.00963	0.00594
χ^2	3.5551	5.9471	1.8055	14.7075***
Log-likelihood	12396.26			
# of observations	911			

Note: *, **, and *** indicate 10%, 5%, and 1% significance levels, respectively.

4.4.3. Directed Acyclic Graphs

The VAR approach does not explicitly provide the information on the causal structures of the four price return series in contemporaneous time. Thus, we present results from the DAGs which provide causal information about the contemporaneous interrelationships among the four markets. From the residuals of the VAR models for each sub-period, we obtain the three identical causal relationship graphs from TETRAD IV's PC algorithms at the 10% significance level[17] (figure 4.2).

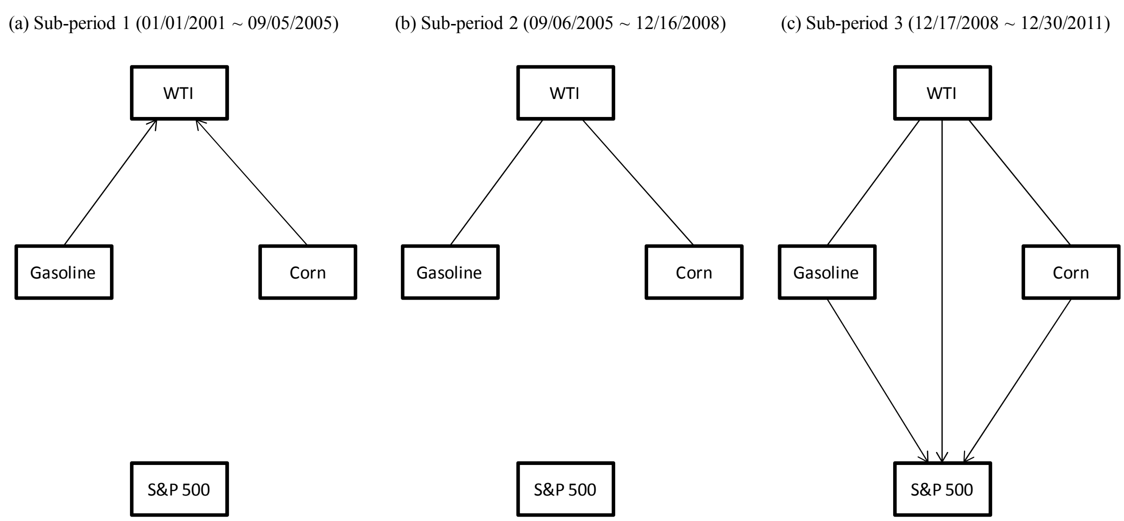


Figure 4.2. Contemporaneous causal relationships among four daily price returns for each sub-period.

Figure 4.2 panel (a) shows the causal linkages among the daily price returns of WTI crude oil, gasoline, and corn. More specifically, gasoline and corn price returns

[17] Contemporaneous casual patterns are the same at the 5% significance level from TETRAD IV's PC algorithm.

directly affect the WTI crude oil price return in contemporaneous time, i.e., the WTI crude oil price returns do not lead any other market, but respond to information signals from gasoline and corn in contemporaneous time. Meanwhile, the S&P 500 return is independent of the other price returns.

Figure 4.2 panel (b) shows that the contemporaneous causal structures for sub-period 2 have patterns similar to the information flow for sub-period 1, i.e., there is significant information flow among the price returns of WTI crude oil, gasoline, and corn in contemporaneous time, whereas the causal link to the S&P 500 return is insignificant. However, the PC algorithm fails to validate the directions of two edges, WTI-gasoline and WTI-corn.

Figure 4.2 panel (c) shows the results of the DAG patterns for sub-period 3. WTI crude oil, gasoline, and corn price returns directly affect the S&P return, whereas the two undirected edges (WTI-gasoline and WTI-corn) still remain. The remarkable finding that more contemporaneous causal relationships appear to be present in sub-periods 1 and 2 implies that information flow is quicker or more efficient within the four price returns in recent times than in the past. Further, before the structural break, the S&P 500 return is causally isolated from information which comes from the price returns of WTI crude oil, gasoline, and corn. Thus, those three price returns have no influence on the S&P 500 return in contemporaneous time. After the structural break, it appears that the revolution of causal relationships occurred in the S&P 500 return where information from WTI crude oil, gasoline, and corn comes together to determine the S&P 500 return, the most endogenous variable.

We note that the PC algorithm does not explain how to assign directed edges in the two undetermined relationships, WTI-gasoline and WTI-corn for sub-periods 2 and 3. In general, the correlation and conditional correlation patterns associated with these three price returns are not enough to assign causal structures for the undetermined edges. We need additional information from existing economic theory, and/or include other markets in the modeling which may give the possibility of more correlation and conditional correlations patterns; without it, we must carefully look at any movements towards a DAG (Bessler et al. 2003). Thus, we consider three possible cases of directed causal relationships for the two undirected edges. Figures 4.3 and 4.4 provide the three equivalent DAGs for sub-periods 2 and 3, respectively. Using the six possible equivalent DAG patterns, we present the results of the innovation accounting analysis in the next sections.

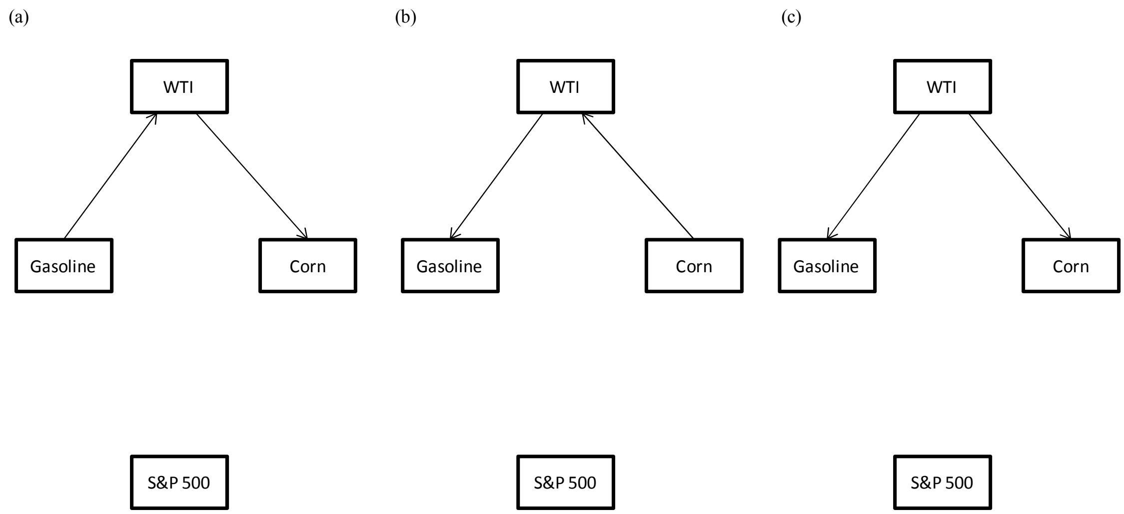


Figure 4.3. The possible equivalent DAGs for sub-period 2

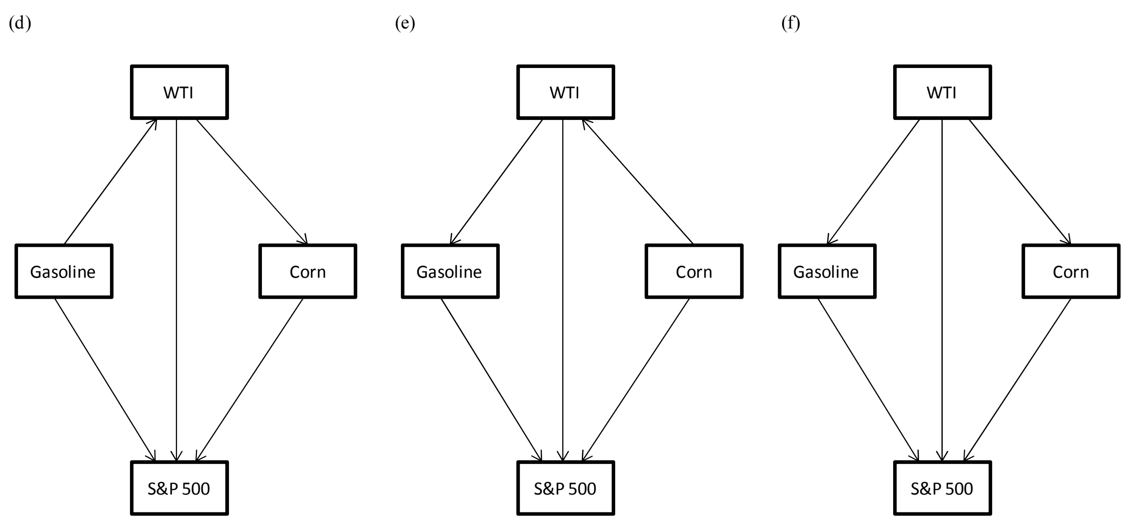


Figure 4.4. The possible equivalent DAGs for sub-period 3

In addition, we compare and proceed to test the equality of variance-covariance matrices between the residuals from the VAR models for each sub-period. Figure 4.5 gives the details.

$$\Sigma_{sub-period\ 1} = \begin{array}{cc} \begin{array}{cccc} R_{WTI} & R_{GASOLINE} & R_{CORN} & R_{S\&P500} \end{array} & \\ \begin{bmatrix} 1.12E-04 & & & \\ 9.04E-05 & 1.75E-04 & & \\ 5.81E-06 & 2.18E-06 & 4.81E-05 & \\ -9.69E-07 & -4.61E-07 & 1.71E-06 & 2.50E-05 \end{bmatrix} & \begin{array}{l} R_{WTI} \\ R_{GASOLINE} \\ R_{CORN} \\ R_{S\&P500} \end{array} \end{array}$$

$$\Sigma_{sub-period\ 2} = \begin{array}{cc} \begin{array}{cccc} R_{WTI} & R_{GASOLINE} & R_{CORN} & R_{S\&P500} \end{array} & \\ \begin{bmatrix} 1.08E-04 & & & \\ 8.88E-05 & 1.58E-04 & & \\ 3.65E-05 & 3.13E-05 & 1.00E-04 & \\ 1.71E-05 & 1.93E-05 & 1.01E-05 & 4.32E-05 \end{bmatrix} & \begin{array}{l} R_{WTI} \\ R_{GASOLINE} \\ R_{CORN} \\ R_{S\&P500} \end{array} \end{array}$$

$$\Sigma_{sub-period\ 3} = \begin{array}{cc} \begin{array}{cccc} R_{WTI} & R_{GASOLINE} & R_{CORN} & R_{S\&P500} \end{array} & \\ \begin{bmatrix} 1.36E-04 & & & \\ 7.75E-05 & 1.11E-04 & & \\ 3.22E-05 & 2.41E-05 & 9.59E-05 & \\ 3.17E-05 & 2.88E-05 & 1.42E-05 & 3.86E-05 \end{bmatrix} & \begin{array}{l} R_{WTI} \\ R_{GASOLINE} \\ R_{CORN} \\ R_{S\&P500} \end{array} \end{array}$$

Figure 4.5. The variance-covariance matrices of the residuals from the VAR models for each sub-period

Next, we test the equality of the three variance-covariance matrices by using the multivariate Box M statistics (Box 1949). The two null hypotheses are:

$H_0{}^1: \Sigma_{sub-period\ 1} = \Sigma_{sub-period\ 2}$ and

$$H_0{}^2: \Sigma_{sub-period\ 2} = \Sigma_{sub-period\ 3} \tag{4.10}$$

Table 4.8 summarizes the statistics of the Box M tests for sub-periods 1 and 2 and sub-periods 2 and 3, respectively.

Table 4.8
Summary of test statistics of the Box M tests

	Box M test statistics	Critical Value at 1% significance level	*p*-value
Between sub-periods 1 and 2	276.0661	23.21	0.000
Between sub-periods 2 and 3	122.6791	23.21	0.000

Note: We adopt the Box M test when the sample size is small, following Mardia and Kent (1979). The Box M test statistic is generated by $M = \gamma \sum_{i=1}^{g}(n_i - 1)log\left|S_u S_{u_i}^{-1}\right|$, where $\gamma = 1 - \frac{2k^2+3k-1}{6(k+1)(g-1)}\left(\sum_{i=1}^{g}\frac{1}{n_i-1} - \frac{1}{n-g}\right)$, $S_u = \frac{n}{n-g}S$, $S_{u_i} = \frac{n_i}{n_i-g}S_i$, $S = \sum_{i=1}^{g}\frac{n_i S_i}{n}$ is the pooled covariance matrix, g is the number of groups with non-singular covariance matrices, $n = n_1 + n_2 + \cdots + n_g$ is the number of the total sample size, n_i is number of the sample size for deriving sample covariance matrix S_i, k is the dimension of the covariance matrix, and $i = 1, 2, \cdots, g$. The Box M test statistic is asymptotically distributed as a Chi-square distribution with the degree of freedom, $k(k+1)/2$.

The results indicate that the test statistic values are 276.0661 and 122.6791 for sub-periods 1 and 2 and sub-periods 2 and 3, respectively. These results suggest that we can reject the null hypothesis at the 1% significance level (*p*-values are 0.000 for both cases), since the critical value is 23.21. Thus, we say that the two covariance matrices from sub-periods 1 and 2 and sub-periods 2 and 3 significantly differ, i.e., the covariance matrices of the residuals between sub-periods 1 and 2 and between sub-periods 2 and 3 are not homogeneous.

4.4.4. Impulse Response Functions

From the analysis of the impulse responses, we can evaluate the dynamic mechanism by which innovations in one market are transmitted to the other markets over time. Here, we verify information on the direction and significance of dynamic responses for 6 days ahead-horizon following an initial shock of each price return. Figures 4.6 through 4.8 show the impulse responses (the effect of a shock in one price return transfer to other price returns) based on the DAGs.[18] The black solid line indicates the estimated response in sub-period 1, equivalent DAG type (a) in sub-period 2, and equivalent DAG type (d) in sub-period 3. The red dashed line represents the estimated response in equivalent DAG types (b) and (e) in sub-periods 2 and 3, respectively. The blue dotted line plotted in graphs for sub-periods 2 and 3 represents the estimated response in equivalent DAG types (c) and (f), respectively.

[18] The contemporaneous causal structures for each sub-period described in Figures 4.2 through 4.4 are used in a Bernanke factorization for orthogonalization to generate impulse responses and forecast error variance decompositions to describe the dynamic structures.

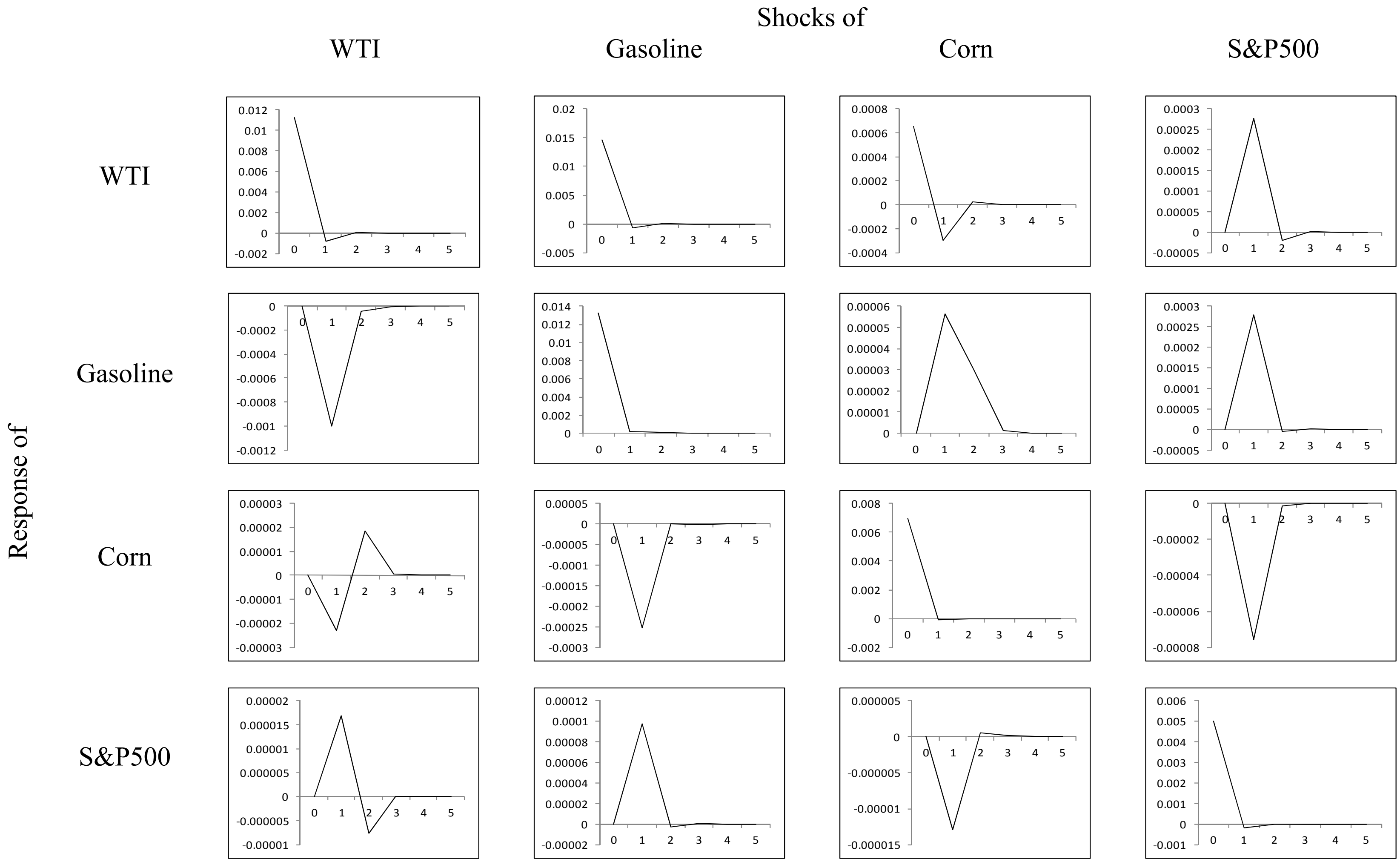


Figure 4.6. Impulse responses for sub-period 1 (2001/01/01 ~ 2005/09/05)

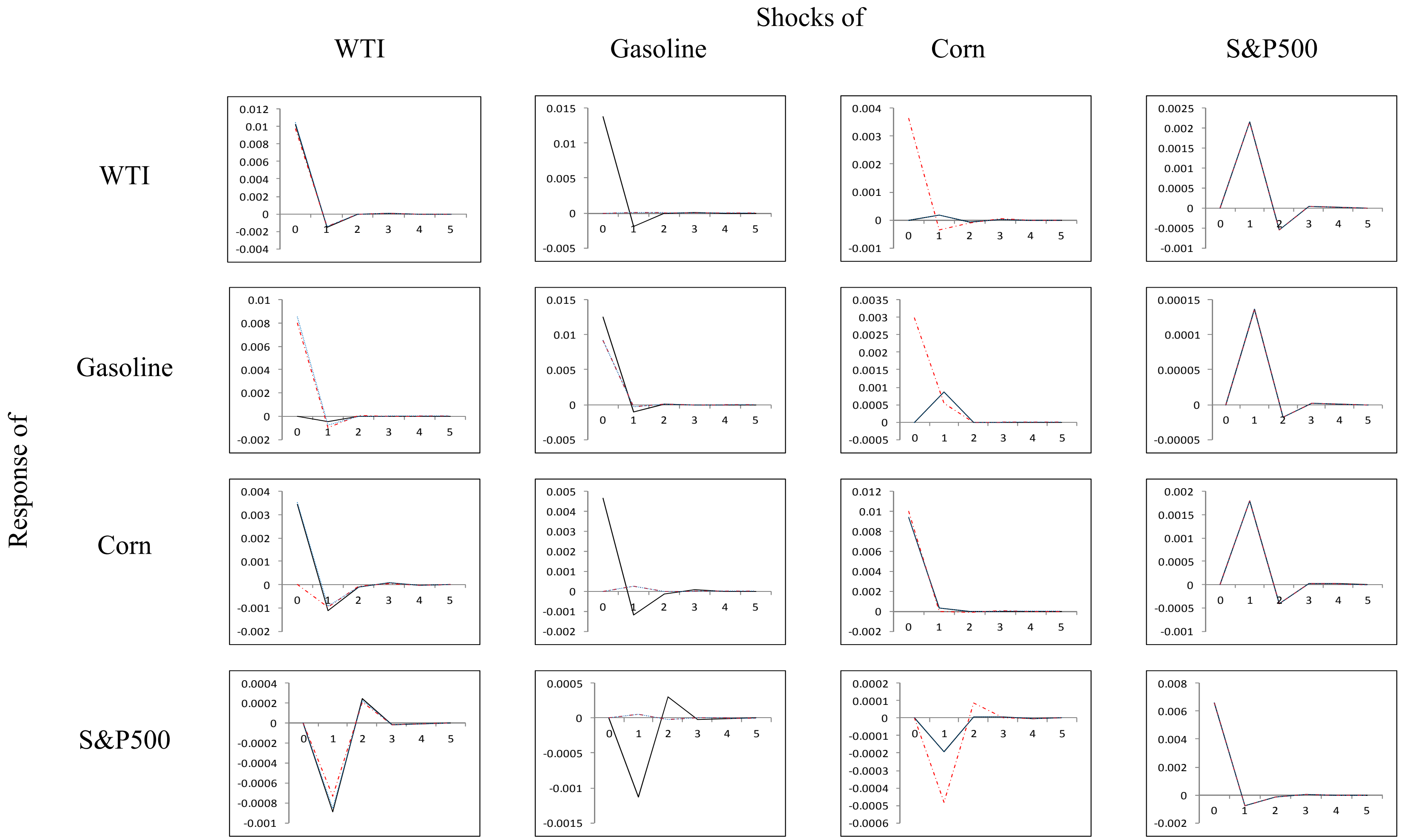


Figure 4.7. Impulse responses for sub-period 2 (2005/09/06 ~ 2008/12/16)
Note: Black solid line, red dashed line, and blue dotted line represent equivalent DAGs (a), (b), and (c), respectively.

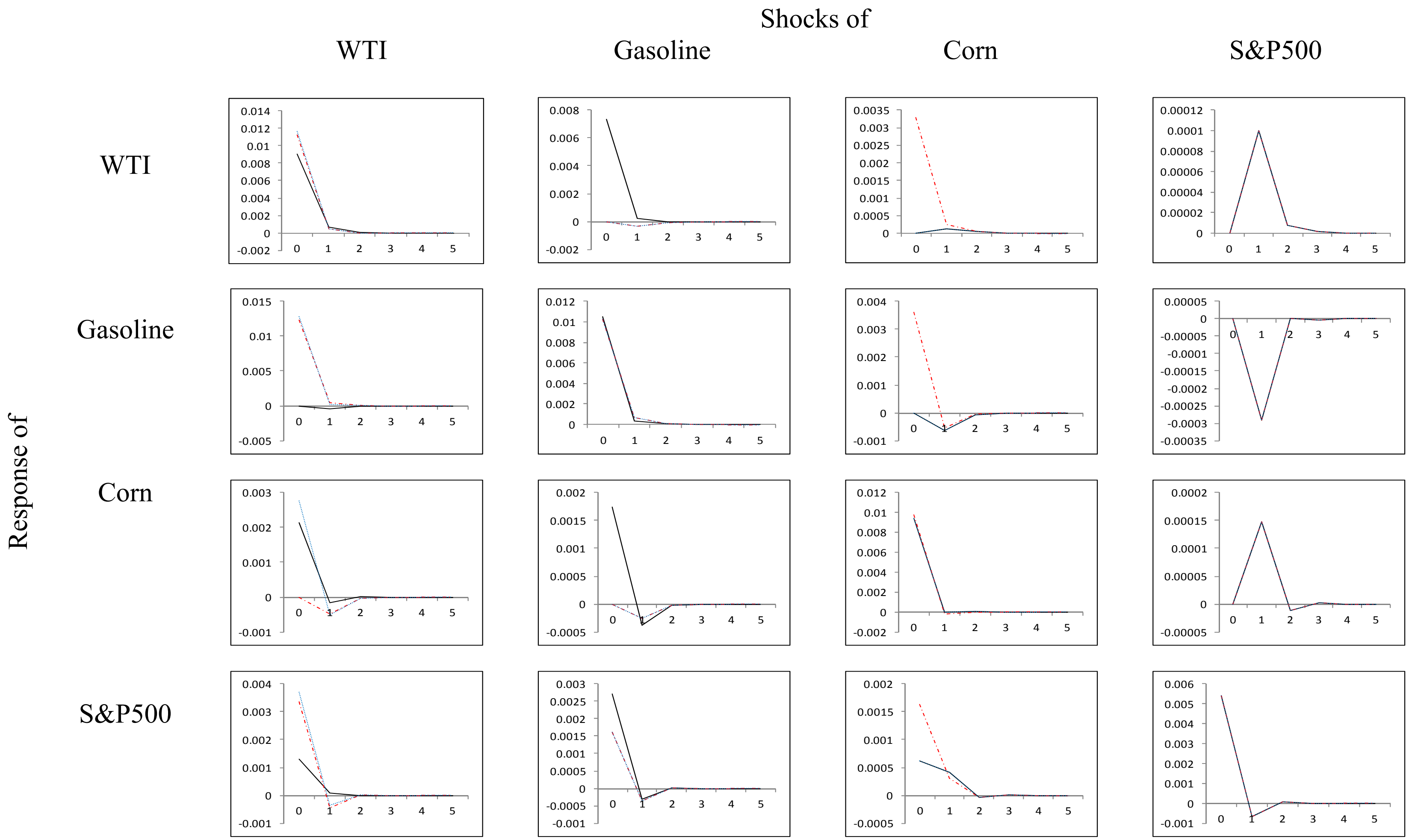


Figure 4.8. Impulse responses for sub-period 3 (2008/12/17 ~ 2011/12/30)
Note: Black solid line, red dashed line, and blue dotted line represented equivalent DAGs (d), (e), and (f), respectively.

As stated, we consider three equivalent DAG types for sub-periods 2 and 3. Thus, it matters how we model difference the contemporaneous causal structures among the price returns of WTI crude oil, gasoline, and corn. The results show that the contemporaneous and dynamic responses in all returns from one shock differ significantly between the sub-periods; hence, there are different responses in all returns from one shock in contemporaneous time among the three equivalent DAG types. However, dynamic responses in the returns of the three equivalent DAG types from one shock are mostly consistent for each sub-period. The specific descriptions for the IRFs are as follows.

A shock to the WTI crude oil price return does not affect the other price returns in sub-period 1 in contemporaneous time, whereas it positively affects the gasoline price returns in equivalent DAG type (b) and (c) in sub-period 2 and (e) and (f) in sub-period 3 in contemporaneous time. Similarly, a shock to the WTI crude oil price return positively affects the corn price returns in equivalent DAG type (a), (c), (d), and (f) and the S&P 500 returns in equivalent DAG type (d), (e), and (f) in contemporaneous time. Dynamically, a WTI crude oil price return shock moves the gasoline, corn, and S&P 500 returns in sub-period 1 up and down for several days before returning to normal; however this shock has no statistically significant effect. In contrast, dynamic responses in returns on gasoline, corn, and the S&P 500 from the WTI crude oil shock are consistent in sub-periods 2 and 3, i.e., they show negative instantaneous responses and recovery dynamics.

A shock to the gasoline price return positively affects the WTI crude oil price returns in sub-period 1 in contemporaneous time, whereas it positively affects the WTI crude oil price returns in equivalent DAG type (a) in sub-period 2 and (d) in sub-period 3. Similarly, a shock to the gasoline price return positively affects the corn price returns in equivalent DAG type (a) and (d), and the S&P 500 returns in equivalent DAG type (d), (e), and (f) in contemporaneous time. Dynamically, a gasoline price return shock moves the WTI crude oil and corn price returns in sub-period 1 down for several days before returning to normal, whereas it move the S&P 500 returns up. In contrast, dynamic responses in returns on WTI crude oil, corn, and the S&P 500 from the gasoline shock for equivalent DAG type (a) in sub-period 2 show negative instantaneous responses and recovery dynamics, whereas their dynamic responses from the gasoline shock for equivalent DAG type (b) and (c) in sub-period 2 show positive instantaneous responses and recovery dynamics. Similarly, dynamic responses in returns on WTI crude oil, corn, and the S&P 500 from the gasoline shock for all equivalent DAG types of sub-period 3, with the exception of the WTI crude oil in equivalent DAG type (d), drop initially and then return to normal.

A shock to the corn price return only affects the WTI crude oil price return in contemporaneous time for sub-period 1; however, it positively affects the WTI crude oil and gasoline price returns in equivalent DAG type (b) and (e) in sub-periods 2 and 3 and the S&P 500 returns in (d), (e), and (f) in sub-period 3 in contemporaneous time. On the other hand, the dynamic responses in return on the WTI crude oil, gasoline, corn, and the S&P 500 from the corn price return shock for each sub-period show dramatic variations. The WTI crude oil returns in sub-period 1 drop initially, then return to normal, whereas the WTI crude oil returns in all equivalent DAG types of sub-periods 2 and 3, with the exception of the equivalent DAG type (c) of sub-period 2, slightly increase. Gasoline price returns in sub-periods 1 and 2 are influenced positively for several days before returning to normal, whereas a corn price return shock negatively impacts gasoline price returns in all equivalent DAG types of sub-period 3. In contrast, the S&P 500 return in sub-periods 1 and 2 drops initially, whereas the dynamic responses in return on the S&P 500 for all equivalent DAG types in sub-period 3 are influenced positively.

As our previous DAG patterns in figures 4.2 through 4.4 show, a shock to the S&P 500 return does not affect the other price returns in contemporaneous time. Dynamically, there are no significant differences among the equivalent DAG types in sub-periods 2 and 3. Specifically, a shock to the S&P 500 return positively affects the WTI crude oil price return in sub-periods 1, 2, and 3. However, this shock positively influences gasoline price returns in sub-periods 1 and 2, but negatively impacts the prices in sub-period 3. The shock also negatively affects corn price returns in sub-period 1, but positively impacts the prices in sub-periods 2 and 3.

4.4.5. Forecast Error Variance Decompositions

Tables 4.9 through 4.11 contain the forecast error variance decompositions at horizons 0, 1, 2, and 12 days ahead under the contemporaneous causal ordering of the residuals of the VAR models for each sub-period as inferred by the DAG in figures 4.2

through 4.4. As before, (a), (b), (c), (d), (e), and (f) represent the equivalent DAGs types for sub-periods 2 and 3.

Table 4.9.
Forecast error variance decomposition for sub-period 1

Variance Decomposition of WTI					
Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
0	0.01838	37.122	62.752	0.126	0.000
1	0.01842	37.178	62.648	0.151	0.023
2	0.01842	37.178	62.648	0.151	0.023
12	0.01842	37.178	62.648	0.151	0.023
Variance Decomposition of Gasoline					
Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
0	0.01324	0.000	100.000	0.000	0.000
1	0.01328	0.571	99.384	0.002	0.044
2	0.01328	0.572	99.382	0.002	0.044
12	0.01328	0.572	99.382	0.002	0.044
Variance Decomposition of Corn					
Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
0	0.00694	0.000	0.000	100.000	0.000
1	0.00694	0.001	0.132	99.855	0.012
2	0.00694	0.002	0.132	99.855	0.012
12	0.00694	0.002	0.132	99.855	0.012
Variance Decomposition of S&P 500					
Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
0	0.00500	0.000	0.000	0.000	100.000
1	0.00500	0.001	0.038	0.001	99.960
2	0.00500	0.001	0.038	0.001	99.960
12	0.00500	0.001	0.038	0.001	99.960

Table 4.10.
Forecast error variance decomposition for sub-period 2

Variance Decomposition of WTI						
Equivalent DAGs type	Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
(a)	0	0.01718	35.285	64.715	0.000	0.000
	1	0.01749	34.781	63.678	0.011	1.530
	2	0.01750	34.748	63.615	0.013	1.624
	12	0.01750	34.748	63.614	0.013	1.625
(b)	0	0.01041	87.742	0.000	12.258	0.000
	1	0.01074	84.305	0.003	11.629	4.062
	2	0.01075	84.084	0.004	11.607	4.305
	12	0.01075	84.082	0.004	11.608	4.306
(c)	0	0.01041	100.000	0.000	0.000	0.000
	1	0.01074	95.907	0.003	0.028	4.062
	2	0.01075	95.658	0.004	0.034	4.305
	12	0.01075	95.657	0.004	0.034	4.306

Variance Decomposition of Gasoline						
Equivalent DAGs type	Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
(a)	0	0.01257	0.000	100.000	0.000	0.000
	1	0.01264	0.130	99.396	0.463	0.012
	2	0.01264	0.130	99.396	0.463	0.012
	12	0.01264	0.130	99.396	0.463	0.012
(b)	0	0.01257	40.391	53.966	5.643	0.000
	1	0.01262	40.654	53.552	5.783	0.012
	2	0.01262	40.654	53.551	5.783	0.012
	12	0.01262	40.654	53.551	5.783	0.012
(c)	0	0.01257	46.034	53.966	0.000	0.000
	1	0.01262	45.972	53.552	0.464	0.012
	2	0.01262	45.972	53.551	0.464	0.012
	12	0.01262	45.972	53.551	0.464	0.012

Table 4.10.
Continued

Variance Decomposition of Corn						
Equivalent DAGs type	Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
(a)	0	0.01103	9.726	17.838	72.436	0.000
	1	0.01130	10.255	18.108	69.094	2.543
	2	0.01131	10.249	18.095	68.982	2.666
	12	0.01132	10.253	18.099	68.982	2.666
(b)	0	0.01000	0.000	0.000	100.000	0.000
	1	0.01023	0.900	0.056	95.942	3.102
	2	0.01024	0.905	0.056	95.787	3.251
	12	0.01024	0.909	0.056	95.783	3.251
(c)	0	0.01000	12.258	0.000	87.742	0.000
	1	0.01023	12.572	0.056	84.270	3.102
	2	0.01024	12.561	0.056	84.131	3.251
	12	0.01024	12.565	0.056	84.127	3.251

Variance Decomposition of S&P 500						
Equivalent DAGs type	Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
(a)	0	0.00658	0.000	0.000	0.000	100.000
	1	0.00677	1.735	2.766	0.081	95.418
	2	0.00679	1.856	2.954	0.080	95.109
	12	0.00679	1.857	2.955	0.080	95.108
(b)	0	0.00658	0.000	0.000	0.000	100.000
	1	0.00668	1.207	0.008	0.516	98.269
	2	0.00668	1.306	0.009	0.532	98.153
	12	0.00668	1.307	0.009	0.532	98.153
(c)	0	0.00658	0.000	0.000	0.000	100.000
	1	0.00668	1.639	0.008	0.083	98.269
	2	0.00668	1.755	0.009	0.083	98.153
	12	0.00668	1.755	0.009	0.083	98.153

Table 4.11.
Forecast error variance decomposition for sub-period 3

Variance Decomposition of WTI						
Equivalent DAGs type	Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
(d)	0	0.01167	60.252	39.748	0.000	0.000
	1	0.01170	60.365	39.616	0.011	0.007
	2	0.01170	60.365	39.615	0.013	0.007
	12	0.01170	60.365	39.615	0.013	0.007
(e)	0	0.01167	92.083	0.000	7.917	0.000
	1	0.01169	92.961	0.089	7.942	0.007
	2	0.01169	92.958	0.092	7.943	0.007
	12	0.01169	92.958	0.092	7.943	0.007
(f)	0	0.01167	100.000	0.000	0.000	0.000
	1	0.01169	99.892	0.089	0.011	0.007
	2	0.01169	99.888	0.092	0.013	0.007
	12	0.01169	99.888	0.092	0.013	0.007

Variance Decomposition of Gasoline						
Equivalent DAGs type	Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
(d)	0	0.01052	0.000	100.000	0.000	0.000
	1	0.01056	0.210	99.347	0.368	0.076
	2	0.01056	0.212	99.339	0.373	0.076
	12	0.01056	0.213	99.339	0.373	0.076
(e)	0	0.01643	56.208	38.959	4.833	0.000
	1	0.01646	56.058	38.985	4.926	0.031
	2	0.01646	56.056	38.986	4.927	0.031
	12	0.01646	56.056	38.986	4.927	0.031
(f)	0	0.01643	61.041	38.959	0.000	0.000
	1	0.01646	60.883	38.985	0.151	0.031
	2	0.01646	60.830	38.986	0.153	0.031
	12	0.01646	60.830	38.986	0.153	0.031

Table 4.11.
Continued

Variance Decomposition of Corn						
Equivalent DAGs type	Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
(d)	0	0.00979	4.770	3.147	92.084	0.000
	1	0.00980	4.783	3.288	91.906	0.022
	2	0.00980	4.783	3.288	91.906	0.023
	12	0.00980	4.783	3.288	91.906	0.023
(e)	0	0.00979	0.000	0.000	100.000	0.000
	1	0.00981	0.237	0.065	99.675	0.022
	2	0.00981	0.238	0.065	99.674	0.023
	12	0.00981	0.238	0.065	99.674	0.023
(f)	0	0.00979	7.917	0.000	92.083	0.000
	1	0.00981	8.152	0.065	91.760	0.022
	2	0.00981	8.153	0.065	91.759	0.023
	12	0.00981	8.153	0.065	91.759	0.023

Variance Decomposition of S&P 500						
Equivalent DAGs type	Period	Std. Errors	WTI	Gasoline	Corn	S&P 500
(d)	0	0.00620	4.419	19.017	0.988	75.576
	1	0.00626	4.358	18.907	1.406	75.330
	2	0.00626	4.357	18.902	1.409	75.332
	12	0.00626	4.357	18.902	1.409	75.332
(e)	0	0.00675	24.821	5.639	5.817	63.723
	1	0.00682	24.791	5.815	5.904	63.490
	2	0.00682	24.787	5.814	5.905	63.494
	12	0.00682	24.787	5.814	5.905	63.494
(f)	0	0.00675	29.805	5.639	0.833	63.723
	1	0.00682	29.510	5.815	1.185	63.490
	2	0.00682	29.504	5.814	1.188	63.494
	12	0.00682	29.504	5.814	1.188	63.494

For sub-period 1, the variances of WTI crude oil are mainly explained by the residuals of gasoline (62.65 to 62.75%) and own contribution (37.12 to 37.18%). By contrast, own contribution (99.38 to 100% for gasoline, 99.85 to 100% for corn, and 99.96 to 100% for the S&P 500) appear to be important for explaining the variations of gasoline, corn, and the S&P 500.

We consider three equivalent DAG types in sub-period 2. Thus, it matters how we model the causal structures among the price returns of WTI crude oil, gasoline, and corn in contemporaneous time. However, the influence of the S&P 500 return on the other three price returns is usually less than 5% under all three equivalent DAG types. For the variation in WTI crude oil price returns, the residual of gasoline (63.61% to 64.72%) and own contribution (34.75% to 35.29%) appear to be important for equivalent DAG type (a). The variation in WTI crude oil price returns is primarily explained by the residuals of corn (11.61% to 12.26%) and own contribution (84.08% to 87.74%) for equivalent DAG type (b). On the other hand, the variation in WTI crude oil price returns is explained by own contribution (95.66% to 100%) in equivalent DAG type (c). For the variation in gasoline, own contribution (99.40% to 100% for equivalent DAG type (a) and 53.55% to 54.0% for equivalent DAG type (b) and (c)) appear to be important in all three equivalent DAG types, whereas the residuals of WTI crude oil (40.39% to 46.03%) and corn (0.46% to 5.78%) are accounted for by the variation in gasoline price returns in equivalent DAG type (b) and (c). The variation in corn price returns is mostly explained by its own residuals (68.98% to 72.44% for equivalent DAG type (a), 95.78% to 54.0% for equivalent DAG type (b), and 84.13% to 87.74% for equivalent DAG type (c)). Also, the residuals of WTI crude oil (9.73% to 10.26% for equivalent DAG type (a), 12.26% to 12.57% for equivalent DAG type (c)) and gasoline (17.84% to 18.11% for equivalent DAG type (a)) appear to be important for explaining the variation in corn price returns. For the uncertainty associated with the S&P 500 return, the variation is mainly explained by the residual of own contribution (95.11% to 100%) for the three equivalent DAG types.

Similarly, in sub-period 3, the variation in WTI crude oil price returns is explained by the residual of gasoline (39.62% to 39.75%) and own contribution (60.25% to 60.37%) for equivalent DAG type (d). The variation in WTI crude oil price return is primarily explained by the residuals of corn (7.92% to 7.94%) and own contribution (92.08% to 92.96%) for equivalent DAG type (e), whereas its own residuals (99.89 to 100%) appear to be important for equivalent DAG type (f). For the variation in gasoline price returns, own contribution (99.34% to 100% for equivalent DAG type (d) and 38.96% to 38.99% for equivalent DAG type (e) and (f)) appears to be important for the three equivalent DAG types. Moreover, the residuals of WTI crude oil (56.06% to 56.21%) and corn (0% to 4.93%) account for the variation in gasoline for DAG type (e) and (f). The variation in corn price returns is mostly explained by own residuals (91.91% to 92.08% for equivalent DAG type (d), 99.67% to 100% for equivalent DAG type (e) and 91.76% to 92.08% for equivalent DAG type (f)). Additionally, the residuals of WTI crude oil (4.77% to 4.78% for equivalent DAG type (d), 7.92% to 8.15% for equivalent DAG type (f)), and gasoline (3.15% to 3.29% for equivalent DAG type (d)) appear to be important for explaining the variation in corn price returns. For the variation in the S&P

500 returns, the variation is mostly explained by the residual of own contribution (63.49 to 75.58%) for equivalent DAG type (d), (e), and (f). However, the residuals of WTI crude oil (4.36% to 29.81%), gasoline (5.64% to 19.02%), and corn (0.83% to 5.91%) also appear to be important for the three equivalent DAG types, contrary to sub-periods 1 and 2.

4.5. Conclusion

In this chapter, we built an econometric model and applied the VAR approach to analyze the dynamic causal relationships among the price returns of WTI crude oil, gasoline, corn, and the S&P 500. To investigate the empirical contemporaneous causal relationships, we used the Directed Acyclic Graph (DAG) approach following Bessler and Lee (2002), Bessler and Yang (2003), Demiralp and Hoover (2003), Moneta (2004), Moneta (2008), Swanson and Granger (1997), and Kim and Bessler (2007). We also implemented innovation accounting time series techniques, such as forecast error variance decomposition and an impulse response function, to analyze the dynamic information transmission among the price returns of WTI crude oil, gasoline, corn, and the S&P 500.

Most important, we tested for structural breaks by using the procedure developed by Bai and Perron (1998) and Bai and Perron (2003). In our empirical application, we found strong evidence of structural breaks in the VAR models of the WTI crude oil, gasoline, corn, and the S&P 500 returns. Additional research found that the two structural break points identified in August 2005 and September 2008 related to

Hurricane Katrina, ethanol production or the Lehman Brothers bankruptcy. Based on the structural break test results, we divided the full sample period into three sub-periods, which enabled us to find consistent sub-sample periods having stable parameters in a given VAR framework.

Using the DAG results, we found strong contemporaneous causal relationships among the residuals from the VAR models of the WTI crude oil, gasoline, corn, and S&P 500 returns. There were significant differences between contemporaneous causal structures for each sub-period. Finding two undirected edges (WTI-gasoline and WTI-corn) in sub-periods 2 and 3 led us to develop three equivalent DAG types for them.

In terms of innovation discovery there are two key findings: (1) WTI crude oil changed from a causal sink in sub-period 1 to a causal parent or channel in sub-periods 2 and 3. This suggests that the WTI crude oil price return has an exogenous role or channel role to transmit the causal influences from the gasoline/agricultural part into the agricultural/gasoline part in contemporaneous time after the first structural break; (2) The S&P 500 return is isolated from other returns in sub-periods 1 and 2, and changed to a causal sink in sub-period 3. This suggests that the S&P 500 return has an endogenous role in contemporaneous time after the second structural break.

From the innovation accounting analysis based on our VAR models for each sub-period, we recognize that the contemporaneous and dynamic responses in returns on the WTI crude oil, gasoline, corn, and S&P 500 from one shock significantly differ between each sub-period. These results also show significantly different responses in each return from one shock in contemporaneous time on all the equivalent DAG types in sub-period

2 and 3. This is consistent with our DAG results. However, dynamic responses in returns on all equivalent DAG types from one shock are mostly consistent for each sub-period.

We assert that identifying contemporaneous causal relationships and their dynamic variations provides important information for future studies of market linkage and/or market integration among energy, agricultural and financial markets, and indeed, other commodity markets. Such information can inform future research related to identifying and verifying the role of markets in aspects of dynamics and causalities.

CHAPTER V

CONCLUSIONS

We have investigated establishing empirical time series models for price dynamics and causation among energy prices, macroeconomic, and financial indicators in this dissertation. We focused on three issues: (1) the contemporaneous interdependencies and information flows among crude oil, natural gas, and electricity prices in the US; (2) the federal fund rate and WTI crude oil price dynamics with macroeconomic and financial indicators and their related information transmission mechanisms; and (3) structural change in mean equations among US energy, agricultural, and financial markets considering contemporaneous causality, dynamics, and structural change with unknown break points.

In Chapter II, we estimated a causal model for the price dynamics for contemporaneous relationships by using the daily price returns of Dated Brent crude oil, Henry Hub natural gas, PJM Electricity firm on peak and COB electricity firm on peak. We assessed both standardized residuals from within-sample-fit and out-of-sample-forecast for modeling information flows in contemporaneous time using the VAR-DCC-GARCH models. From these processes as well as comparing forecast performance and variance-covariance matrices, we found that the within-sample-fit and out-of-sample-forecast statistically perform similarly, whereas the within-sample-fit model generally outperforms the out-of-sample-forecast and contains small elements in variance-covariance matrix. Using the PC algorithm in TETRAD IV, we demonstrated that both methods for calculating the standardized residuals show the same graphical patterns. However, the price returns of PJM and COB electricity firm on peak in the DAG revealed an ambiguous direction of information flows from the given information. Therefore, we confirm the proposition that causal flows based on both residuals from within-sample-fit and out-of-sample-forecast exhibit consistency. Moreover, the test results for homogeneity of variance-covariance matrices and forecast performance (accuracy) support this conclusion. As a consequence, we can be confident in the out-of-sample-forecast and its causal results.

In Chapter III, we built econometric models for the federal fund rate and the WTI crude oil price return using a large panel of macroeconomic time series. We summarized the information with a few estimated factors using PCA, which allowed us to interpret and identify the underlying factors. We augmented these factors as regressors in a VAR framework to assess the effects of the federal fund rate and WTI crude oil price shocks in the US. Using the GES algorithm, we inductively inferred the contemporaneous causal relationships among innovations of both FAVAR models. We found that the federal fund rate shock is exogenous in contemporaneous time as the identifying assumption in the VAR framework of the monetary shock transmission mechanism. This finding is consistent with the identification assumption of Bernanke et al. (2005) and the stylized fact of Kwon (2007). However, we found that the WTI crude oil price return is not exogenous in contemporaneous time. Thus, we argue that the oil price shock transmission mechanisms identified in this causal information can be inferred from the data. Our innovation accounting results generally align with previous literature (Bernanke et al. 2005; Christiano et al. 1999; Cuaresma et al. 2004; Estrella 2005; Fair 2002; Kwon 2007; Tufte and Wohar 1999) and appear to make economic sense. Notably, we identify that the price puzzle (Sims 1992) is considerably reduced and prices eventually drop to the federal fund rate shocks as Bernanke et al. (2005) and Kwon (2007) found. Thus, we conclude that the inclusion into the model of the information captured by the factors succeeds in mitigating the price puzzle. Consequently, it is advantageous to use the larger macroeconomic information set for analyzing monetary policy and oil price shock transfer mechanisms; the results from our out-of-sample-forecasts of the federal fund rate and WTI crude oil price return emphasize this finding. The forecasting performance of FAVAR models based on common factors clearly outperforms a model based on individual variables. More importantly, the FAVAR model shows a striking superiority with respect to univariate AR models in out-of-sample-forecasts.

In Chapter IV, we built an econometric model and applied the VAR approach to analyze the dynamic causal relationships among the returns of WTI crude oil, gasoline, corn, and the S&P 500. We tested for structural breaks using the procedure developed by Bai and Perron (1998) and Bai and Perron (2003). Empirically, we found strong evidence of two structural breaks. Based on the structural break test results, we divided the full sample period into three sub-periods in order to identify consistent sub-sample periods having stable parameters in a given VAR framework. We found strong contemporaneous causal relationships among the residuals from the VAR models and significant differences between contemporaneous causal structures for each sub-period.

The discovery of two undirected edges (WTI-gasoline and WTI-corn) in sub-periods 2 and 3 prompted us to provide three equivalent DAG types for the two undirected edges. Although it makes little difference how the contemporaneous causal structures among the four price returns are modeled, we made several key findings in terms of innovation discovery: (1) The WTI crude oil price return change from a causal sink in sub-period 1 to a causal parent or channel in sub-periods 2 and 3 suggests that this price return has an exogenous role or channel role to transmit the causal influences from the gasoline/agricultural part into the agricultural/gasoline part in contemporaneous time after the first break; (2) The isolated S&P 500 return in sub-periods 1 and 2, which then changes to a causal sink, suggests that this price return has an endogenous role in contemporaneous time after the second break. Moreover, the contemporaneous and dynamic responses on the WTI crude oil, gasoline, corn, and S&P 500 returns from one shock differ significantly between sub-periods. Our finding of significantly different responses in each return from one shock in contemporaneous time among each equivalent DAG type for each sub-period coincides with our DAG results. However, the dynamic responses in returns on all equivalent DAG types from one shock are mostly consistent for each period.

This dissertation has two limitations. First, the PC algorithm assumes no latent common causes of the considered variables. An alternative is the FCI algorithm (Spirtes et al. (2000), which does not make this assumption, although we note that its conclusions are weaker than the PC algorithm. Second, we only consider four energy spot prices for identifying the proposition that causal flows based on both residuals from within-sample-fit and out-of-sample-forecast exhibit consistency. We suggest that future research on other US commodity markets can obtain more comprehensive results.

REFERENCES

Andrews, D.W.K. 1993. "Tests for Parameter Instability and Structural-Change with Unknown Change-Point." *Econometrica* 61(4):821-856.

Andrews, D.W.K., and W. Ploberger. 1994. "Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative." *Econometrica* 62(6):1383-1414.

Bai, J.S, and P. Perron. 2003. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics* 18(1):1-22.

Bai, J.S., and S. Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70(1):191-221.

Bai, J.S., and P. Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66(1):47-78.

Barsky, R., and L. Kilian. 2002. *Do We Really Know That Oil Caused the Great Stagflation? A Monetary Alternative*: The MIT Press, Cambridge, MA, USA.

Barsky, R., B. Robert, and L. Kilian. 2004. "Oil and the Macroeconomy since the 1970s." *Journal of Economic* 18(4):115-134.

Bernanke, B.S. 1986. "Alternative Explanations of the Money-Income Correlation." NBER Working Paper 1842, National Bureau of Economic Research.

Bernanke, B.S., and J. Boivin. 2003. "Monetary Policy in a Data-Rich Environment." *Journal of Monetary Economics* 50(3):525-546.

Bernanke, B.S., J. Boivin, and P. Eliasz. 2005. "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics* 120(1):387-422.

Bernanke, B.S., M. Gertler, and M. Watson. 1997. "Systematic Monetary Policy and the Effects of Oil Price Shocks." *Brooking Papers on Economic Activity* 1997(1):91-157.

Bessler, D.A., and S. Lee. 2002. "Money and Prices: US Data 1869–1914 (a Study with Directed Graphs)." *Empirical Economics* 27(3):427-446.

Bessler, D.A., and J. Yang. 2003. "The Structure of Interdependence in International Stock Markets." *Journal of International Money and Finance* 22(2):261-287.

Bessler, D.A., J. Yang, and M. Wongcharupan. 2003. "Price Dynamics in the International Wheat Market: Modeling with Error Correction and Directed Acyclic Graphs." *Journal of Regional Science* 43(1):1-33.

Bianchi, F., H. Mumtaz, and P. Surico. 2009. "The Great Moderation of the Term Structure of UK Interest Rates." *Journal of Monetary Economics* 56(6):856-871.

Blanco, C., and D. Soronow. 2001. "Jump Diffusion Processes-Energy Price Processes Used for Derivatives Pricing and Risk Management." *Commodities Now* 37:83-87.

Boivin, J., M.P. Giannoni, and I. Mihov. 2009. "Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data." *American Economic Review* 99(1):350-384.

Bollerslev, T. 1990. "Modeling the Coherence in Short-Run Nominal Exchange-Rates - a Multivariate Generalized Arch Model." *Review of Economics and Statistics* 72(3):498-505.

Box, G.E. 1949. "A General Distribution Theory for a Class of Likelihood Criteria." *Biometrika* 36(3-4):317-346.

Box, G.E., G.M. Jenkins, and G.C. Reinsel. 1976. *Time Series Analysis*: Holden-day, San Francisco, CA, USA.

Brown, R., J. Durbin, and J.M. Evans. 1975. "Techniques for Testing the Constancy of Regression Relationships over Time." *Journal of the Royal Statistical Society. Series B (Methodological)* 37(2):149-192.

Brown, S., and M.K. Yücel. 2002. "Energy Prices and Aggregate Economic Activity: An Interpretative Survey." *Quarterly Review of Economics and Finance* 42(2):193-208.

Cantwell, M. 2012. Statement of Senator Cantwell on Oil Should Be Based on Supply and Demand, Not Wall Street Bets: Press Release.

Chatfield, C. 2001. *Time-Series Forecasting*: CRC Press, London, England.

Chickering, D.M. 2002. "Learning Equivalence Classes of Bayesian-Network Structures." *Journal of Machine Learning Research* 2(3):445-498.

---. 2003. "Optimal Structure Identification with Greedy Search." *Journal of Machine Learning Research* 3:507-554.

Chow, G.C. 1960. "Tests of Equality between Sets of Coefficients in Two Linear Regressions." *Econometrica: Journal of the Econometric Society* 28(3):591-605.

Christiano, L.J., M. Eichenbaum, and C.L. Evans. 1999. "Monetary Policy Shocks: What Have We Learned and to What End?" *Handbook of Macroeconomics* 1:65-148.

Clemen, R.T. 1989. "Combining Forecasts - a Review and Annotated-Bibliography." *International Journal of Forecasting* 5(4):559-583.

Clements, M.P., and D.F. Hendry. 1993. "On the Limitations of Comparing Mean Square Forecast Errors." *Journal of Forecasting* 12(8):617-637.

Cuaresma, J., E. Gnan, and D. Ritzberger-Gruenwald. 2004. "Searching for the Natural Rate of Interest: A Euro Area Perspective." *Empirica* 31(2):185-204.

Demiralp, S., and K.D. Hoover. 2003. "Searching for the Causal Structure of a Vector Autoregression." *Oxford Bulletin of Economics and Statistics* 65:745-767.

Diebold, F.X., and J.A. Lopez. 1996. "Forecast Evaluation and Combination." NBER Working Paper 192, National Bureau of Economic Research.

Diebold, F.X., and R.S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 13(3):253-263.

Eickmeier, S., W. Lemke, M. Marcellino, and C.F.E.P. Research. 2011. *The Changing International Transmission of Financial Shocks: Evidence from a Classical Time-Varying Favar*: Centre for Economic Policy Research, London, UK.

Elliott, G., T.J. Rothenberg, and J.H. Stock. 1996. "Efficient Tests for an Autoregressive Unit Root." *Econometrica* 64(4):813-836.

Engle, R. 2001. "Garch 101: The Use of ARCH/GARCH Models in Applied Econometrics." *Journal of Economic Perspectives* 15(4):157-168.

---. 2002. "Dynamic Conditional Correlation—a Simple Class of Multivariate Garch Models." *Journal of Business and Economic Statistics* 20:339-350.

Engle, R., and J. Mezrich. 1996. "Garch for Groups: A Round-up of Recent Developments in Garch Techniques for Estimating Correlation." *Risk* 9:36-40.

Engle, R., and K. Sheppard. 2001. "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate Garch." NBER Working Paper 8554, National Bureau of Economic Research.

Estrella, A. 2005. "Why Does the Yield Curve Predict Output and Inflation?" *Economic Journal* 115(505):722-744.

Fair, R.C. 2002. "On Modeling the Effects of Inflation Shocks." *Journal of Macroeconomics* 2(1):article 3.

Forni, M., and L. Gambetti. 2010. "The Dynamic Effects of Monetary Policy: A Structural Factor Model Approach." *Journal of Monetary Economics* 57(2):203-216.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin. 2000. "The Generalized Dynamic-Factor Model: Identification and Estimation." *Review of Economics and Statistics* 82(4):540-554.

---. 2003. "Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area?" *Journal of Monetary Economics* 50(6):1243-1255.

---. 2004. "The Generalized Dynamic Factor Model Consistency and Rates." *Journal of Econometrics* 119(2):231-255.

Forni, M., and M. Lippi. 2001. "The Generalized Dynamic Factor Model: Representation Theory." *Econometric Theory* 17(6):1113-1141.

Forni, M., and L. Reichlin. 1998. "Let's Get Real: A Factor Analytical Approach to Disaggregated Business Cycle Dynamics." *Review of Economic Studies* 65(3):453-473.

Gilchrist, S., V. Yankov, and E. Zakrajsek. 2009. "Credit Market Shocks and Economic Fluctuations: Evidence from Corporate Bond and Stock Markets." *Journal of Monetary Economics* 56(4):471-493.

Granger, C.W.J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37(3):424-438.

---. 1980. "Testing for Causality - a Personal Viewpoint." *Journal of Economic Dynamics & Control* 2(4):329-352.

---. 1988. "Some Recent Development in a Concept of Causality." *Journal of Econometrics* 39(1-2):199-211.

---. 1989. "Invited Review Combining Forecasts—Twenty Years Later." *Journal of Forecasting* 8(3):167-173.

Greene, W.H. 2007. *Econometric Analysis 6th Edition*: Prentice Hall, New Jersey, NJ, USA.

Haan, W.J., and A.T. Levin. 2000. "Robust Covariance Matrix Estimation with Data-Dependent VAR Prewhitening Order." NBER Working Paper 255, National Bureau of Economic Research.

Hafner, C.M., and P.H.B.F. Franses. 2003. "A Generalized Dynamic Conditional Correlation Model for Many Asset Returns." Econometric Institute Report EI 2003-18, Erasmus School of Economics.

Hamilton, J.D. 1983. "Oil and the Macroeconomy since World-War-II." *Journal of Political Economy* 91(2):228-248.

---. 1996. "This Is What Happened to the Oil Price Macroeconomy Relationship." *Journal of Monetary Economics* 38(2):215-220.

---. 2003. "What Is an Oil Shock?" *Journal of Econometrics* 113(2):363-398.

Hanson, K., S. Robinson, and G. Schluter. 1993. "Sectoral Effects of a World Oil Price Shock: Economywide Linkages to the Agricultural Sector." *Journal of Agricultural and Resource Economics* 18(1):96-116.

Harri, A., L. Nalley, and D. Hudson. 2009. "The Relationship between Oil, Exchange Rates, and Commodity Prices." *Journal of Agricultural and Applied Economics* 41(2):501-510.

Harvey, D., S. Leybourne, and P. Newbold. 1997. "Testing the Equality of Prediction Mean Squared Errors." *International Journal of Forecasting* 13(2):281-291.

Helbling, T., R. Huidrom, M. Ayhan Kose, and C. Otrok. 2010. "Do Credit Shocks Matter? A Global Perspective." *European Economic Review* 55(3):340-353.

Hickman, B.G., H.G. Huntington, and J.L. Sweeney. 1987. *Macroeconomic Impacts of Energy Shocks*: Elsevier Science Pub. Co. Inc.,New York, NY, USA.

Hoover, K.D. 2005. "Automatic Inference of the Contemporaneous Causal Order of a System of Equations." *Econometric Theory* 21(1):69-77.

Hoover, K.D., and S.J. Perez. 1994. "Post Hoc Ergo Propter Once More an Evaluation of 'Does Monetary Policy Matter?' In the Spirit of James Tobin." *Journal of Monetary Economics* 34(1):47-74.

Jiménez-Rodríguez, R., and M. Sánchez. 2005. "Oil Price Shocks and Real GDP Growth: Empirical Evidence for Some Oecd Countries." *Applied Economics* 37(2):201-228.

Johansen, S. 1991. "Estimation and Hypothesis-Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica* 59(6):1551-1580.

Jones, D.W., P.N. Leiby, and I.K. Paik. 2004. "Oil Price Shocks and the Macroeconomy: What Has Been Learned since 1996." *Energy Journal* 25(2):1-32.

Kaltalioglu, M., and U. Soytas. 2009. "Price Transmission between World Food, Agricultural Raw Material, and Oil Prices." Paper presented at GBATA meeting. Prague, Czech Republic.

Kilian, L. 2008. "Exogenous Oil Supply Shocks: How Big Are They and How Much Do They Matter for the US Economy?" *Review of Economics and Statistics* 90(2):216-240.

Kim, J.W., and D.A. Bessler. 2007. "The Causal Modelling on Equity Market Innovations: Fit or Forecast?" *Applied Financial Economics* 17(8):635-646.

Krämer, W., W. Ploberger, and R. Alt. 1988. "Testing for Structural Change in Dynamic Models." *Econometrica: Journal of the Econometric Society* 56(6):1355-1369.

Kwon, D.H. 2007. "Causality and Aggregation in Economics: The Use of High Dimensional Panel Data in Micro-Econometrics and Macro-Econometrics."

Unpublished Ph. D. dissertation, Department of Agricultural Economics, Texas A&M University, College Station, TX.

Lagana, G., and A. Mountford. 2005. "Measuring Monetary Policy in the UK: A Factor-Augmented Vector Autoregression Model Approach." *The Manchester School* 73:77-98.

Lagana, G., and P.M. Sgro. 2011. "A Factor-Augmented VAR Approach: The Effect of a Rise in the US Personal Income Tax Rate on the US and Canada." *Economic Modelling* 28(3):1163-1169.

LaMotte, L.R., and A. McWhorter Jr. 1978. "An Exact Test for the Presence of Random Walk Coefficients in a Linear Regression Model." *Journal of the American Statistical Association* 73(364):816-820.

Leeper, E.M., C.A. Sims, T. Zha, R.E. Hall, and B.S. Bernanke. 1996. "What Does Monetary Policy Do?" *Brooking Papers on Economic Activity* 1996(2):1-78.

Lütkepohl, H. 1999. *Vector Autoregressions*: Blackwell Publishers, Wiley Online Library.

Lütkepohl, H., and H.E. Reimers. 1992. "Granger-Causality in Cointegrated VAR Processes the Case of the Term Structure." *Economics Letters* 40(3):263-268.

Mardia, K.V., and J.T. Kent. 1979. *Multivariate Analysis*: Academic Press, San Diego, CA, USA.

Meek, C. 1997. "Graphical Models: Selecting Causal and Statistical Models." Unpublished Ph. D. dissertation, Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA.

Moench, E., and S. Ng. 2011. "A Hierarchical Factor Analysis of US Housing Market Dynamics." *Econometrics Journal* 14(1):C1-C24.

Moneta, A. 2004. "Identification of Monetary Policy Shocks: A Graphical Causal Approach." *Notas Económicas* 20:39-62.

---. 2008. "Graphical Causal Models and Vars: An Empirical Assessment of the Real Business Cycles Hypothesis." *Empirical Economics* 35(2):275-300.

Mork, K.A., and R.E. Hall. 1980. "Energy Prices, Inflation, and Recession, 1974-1975." *Energy Journal* 1(3):31-64.

Mutuc, M., S. Pan, and D. Hudson. 2010. "Response of Cotton to Oil Price Shocks." Paper presented at SAEA meeting. Birmingham, AL, USA.

NASS. 2011. National Agricultural Statistics Service, vol. 2011.

Ng, S., and E. Moench. 2009. "A Hierarchical Factor Analysis of US Housing Market Dynamics." *Econometrics Journal* 14(1):C1-C24.

Nyblom, J. 1989. "Testing for the Constancy of Parameters over Time." *Journal of the American Statistical Association* 84(405):223-230.

Nyblom, J., and T. Makelainen. 1983. "Comparisons of Tests for the Presence of Random-Walk Coefficients in a Simple Linear-Model." *Journal of the American Statistical Association* 78(384):856-864.

Pearl, J. 1985. *A Constraint-Propagation Approach to Probabilistic Reasoning*: Computer Science Department, University of California, Los Angeles, CA, USA.

---. 1995. "Causal Diagrams for Empirical Research." *Biometrika* 82(4):669-688.

---. 2000. *Causality: Models, Reasoning and Inference*: Cambridge University Press, Cambridge, England.

Ploberger, W., and W. Kramer. 1992. "The Cusum Test with Ols Residuals." *Econometrica* 60(2):271-285.

Ploberger, W., W. Krämer, and R. Alt. 1989. "A Modification of the Cusum Test in the Linear Regression Model with Lagged Dependent Variables." *Empirical Economics* 14(2):65-75.

Quandt, R.E. 1960. "Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes." *Journal of the American Statistical Association* 55(290):324-330.

Rasche, R.H., and J.A. Tatom. 1977. "The Effects of the New Energy Regime on Economic Capacity, Production, and Prices." *Federal Reserve Bank of St. Louis Review* 59(4):2-12.

RFA. 2011. Renewable Fuels Association, vol. 2011.

Sadorsky, P. 1999. "Oil Price Shocks and Stock Market Activity." *Energy Economics* 21(5):449-469.

Sims, C.A. 1980. "Macroeconomics and Reality." *Econometrica* 48(1):1-48.

---. 1986. "Are Forecasting Models Usable for Policy Analysis?" *Federal Reserve Bank of Minneapolis Quarterly Review* 10(1):2-16.

---. 1992. "Interpreting the Macroeconomic Time-Series Facts - the Effects of Monetary-Policy." *European Economic Review* 36(5):975-1000.

Soytas, U., R. Sari, S. Hammoudeh, and E. Hacihasanoglu. 2009. "World Oil Prices, Precious Metal Prices and Macroeconomy in Turkey." *Energy Policy* 37(12):5557-5566.

Spirtes, P., C.N. Glymour, and R. Scheines. 2000. *Causation, Prediction, and Search*: The MIT Press, Cambridge, MA, USA.

Stock, J.H., and M.W. Watson. 2001. "Vector Autoregressions." *Journal of Economic Perspectives* 15(4):101-115.

---. 2002a. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97(460):1167-1179.

---. 2002b. "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business & Economic Statistics* 20(2):147-162.

---. 2005. "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper 11467, National Bureau of Economic Research.

Swanson, N.R., and C.W.J. Granger. 1997. "Impulse Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions." *Journal of the American Statistical Association* 92(437):357-367.

Tse, Y.K., and A.K.C. Tsui. 2002. "A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations." *Journal of Business & Economic Statistics* 20(3):351-362.

Tufte, D., and M.E. Wohar. 1999. "Models with Unexpected Components: The Case for Efficient Estimation." *Review of Quantitative Finance and Accounting* 13(3):295-313.

Yu, T.H., D.A. Bessler, and S. Fuller. 2006. "Cointegration and Causality Analysis of World Vegetable Oil and Crude Oil Prices." Paper presented at AAEA meeting. Long Beach, CA, USA

Zhang, Z., L. Lohr, C. Escalante, and M. Wetzstein. 2010. "Food Versus Fuel: What Do Prices Tell Us?" *Energy Policy* 38(1):445-451.

APPENDIX A

DATA DESCRIPTION

Table A.1 lists the short name of each series, the transformation applied to the series, and a brief data description. In the transformation column (Tcode), 1 denotes no transformation, 2 denotes first difference, 3 denotes second differences, 4 denotes logarithm, 5 denotes first difference of logarithm and 6 denotes second difference of logarithm. In the slow code column, 1 denotes a variable assumed to be "slow-moving" in the estimation, 0 denotes otherwise.

No.	Series ID	Tcode	Slow Code	Description
1	PI	5	1	Personal Income
2	PI less transfers	5	1	Personal Income Less Transfer Payments
3	Consumption	5	1	Real Consumption
4	M&T sales	5	1	Manufacturing & Trade Sales
5	Retail sales	5	1	Sales of Retail Stores
6	IP:total	5	1	Industrial Production Index - Total Index
7	IP: products	5	1	Industrial Production Index - Products, Total
8	IP: final prod	5	1	Industrial Production Index - Final Products
9	IP: cons gds	5	1	Industrial Production Index - Consumer Goods
10	IP: cons dble	5	1	Industrial Production Index - Durable Consumer Goods
11	IP: cons nondble	5	1	Industrial Production Index - Nondurable Consumer Goods
12	IP: bus eqpt	5	1	Industrial Production Index - Business Equipment

13	IP: matls	5	1	Industrial Production Index - Materials
14	IP: dble matls	5	1	Industrial Production Index - Durable Goods Materials
15	IP: nondble matls	5	1	Industrial Production Index - Nondurable Goods Materials
16	IP: mfg	5	1	Industrial Production Index - Manufacturing
17	IP: res util	5	1	Industrial Production Index - Residential Utilities
18	IP: fuels	5	1	Industrial Production Index - Fuels
19	NAPM prodn	1	1	Napm Production Index (Percent)
20	Cap util	2	1	Capacity Utilization
21	Emp CPS total	5	1	Civilian Labor Force: Employed, Total
22	U: all	2	1	Unemployment Rate: All Workers, 16 Years & Over
23	U: mean duration	2	1	Unemployment Rate by Duration: Average (Mean) Duration in Weeks
24	U<5 wks	5	1	Unemployment Rate by Duration: Persons Unempl.Less Than 5 Wks
25	U 5-14 wks	5	1	Unemployment Rate by Duration: Persons Unempl. 5 To 14 Wks
26	U 15+ wks	5	1	Unemployment Rate by Duration: Persons Unempl. 15 Wks +
27	U 15-26 wks	5	1	Unemployment Rate by Duration: Persons Unempl. 15 to 26 Wks
28	U 27+ wks	5	1	Unemployment Rate by Duration: Persons Unempl. 27 Wks +
29	UI claims	5	1	Average Weekly Initial Claims, Unemployment Insurance
30	Emp: total	5	1	Employees on Nonfarm Payrolls: Total Private
31	Emp: gds prod	5	1	Employees on Nonfarm Payrolls - Goods-Producing
32	Emp: mining	5	1	Employees on Nonfarm Payrolls - Mining
33	Emp: const	5	1	Employees on Nonfarm Payrolls - Construction
34	Emp: mfg	5	1	Employees on Nonfarm Payrolls - Manufacturing
35	Emp: dble gds	5	1	Employees on Nonfarm Payrolls - Durable Goods
36	Emp: nondble gds	5	1	Employees on Nonfarm Payrolls - Nondurable Goods

37	Emp: services	5	1	Employees on Nonfarm Payrolls - Service-Providing
38	Emp: TTU	5	1	Employees on Nonfarm Payrolls - Trade, Transportation & Utilities
39	Emp: wholesale	5	1	Employees on Nonfarm Payrolls - Wholesale Trade
40	Emp: retail	5	1	Employees on Nonfarm Payrolls - Retail Trade
41	Emp: FIRE	5	1	Employees on Nonfarm Payrolls - Financial Activities
42	Emp: Govt	5	1	Employees on Nonfarm Payrolls - Government
43	Avg hrs: gds prod	1	1	Avg Weekly Hrs of Prod or Nonsup Workers on Private Nonfarm Payrolls - Goods-Producing
44	Overtime: mfg	2	1	Avg Weekly Hrs of Prod or Nonsup Workers on Private Nonfarm Payrolls - Mfg Overtime Hours
45	Avg hrs: mfg	1	1	Average Weekly Hours, Mfg.
46	NAPM empl	1	1	Napm Employment Index (Percent)
47	Starts: nonfarm	4	1	Housing Starts: Total Farm & Nonfarm
48	Starts: NE	4	0	Housing Starts: Northeast
49	Starts: MW	4	0	Housing Starts: Midwest
50	Starts: South	4	0	Housing Starts: South
51	Starts: West	4	0	Housing Starts: West
52	BP: total	4	0	Housing Authorized: Total New Priv Housing Units
53	BP: NE	4	0	Houses Authorized by Build. Permits: Northeast
54	BP: NW	4	0	Houses Authorized by Build. Permits: Midwest
55	BP: South	4	0	Houses Authorized by Build. Permits: South
56	BP: West	4	0	Houses Authorized by Build. Permits: West
57	PMI	1	0	Purchasing Managers' Index
58	NAPM new orders	1	0	Napm New Orders Index (Percent)
59	NAPM vendor del	1	0	Napm Vendor Deliveries Index (Percent)
60	NAPM invent	1	0	Napm Inventories Index (Percent)
61	Orders: cons	5	0	Mfrs' New Orders, Consumer Goods &

	gds			Materials
62	Orders: dble gds	5	0	Mfrs' New Orders, Durable Goods Industries
63	Orders: cap gds	5	0	Mfrs' New Orders, Nondefense Capital Goods
64	Unf orders: dble	5	0	Mfrs' Unfilled Orders, Durable Goods Indus.
65	M&T invent	5	0	Manufacturing & Trade Inventories
66	M&T invent/sales	2	0	Ratio, Mfg. & Trade Inventories To Sales
67	M1	6	0	Money Stock: M1(Curr,Trav.Cks,Dem Dep,Other Ck'able Dep)
68	MZM	6	0	Mzm frb St. Louis
69	M2	6	0	Money Stock:M2(M1+O'nite Rps,Euro$,G/P&B/D Mmmfs&Sav&Sm Time Dep)
70	MB	6	0	Monetary Base, Adj for Reserve Requirement Changes
71	Reserve tot	6	0	Depository Inst Reserves: Total, Adj forReserve Req Chgs
72	Reserves nonbor	6	0	Depository Inst Reserves: Nonborrowed,Adj Res Req Chgs
73	Bus loans	6	0	Commercial & industrial loans at all commercial banks
74	Cons credit	6	0	Consumer Credit Outstanding - Nonrevolving
75	Inst cred/PI	2	0	Ratio, Consumer Installment Credit to Personal Income
76	S&P 500	5	0	S&P Common Stock Price Index: Composite
77	S&P: indust	5	0	S&P Common Stock Price Index: Industrials
78	S&P div yield	2	0	S&P Composite Common Stock: Dividend Yield
79	S&P PE ratio	5	0	S&P Composite Common Stock: Price-Earnings Ratio
80	DJIA	5	0	Common stock prices: Dow Jones Industrial average
81	Fed Funds	2	0	Interest Rate: Federal Funds (Effective)
82	3mo T-bill	2	0	Interest Rate: US Treasury Bills,Sec Mkt,3-Mo.
83	6mo T-bill	2	0	Interest Rate: US Treasury Bills,Sec Mkt,6-Mo.
84	1yr T-bond	2	0	Interest Rate: US Treasury Const Maturities,1-Yr.

85	5yr T-bond	2	0	Interest Rate: US Treasury Const Maturities,5-Yr.
86	10yr T-bond	2	0	Interest Rate: US Treasury Const Maturities,10-Yr.
87	Aaa bond	2	0	Bond Yield: Moody's Aaa Corporate
88	Baa bond	2	0	Bond Yield: Moody's Baa Corporate
89	3mo FF spread	1	0	fygm3-fyff
90	6mo FF spread	1	0	fygm6-fyff
91	1yr FF spread	1	0	fygt1-fyff
92	5yr FF spread	1	0	fygt5-fyff
93	10yr FF spread	1	0	fygt10-fyff
94	Aaa FF spread	1	0	fyaaac-fyff
95	Baa FF spread	1	0	fybaac-fyff
96	Ex rate: avg	5	0	United States: Effective Exchange Rate
97	Ex rate: Switz	5	0	Foreign Exchange Rate: Switzerland (Swiss Franc per US$)
98	Ex rate: Japan	5	0	Foreign Exchange Rate: Japan (Yen per US$)
99	Ex rate: UK	5	0	Foreign Exchange Rate: United Kingdom (Cents per Pound)
100	Ex rate: Canada	5	0	Foreign Exchange Rate: Canada (Canadian$ per US$)
101	PPI: fin gds	6	0	Producer Price Index: Finished Goods
102	PPI: cons gds	6	0	Producer Price Index: Finished Consumer Goods
103	PPI: int matls	6	0	Producer Price Index: Intermed. Mat. Supplies & Components
104	PPI: crud matls	6	0	Producer Price Index: Crude Materials
105	PPI: crude petroleum	6	0	Producer Price Index: Crude Petroleum
106	CPI-U:all	6	1	Cpi-U: All Items
107	CPI-U: apparel	6	1	Cpi-U: Apparel & Upkeep
108	CPI-U: transp	6	1	Cpi-U: Transportation
109	CPI-U:	6	1	Cpi-U: Medical Care

	medical			
110	CPI-U: comm	6	1	Cpi-U: Commodities
111	CPI-U: dbles	6	1	Cpi-U: Durables
112	CPI-U:services	6	1	Cpi-U: Services
113	CPI-U:core	6	1	Cpi-U: All Items Less Food and Energy
114	CPI-U: less shelter	6	1	Cpi-U: All Items Less Shelter
115	CPI-U: less med	6	1	Cpi-U: All Items Less Medical Care
116	PCE defl	6	1	Pce, Impl Pr Defl:Pce
117	PCE defl: dbles	6	1	Pce, Impl Pr Defl:Pce; Durables
118	PEC defl: nondbles	6	1	Pce, Impl Pr Defl:Pce; Nondurables
119	PCE defl: service	6	1	Pce, Impl Pr Defl:Pce; Services
120	AHE: gds	6	1	Avg Hourly Earnings of Prod or Nonsup Workers on Private Nonfarm Payrolls - Goods-Producing
121	AHE: const	6	1	Avg Hourly Earnings of Prod or Nonsup Workers on Private Nonfarm Payrolls - Construction
122	AHE: mfg	6	1	Avg Hourly Earnings of Prod or Nonsup Workers on Private Nonfarm Payrolls - Manufacturing
123	RAHE: gds	5	1	Real avg hrly earnings, prod wrkrs, nonfarm - goods-producing
124	RAHE: const	5	1	Real avg hrly earnings, prod wrkrs, nonfarm - construction
125	RAHE: mfg	5	1	Real avg hrly earnings, prod wrkrs, nonfarm - mfg
126	Consumer expect	2	0	University of Michigan Index of Consumer Expectations