ECON 4101 Econometrics
Reproducibility Term Paper

Pranav Singh

April 28, 2017

Contents

I. Background Information

There are two main competing views of the global market for crude oil:

1. The price of oil is determined by desired stocks. Shifts in the expectations of forward-looking traders (hedgers and speculators) are reflected in changes in the real price of oil and in changes in oil inventories.
2. The real price of oil is determined by shocks to the amount of oil coming out of the ground ("flow supply of oil") and the amount of oil being consumed ("flow demand for oil").

Recently, there has been increasing recognition that both elements of price determination matter. In *The Role of Inventories and Speculative Trading in the Global Market for Crude Oil (2013)*, Kilian and Murphy develop a structural vector autoregressive (VAR) model of the global market for crude oil that nests these two theories together and quantifies the effects of shocks to the speculative demand for oil as well as shocks to flow demand and flow supply of oil. The study stresses the importance of oil inventories data for building a model for price of crude oil. The paper attempts to indirectly identify the effects of shifts in expectations (i.e. without explicit measures of expectations) by including changes in oil inventories in their econometric model. They do this because reliable and substantial data on expectation shifts is not readily available, and because how subjective expectations are formed is too complicated of a nonlinear function to be practical in modeling. The model is built off of four main features (measured monthly from 1973:2 to 2009:8):

1. Percent change in global crude oil production
2. Index of global real activity in deviations from trend
3. Real price of oil
4. Change in above-ground global crude oil inventories

The authors assert that the variables are mutually endogenous. They include two years worth of lags in the model. The paper models four shocks as its responses:

1. Flow supply shock - shock to the flow of crude oil production
2. Flow demand shock - shock to the demand for crude oil driven by the global business
3. Speculative demand shock - shock to the demand for above-ground oil inventories arising from forward-looking behavior
4. Residual shock - captures all structural shocks not otherwise accounted for

II. Methods & Procedures

VAR Methodology

The analysis is based on a dynaic simultaneous equation model in the form of a structural vector autoregression (VAR). Let y_t be a vector of endogenous variables including the percent change in global crude oil production, a measure of global real activity, the real price of crude oil, and the change in global oil inventories above the ground. All data are aggregated monthly. The sample time period is 1973:2 to 2009:8. Seasonal variation is removed by including seasonal dummies in the VAR model.

Data

The dry cargo shipping rate index (as developed by Kilian, 2009a) are used as the measure for fluctuations in global real activity. This index is designed to capture shifts in the global use of industrial commodities. Data on global crude oil production are gathered from the *Monthly Energy Review* of the Energy Information Administration (EIA). Oil production is expressed in period-over-period percent changes in the model. The real price of oil is defined as the US refiners' acquisition cost for imported crude oil, as reported by the EIA, extrapolated from 1974:1 back to 1973:1 as in Barsky and Kilian (2002) and deflated by the US CPI. The data for total US crude oil inventories is also provided by the EIA. These data are scaled by the ratio of OECD petroleum stocks over US petroleum stocks, also obtained from the EIA. That scale factor ranges from about 2.23 to 2.59 in the paper's data sample. The resulting proxy for global crude oil inventories is expressed in unit changes rather than percent changes for two main reasons. Firstly, percent changes do not appear to be covariance stationary, whereas the unit changes do. Secondly, expressing in unit changes was deemed necessary for properly computing the oil demand elasticity.

Structural VAR model for Global Market for Crude Oil

The reduced-form model allows for two years' worth of lags. The corrresponding structural model of the global oil market wass written as:

$$\beta_0 y_t = \sum_{i=1}^{24} \beta_1 y_{t-i} + \epsilon_t$$

where ϵ_t is the vector of orthogonal structural innovations and β_i denotes the coefficient matrices for $i = 0, \ldots, 24$. The seasonal dummies were suppressed for notaional convenience. The vector ϵ_t consists of our four structural shocks:

1. Flow Supply Shock - incorporates supply disruptions associated with exogenous political events in oil-producing countries as well as unexpected politically motivated supply decisions by OPEC members and other flow supply shocks
2. Flow Demand Shock - incorporates to the demand for crude oil and other industrial commodites that is associated with unexpected fluctuations in the global business cycle
3. Speculative Demand Shock - captures shifts in the demand for above-ground oil inventories arising from forward-looking behavior not otherwise captured by the model
4. Residual Shock - captures idiosyncratic oil demand shocks no otherwise acounted for (e.g. weather shocks, changes in technology or prefernces)

Imposed Restrictions for Admissible Models

A flow supply shock is expected to disruput oil production, raise the real price of crude oil and lower global real activity within the same month. Its impact on oil inventories is ambiguous; it may draw down inventories in an effort to smooth consumption, but it may also raise demand for inventories to the extent that it triggers a predictable increase in the real price of oil. A positive flow demand shock raises the real price of oil and stimulates global oil production within the same month. Its impact on inventories is similarly ambiguous. Rather than being associated only with future oil supply conditions or only with future oil demand conditions, speculative demand shocks reflect the expected shortfall of future oil supply relative to future oil demand. A positive speculative demand shock will shift the demand for above-ground oil inventories, causing in equilibrium the level of inventories and the real price of oil to increase on impact.

The aforementioned sign restrictions on the impact responses of oil production, real activity, real price, and inventories to the different shocks are summarized in the following table:

Table I. Sign restrictions on impact responses in VAR model

	Flow supply shock	Flow demand shock	Speculative demand shock
Oil production	−	+	+
Real activity	−	+	−
Real price of oil	+	+	+
Inventories			+

Note: All structural shocks have been normalized to imply an increase in the real price of oil. Missing entries mean that no sign restriction is imposed.

Impact sign restrictions alone are typically too weak to be informative about the effects of oil demand and oil supply shocks. A better set of restrictions relates to bounds on impact price elasticities of oil demand and oil supply. The price elasticity of oil supply depends on the slope of the oil supply curve. For example, a vertical short-run oil supply curve would imply a price elasticity of zero. The impact price elasticity of oil supply is estimated by the ratio of the impact responses of oil production and of the real price of oil to an unexpected increase in flow demand or in speculative demand. Based off a consesnsus in the literature, a fairly stringent bound of 0.025 on the impact price elasticity of oil supply was used for this paper. The *oil demand elasticity in use* is the sum of the flow of oil production and the depletion of oil inventories triggered by an oil supply shock. The paper makes the reasonable assumption that the impact price elasticity of oil demand is lower in magnitude than the corresponding long-run price elasticity of oil demand. Based off results from relevant literature, this paper uses a bound of -0.8 and 0 for the impact price elasticity of demand in use.

The final set of restrictions imposed is that the response of the real price of oil to a negative flow supply shock must be positive for at least 12 months, starting in the impact period. This restriction is necessary to rule out structural models in which unanticipated flow supply disruptions cause a decline in the real price of oil below its starting level, which would be at odds with conventional views of the effects of unanticipated oil supply disruptions. This restriction necessitates that global real activity responds negatively to oil supply shocks, since that's the only way for the oil market to experience higher prices and lower quantities in practice. This implies a joint set of sign restrictions such that the responses of oil production and global real activity to an unanticipated flow supply disruption are negative for the first 12 months, while the response of the real price of oil is positive.

Simulation and Identification Procedure for Admissible Models

The procedure consists of the following steps: 1. Draw an $N \times N$ matrix K of $NID(0,1)$ random variables. Derive the QR decomposition of K such that $K = Q \cdot R$ and $QQ' = I_N$. 2. Let $D = Q'$. Compute impulse responses using the orthogonalization $\tilde{B} = BD$, where Σ_e is the variance-covariance matrix for the vector e_t of white noise reduced-form innovations and B satisfies $\Sigma_e = BB'$. If all implied impulse response functions satisfy the identifying restrictions, retain D. Otherwise discard D. 3. Repeat the first two steps a large number of times (5 million), recording each D that satisfies the restrictions and record the corresponding impulse response functions.

The resulting set $\tilde{\mathbb{B}}$ corresponds to the set of all admissible structural VAR models. The estimation uncertainty underlying these structural impulse response estimates is assesed by Bayesian methods by specifying a diffuse Gaussian-inverse Wishart prior distribution for the reduced-form parameters and a Haar distribution for the rotation matrix. The posterior distribution of the structural impulse responses is obtained by applying our identification procedure to each draw of $A(L)$ and Σ_e from their posterior distribution. The results in this paper are solely focused on the single model among the admissible structural models that yields an impact price elasticity of demand in use closest to the posterior median of this elasticity among the candidate models that satisfy all identifying restrictions.

Our replication study obtained the set $\tilde{\mathbb{B}}$ of all admissible structural VAR models by running the provided MATLAB code. All subsequent analysis, such as obtaining obtaining relevant impact response functions for Figure 1 and computing cumulative effects to the structural shocks for Figure 2 were performed in R on a Ubuntu 16.04 64-bit x84 linux platform.

III. Replication Results

Part I: Replication of Structural Impulse Responses of Shocks

```
bayesPosterior <- readMat('../BayesPosterior.mat'); IRMposs <- bayesPosterior$IRMposs
IRFelas <- readMat('../IRFelas.mat'); IRFelas <- IRFelas$IRFelas;
findex <- readMat('../findex.mat'); findex <- findex$findex;

xmax = 17
mindist <- 0.0061

IRF <- IRFelas[,,findex]
time <- c(0:xmax);
CI <- apply(IRMposs, c(1,2), quantile, probs = c(.16,.84))
CI1458912=apply(apply(IRMposs, c(1,3), cumsum), c(1,2), quantile, probs = c(.16,.84))
for (i in c(1, 4, 5, 8, 9, 12)) {
  CI[, i, ] <- CI1458912[, , i]
}

CI5 = apply(IRMposs, c(1,2), quantile, probs=c(.025, .975))
CI5_1458912=apply(apply(IRMposs, c(1,3), cumsum), c(1,2), quantile, probs = c(.025, .975));
for (i in c(1, 4, 5, 8, 9, 12)) {
  CI5[, i, ] <- CI5_1458912[, , i]
}
```

```
fn <- function(title, Months, ylabel, y, yl, yu, ylim) {
  df <- data.frame(Months=Months, ylabel = y, yl = yl, yu = yu)
  ggplot(df, aes(x=Months)) +
    geom_line(aes(y=ylabel), color='red') +
    geom_line(aes(y=yl), color='blue', linetype='dashed') +
    geom_line(aes(y=yu), color='blue', linetype='dashed') +
    geom_hline(aes(yintercept=0)) +
    scale_y_continuous(limits = ylim) +
    ggtitle(title) +
    labs(y = ylabel) +
    theme(axis.text=element_text(size=6),
          axis.title=element_text(size=8),
          plot.title=element_text(size=8))
}
```

```r
g1 <- fn('Flow Supply Shock', time
         , 'Oil Production', -cumsum(IRF[1,])
         , -CI[1,1,], -CI[2,1,], c(-2,2))
g5 <- fn('Flow Demand Shock', time
         , 'Oil Production', cumsum(IRF[5,])
         , CI[1,5,], CI[2,5,], c(-2,2))
g9 <- fn('Speculative Demand Shock', time
         , 'Oil Production', cumsum(IRF[9,])
         , CI[1,9,], CI[2,9,], c(-2,2))

g2 <- fn('Flow Supply Shock', time
         , 'Real Activity', -IRF[2,]
         , -CI[1,2,], -CI[2,2,], c(-10,10))
g6 <- fn('Flow Demand Shock', time
         , 'Real Activity', IRF[6,]
         , CI[1,6,], CI[2,6,], c(-10,10))
g10 <- fn('Speculative Demand Shock', time
          , 'Real Activity', IRF[10,]
          , CI[1,10,], CI[2,10,], c(-10,10))

g3 <- fn('Flow Supply Shock', time
         , 'Real Price of Oil', -IRF[3,]
         , -CI[1,3,], -CI[2,3,], c(-10,10))
g7 <- fn('Flow Demand Shock', time
         , 'Real Price of Oil', IRF[7,]
         , CI[1,7,], CI[2,7,], c(-10,10))
g11 <- fn('Speculative Demand Shock', time
          , 'Real Price of Oil', IRF[11,]
          , CI[1,11,], CI[2,11,], c(-10,10))

g4 <- fn('Flow Supply Shock', time
         , 'Inventories', -cumsum(IRF[4,])
         , -CI[1,4,], -CI[2,4,], c(-20,20))
g8 <- fn('Flow Demand Shock', time
         , 'Inventories', cumsum(IRF[8,])
         , CI[1,8,], CI[2,8,], c(-20,20))
g12 <- fn('Speculative Demand Shock', time
          , 'Inventories', cumsum(IRF[12,])
          , CI[1,12,], CI[2,12,], c(-20,20))

globs <- list(g1,g2,g3,g4,g5,g6,g7,g8,g9,g10,g11,g12)
```

Original Figure 1: Structural Impulse Responses

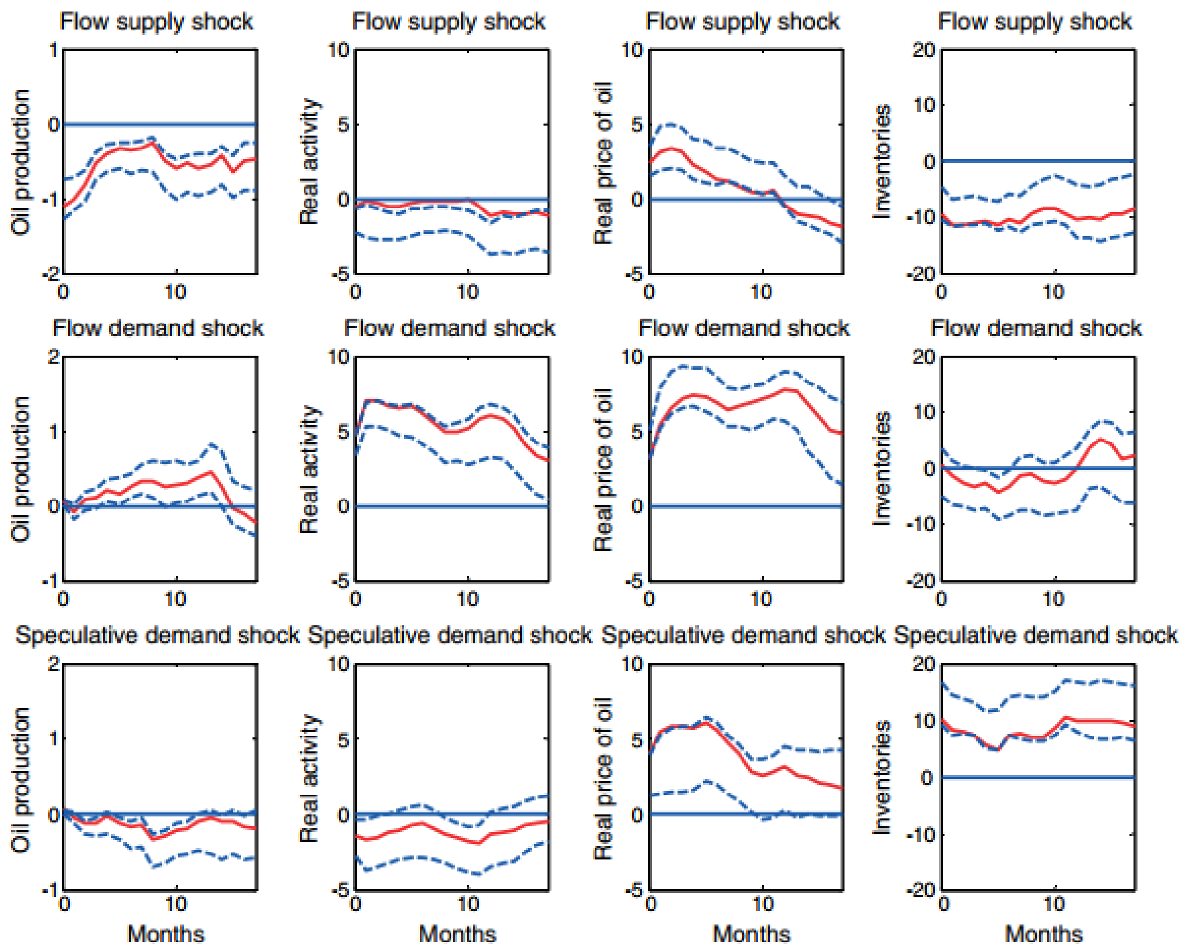
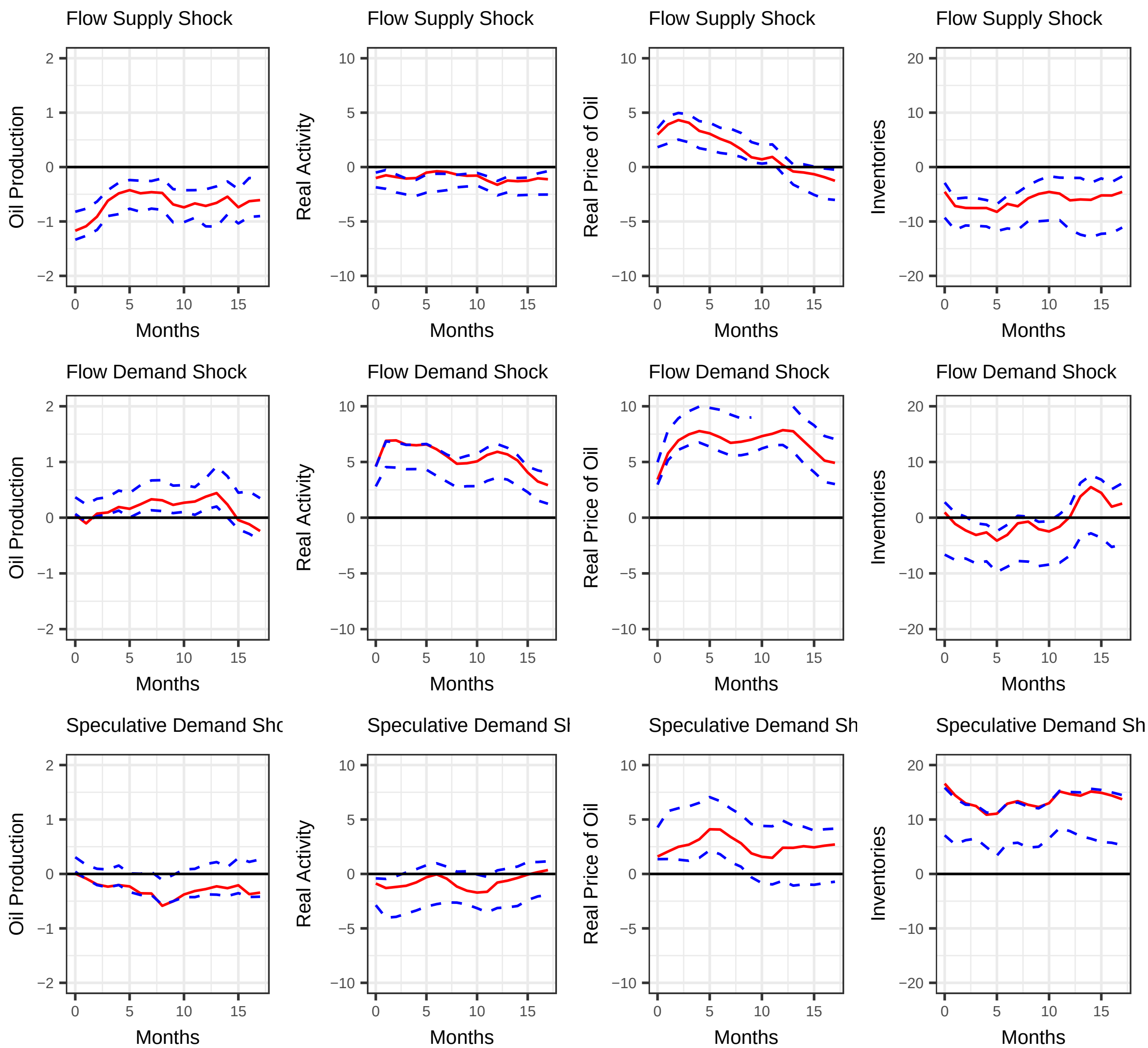


Figure 1. Structural impulse responses: 1973:2–2009:8. Solid lines indicate the impulse response estimates for the model with an impact price elasticity of oil demand in use closest to the posterior median of that elasticity among the admissible structural models obtained conditional on the least-squares estimate of the reduced-from VAR model. Dashed lines indicate the corresponding pointwise 68% posterior error bands. 'Oil production' refers to the cumulative percent change in oil production and inventories to cumulative changes in inventories

Replicated Figure 1: Structural Impulse Responses

```r
grid.arrange(grobs=globs, layout_matrix = matrix(1:12, byrow = T,nrow=3))
```



Comparison of Replication Results for Figure 1

Figure 1 plots the responses of each variable in this benchmark model to the three oil supply and oil demand shocks, along with the corresponding pointwise 68% poseterior error bands obtained by drawing from the reduced-form posterior distribution. All shocks have been normalized such that they imply an increase in the real price of oil.

Response of Oil Production

Our replicated responses of oil production to the three shocks closely mirror those reported in the paper, both in magnitude of effect and in trend over time. As such, our replicated results support the paper's findings on oil production responses: a negative flow spply shock is associated with a persistent decrease; a positive flow demand shock causes a temporary increase; and a positive speculative demand shock causes a negative, but small, effect.

Response of Real Activity

Our replicated responses of real activity levels to the three shocks closely mirror those reported in the paper, both in magnitude of effect and in trend over time. As such, our replicated results support the paper's findings on real activity responses: a negative flow supply shock causes a small but persistent reduction; a positive flow demand shock causes a sizeable persistent increase; and a positive speculative demand shock casues a small negative effect.

Response of Real Price of Oil

Our replicated responses of the real price of oil to flow supply and demand shocks closely mirror those reported in the paper, both in magnitude of effect and in trend over time. However, our replicated response of the real price of oil to speculative demand differs significantly from that which was reported in the paper. In particular, our replicated findings support the paper's claim that a negative flow supply shock temporarily increases the real price of oil, but returns to its original price within a year and continues to fall as real activity continues to decrease. Our replicated findings also support the paper's claim that a positive flow demand causes a hump-shaped rise in the real price of oil which peaks about one year in. However, whereas the paper reported that a positive speculative demand shock caused an immediate significant jump in oil prices, our replicated findings didn't reflect this. Furthermore, the paper reported that this immediate jump was an overshooting of the price and was followed by a steady gradual decline in the real price of oil. While our replicated results also indicate that a positive speculative demand shock causes an overshooting of the real price of oil, our results suggest it happens much later than what was reported in the paper, and our results don't support the subsequent steady gradual decline effect.

Response of Inventories

While our replicated results for the response of inventory to flow supply and speculative demand shocks differ insignificantly in magnitude, the response over the next 18 months following the shock follows the same general trend as that reported in the paper. Our replicated response of inventory to flow demand shocks closely mirrors the paper's reported one. Moreover, the paper's main takeaway for the inventory responses remains visible in our replicated results: a flow supply shock yields a negative response; a flow demand shock causes almost no change on impact, followed by a temporary drawdown of oil inventories and a gradual recovery within 1 year; and a speculative demand shock causes a persistent increase in inventories.

Part II: Replication of Historical Decompositions of Cumulative Effects of Shocks

```r
bayesPosterior <- readMat('../BayesPosterior.mat'); IRMposs <- bayesPosterior$IRMposs
IRFelas <- readMat('../IRFelas.mat'); IRFelas <- IRFelas$IRFelas;
findex <- readMat('../findex.mat'); findex <- findex$findex;
U <- readMat('../U.mat'); U <- U$U
BETAnc <- readMat('../BETAnc.mat'); BETAnc <- BETAnc$BETAnc

IdentMat <- matrix(IRFelas[,1,findex], nrow=4)
Uhat <- U
p=24;
t=439; # length(kmData)
K <- nrow(IdentMat)
q <- ncol(IdentMat)

# Compute structural multipliers
A = rbind(BETAnc, cbind(diag(K*(p-1)), diag(x=0, K*(p-1), K)))
J = cbind(diag(K), diag(x=0, K, K*(p-1)))
IRF = matrix(J %*% (A %^% 0) %*% t(J) %*% IdentMat, nrow = K^2, ncol = 1)
for (i in 1:(t-p-1)) {
  IRF = cbind(IRF, matrix(J %*% (A %^% i) %*% t(J) %*% IdentMat, nrow = K^2, ncol = 1))
}

# Compute structural shocks Ehat from reduced form shocks Uhat
Ehat = MASS::ginv(IdentMat) %*% Uhat[1:q,];

# Cross-multiply the weights for the effect of a given shock on the real
# oil price (given by the relevant row of IRF) with the structural shock
# in question
yhat1 = diag(x=0,t-p,1);
yhat2 = diag(x=0,t-p,1);
yhat3 = diag(x=0,t-p,1);
yhat4 = diag(x=0,t-p,1);
for (i in 1:(t-p)) {
  yhat1[i,] = IRF[3, 1:i] %*% Ehat[1, i:1]
  yhat2[i,] = IRF[7, 1:i] %*% Ehat[2, i:1]
  yhat3[i,] = IRF[11, 1:i] %*% Ehat[3, i:1]
  yhat4[i,] = IRF[15, 1:i] %*% Ehat[4, i:1]
}

time = seq(from = (1973+2/12+1/12*p), to = 2009+8/12, by = 1/12); # starts at 1975.2
cumshock = yhat1 + yhat2 + yhat3 + yhat4;
```

```r
df <- data.frame(Years=time, CumEffect=yhat1)
g <- ggplot(mapping=aes(x=time)) + geom_vline(aes(xintercept=1978+9/12)) +
  geom_vline(aes(xintercept=1980+9/12)) +
  geom_vline(aes(xintercept=1985+12/12)) +
  geom_vline(aes(xintercept=1990+7/12)) +
  geom_vline(aes(xintercept=1997+7/12)) +
  geom_vline(aes(xintercept=2002+11/12)) +
  geom_vline(aes(xintercept=1980), linetype = 'dashed') +
  geom_vline(aes(xintercept=1985), linetype = 'dashed') +
  geom_vline(aes(xintercept=1990), linetype = 'dashed') +
```

```
  geom_vline(aes(xintercept=1995), linetype = 'dashed') +
  geom_vline(aes(xintercept=2000), linetype = 'dashed') +
  geom_vline(aes(xintercept=2005), linetype = 'dashed') +
  geom_hline(aes(yintercept=0), linetype = 'dashed') +
  geom_hline(aes(yintercept=-50), linetype = 'dashed') +
  geom_hline(aes(yintercept=50), linetype = 'dashed') +
  scale_y_continuous(limits = c(-100, 100)) +
  scale_x_continuous(breaks = seq(1980, 2005, by = 5)) +
  ylab('')
g1 <- g + geom_line(aes(y=yhat1), color='blue') +
  ggtitle('Cumulative Effect of Flow Supply Shock on Real Price of Crude Oil')
g2 <- g + geom_line(aes(y=yhat2), color='blue') +
  ggtitle('Cumulative Effect of Flow Demand Shock on Real Price of Crude Oil')
g3 <- g + geom_line(aes(y=yhat3), color='blue') +
  ggtitle('Cumulative Effect of Speculative Demand Shock on Real Price of Crude Oil')
```

Original Figure 2: Historical Decompositions

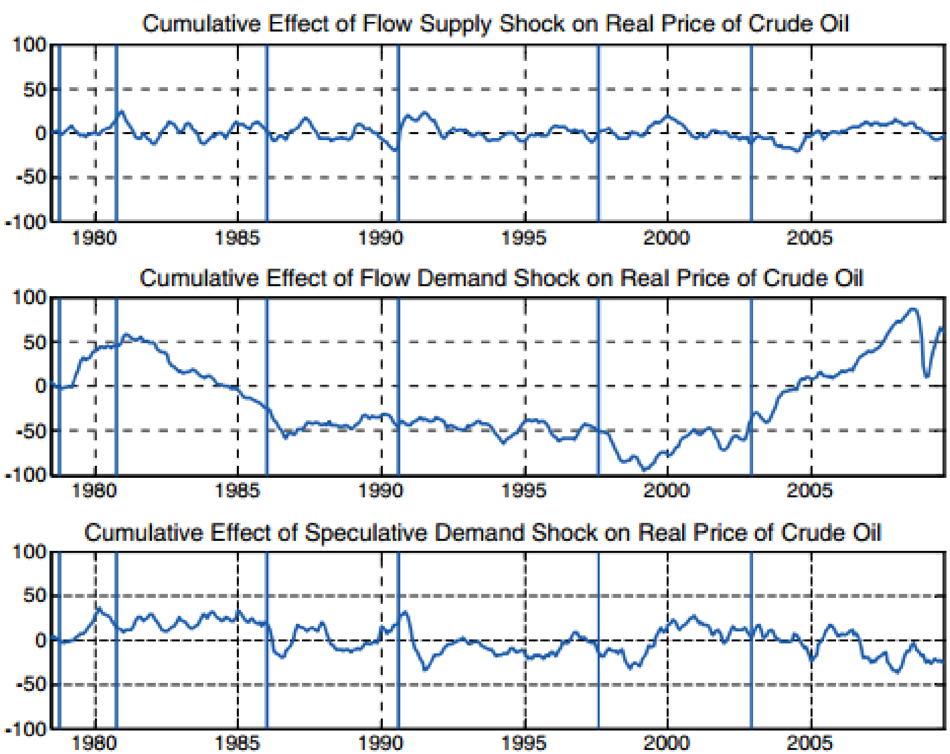
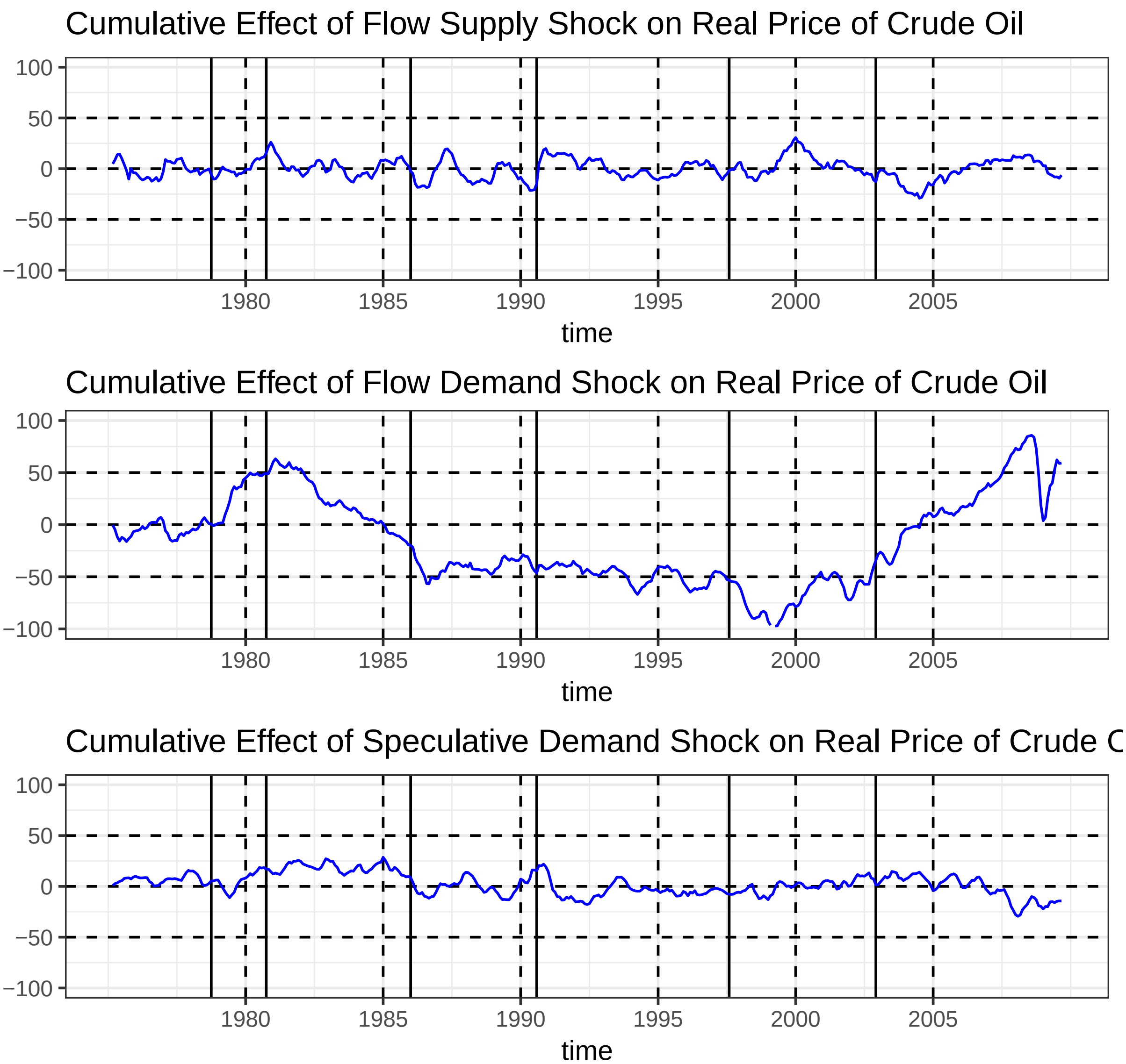


Figure 2. Historical decompositions for 1978:6–2009:8. Based on benchmark estimate as in Figure 1. The vertical bars indicate major exogenous events in oil markets, notably the outbreak of the Iranian Revolution in 1978:9 and of the Iran–Iraq War in 1980:9, the collapse of OPEC in 1985:12, the outbreak of the Persian Gulf War in 1990:8, the Asian Financial Crisis of 1997:7, and the Venezuelan crisis in 2002:11, which was followed by the Iraq War in early 2003. In constructing the historical decomposition we discard the first five years of data in an effort to remove the transition dynamics

Replicated Figure 2: Historical Decompositions

```r
grid.arrange(grobs=list(g1,g2,g3), nrow=3)
```



Comparision of Replication Results for Figure 2

Our replicated results for Figure 2 match very closely with the paper's results. Thus, our replicated results support the paper's analyis about what caused the oil price shock of 2003-2008. The standard interpretation for this sharp increase is (a) that there was an exogenous influx of financial investors into the oil futures market, (b) that this influx drove up oil futures prices, and (c) that the increase in oil futures prices was viewed by spot market praticipants as a signal of an expected increase in the price of oil, shifting inventory demand and hence causng the real spot price to increase. This explanation would necessitate that speculative demand shocks inthe structural model should explain the bulk of the price surge between 2003 and 2008. However, as the bottom graph of Figure 3 shows, there is no evidence to support this view; the cumulative effect of speculative demand shocks remained close to zero during that time period. A competing explanation posits that the price surge is a result of OPEC holding back its production after 2001 in anticipation of even higher oil prices in the future. In our model, such a move would be observationally equivalent to a negative flow supply shock. However, Figure 2 provides no indication that negative flow supply shocks were an important determinant of the real price of oil between 2003 and 2008. The small spike in flow supply shocks' cumulative effect on the real price of oil is dwarfed in comparison to flow demand shocks' cumulative effects during that period. Thus, our replicated findings support the paper's alternative explanation to price surge: the price surge between 2003 and 2008 was mainly caused by shifts in the flow demand for crude oil associated with the global business cycle. This finding is important because it implies that further regulation of oil markets would have done little to stem the increase in the price of oil. That is, there is no basis for the premise that such regulation is required to lower the real price of oil or that it would be helpful. Furthermore, the paper's structural model also implies that even dramatic increases in US oil production would not lower the real price of oil substantially at the global level.