Endogenous monetary policy and effects of oil shocks revisited using structural FAVAR approach *

Konstantin Styrin

This draft: January 31, 2009
This draft: April 20, 2009

Abstract

It has been observed that oil price increases are followed by recessions in the US and other countries. One leading explanation suggests that it is endogenous monetary policy response aiming to eliminate inflationary consequences of oil shocks that triggers recessions. This study re-examines the role of systematic monetary policy in amplification of shocks in world oil supply using structural Factor Augmented VAR (FAVAR). I estimate a structural FAVAR on a panel of 132 US macroeconomic time series assuming that all of them are driven by a small number of exogenous structural shocks. I identify the oil supply, oil demand and US monetary policy shocks and run a counterfactual policy experiment shutting off endogenous monetary policy response in the US to oil shocks. Unlike most of the literature, my identification procedure distinguishes between oil demand and supply shocks. Contrary to earlier studies based on conventional small-size VAR, I find that the systematic monetary policy response has been contractionary with respect to positive oil demand shocks and accommodating with respect to adverse supply shocks. This implies that holding interest rates fixed in response to OPEC I and II shocks would have produced even deeper recessions in the 1970s. My results give support to “good luck” interpretation of the weakened effect of oil shocks on the economy as documented in the literature.

JEL Classification: E4, E5, Q4

Key words: Oil shocks; Systematic monetary policy; Factor-augmented vector autoregressions; Structural shocks.

1 Introduction

It has been observed that, in the post-World War II period, oil price increases were followed by recessions. Hamilton (1983) finds that a 10% increase in the price of petroleum above its maximum level over the past 12 months is associated with a subsequent decrease in the US GDP with a peak of minus 2% happening in four quarters after the shock. What is puzzling is the size of the negative effect given the fact that oil accounts for 3 – 4% of total expenditures. Indeed, the most one could expect to observe is $10\% \times 4\% = 0.4\%$ decrease in output. One way to reconcile such a low share of oil in expenditures with its non-negligible effect on

---

*I am grateful to Kenneth Rogoff and James Stock for invaluable advice and encouragement. I thank seminar participants at Harvard for helpful comments. All errors are mine.
the economy would be to identify some sort of magnification mechanism. The literature took this route and several solutions to the “oil puzzle” were suggested. Rogoff (2006) and Hamilton (2005) survey this literature.

This paper addresses empirical validity of only one of the competing explanations that emphasizes the role of endogenous monetary policy (Bohi, 1989). Observing an increase in the price of petroleum, the Federal Reserve may be willing to tighten monetary policy by raising interest rates in an attempt to reduce future inflation that such a shock is very likely to produce. As a result of increased interest rates, the economy will enter a recession. To a certain extent, the after-shock recession observed in the data might have been thus produced by the endogenous monetary policy response rather than the oil shock per se.

Bernanke, Gertler, and Watson (1997) try to figure out empirically to what extent the endogenous monetary policy is responsible for the recessions in the aftermath of oil price increases. They estimate a seven-variable VAR where identification is obtained through ordering variables as real GDP, GDP deflator, commodity price index, oil, federal funds rate, 3-month T-bill rate, and 10-year T-bond rate. Using the estimated structural VAR, they run several counterfactuals and conclude that endogenous monetary policy has been the main cause of recessions in the aftermath of oil price shocks. If one shuts off the monetary policy response to oil price shocks, the recessionary consequences of an oil price increase almost disappear.

Hamilton and Herrera (2004) offer two important criticisms to Bernanke, Gertler, and Watson (1997): first, the VAR order is chosen incorrectly and, second, the counterfactual experiments involving a permanent change in the monetary policy rule are subject to Lucas critique. In their response, Bernanke, Gertler, and Watson (2004) re-estimate their VAR on quarterly as opposed to monthly data that they used in their original study, which allows them to avoid including too many lags and thus reduce sample uncertainty. Furthermore, they assume that changes in the systematic part of monetary policy are not permanent, although long-lasting, which reduces the severity of Lucas critique. They conclude that, although the “residual” effect of oil shock on GDP becomes non-negligible under such circumstances, the monetary policy response accounts for about half of the size of a decrease in GDP.

This paper revisits the explanation of the endogenous monetary policy explanation of the “oil price puzzle”. I exploit a large monthly dataset of 132 macroeconomic time series covering 44 years, 1960:1-2003:12, augmented by nominal oil price. The core of my empirical methodology is Factor-Augmented Vector Autoregression (FAVAR) that assumes that all macro variables are driven by a relatively small number of structural shocks (Bernanke, Boivin, and Eliasz, 2005; Stock and Watson, 2005).

Historically, the FAVAR as well as their precursor Dynamic Factor Models (DFM) was introduced in the macro forecasting literature as an attempt to expand the information set used by the forecaster beyond
the limits of univariate ARIMA type model or a VAR. An appealing feature of FAVAR from the “structural” perspective is that it places the econometrician on an equal ground with economic agents who normally use much more information in their decision making than that contained in a small number of series included into a conventional low-scale VAR. This significantly reduces the possibility of an omitted variable bias in estimated structural impulse responses.

I start by estimating a reduced-form FAVAR where factors are estimated as principal components of the (pre-filtered) macro time series. Then, in order to separate between the direct effect of an adverse price shock and the effect caused by endogenous monetary policy response, I need to identify two structural shocks, the monetary policy (MP) shock and the oil shock.

When identifying the MP shock, I follow Bernanke, Boivin, and Eliasz (2005) and use the Stock and Watson’s (2005) version of their block-diagonal Wold-like identification scheme. It is assumed that all variables can be split into slow moving (quantities and consumer prices), federal funds rate (FF) and fast moving (producer prices, expectations, and financial market indicators). In their turn, all factors can be split into slow moving, MP, and fast moving. Slow variables can respond to shocks in MP or fast moving factors only with a lag but not contemporaneously. The FF rate, which serves as an indicator of the monetary policy stance, can respond to slow moving factors contemporaneously but to fast moving factors only with a lag. In other words, the Fed is able to respond to innovations in the slow moving variables (real sector) within a month but to those in the fast moving variables (financial markets) with a lag. These assumptions suffice to identify the exogenous MP shock.

Unlike most of the empirical macro literature on oil, the identification of the oil shock that I use is not based on Wold ordering. Existing VAR studies involving oil differ in terms of the Wold order they assign to oil. For example, Bernanke, Gertler, and Watson (1997) place oil after the quantity variables and goods prices and right before the FF rate. The identifying assumption is that quantity variables and goods prices do not respond contemporaneously to oil shocks while the Fed and financial markets can react within a month. Blanchard and Galí (2007), instead, place the oil price first thus assuming that the price of oil price does not respond to innovations in all other variables they use (three measures of inflation, GDP, and employment) within a period (month, quarter?). These two Wold orderings seem to neglect the fact that the oil price like all other commodity prices contains an important forward-looking component. These two schemes might be appropriate for the time period of posted prices by OPEC, i.e. before 1983. Afterwards, i.e. for a considerable fraction of samples used in the two above-mentioned studies, oil prices were determined on a (relatively) free market. It is worth noting that Bernanke, Boivin, and Eliasz (2005), although neither aim to identify the oil shock nor include the price of oil into their FAVAR, place a variable called “Sensitive
materials price index”, which contains oil price as one of its components, into the group of fast moving variables.

In this study, I use a conventional identification procedure for oil shocks and develop an alternative one. The first scheme based on the assumption that, within one month, oil price responds only to oil supply shocks and it reacts to other developments in the economy only with a lag. The new scheme that I offer distinguishes between oil supply and oil demand shocks.

The contribution of the systematic monetary policy to a post-oil-shock evolution of the economy can be estimated in the way similar to Bernanke, Gertler, and Watson (1997). First, we estimate our preferred specification and extract the time paths of structural shocks. Second, we simulate the estimated structural FAVAR assuming that the systematic monetary policy is shut off, i.e. the FF rate does not respond to changes in the price of oil. Finally, we compare the difference between the actual and simulated time paths of macroeconomic times series. I have to decide how exactly I will shut off the systematic MP. Should I keep the FF rate fixed in my simulations or should I set the coefficient on the identified oil structural shock set equal zero? The former assumption seems very restrictive while the latter too soft (check with BGW for details!) while both counterfactual experiments are subject to Lucas critique. Indeed, learning that the Fed has switched from one MP rule to another will affect expectations of economic agents and, as a result, the coefficients of FAVAR as well as impulse responses will change. Bernanke, Gertler, and Watson (1997) address this issue by allowing the term premium of government bonds change in response to the regime switch. Bernanke, Gertler, and Watson (2004) assume that the regime switch is not permanent although very persistent and therefore the bias caused by the changed expectations should not be too large.

My findings based on the new identification scheme for oil shocks are opposite to the consensus view. Adverse oil supply shocks did not trigger monetary tightening. Instead, the monetary policy response was accommodating.

I identify separately oil demand and supply shocks. Impulse responses to the two identified shocks make sense. Specifically, oil supply shocks induce recession, dampen stock market, and trigger a decrease in interest rates. Oil demand shocks, on the contrary, are followed by a boom both in real activity and stock market and higher interest rates.

My identification procedure is based on the Stock and Watson’s (2005) version of Bernanke, Boivin, and Eliasz’s (2005) procedure for the monetary shock. I add only one new element on the top of it. The idea is very simple. Oil is a storable commodity and, therefore, its price responds to expectations of higher demand for oil in the future. If the public learn good news about future productivity growth, then they expect higher demand for oil in the future, and this will drive up the price for oil. Assume that the majority of oil demand
shocks happen due to changes in expectations rather than in current demand (e.g., cold winter). Then, in the jargon of BBE’s procedure, shifts in oil demand should be driven mostly by fast shocks, i.e. by a subgroup of shocks that affect fast moving variables (stock market, expectation surveys, commodity prices) on impact but have a delayed effect on slow moving series (output, employment). Technically, I identify the oil supply (demand) shock as the OLS projection of FAVAR innovation for oil price on the subspace of slow (MP and fast) shocks, which is delivered by the BBE’s scheme.

As I mentioned, the response of the systematic MP is different for the two kinds of oil shocks. It is accommodating for oil supply shocks and anti-inflationary for oil demand shocks. In that respect, the results look quite opposite to Bernanke, Gertler, and Watson (1997), which are based on block-recursive identification and do not separate oil demand from oil supply shocks. My findings imply that if the Fed had not responded to oil (supply) shocks (i.e. if had not accommodate them) by cutting interest rates then the recessions in the US in the aftermath of those shocks would have been even deeper. This is consistent with the view that, in the 1970s and early 1980s, the MP was more accommodating than in the subsequent period (i.e. during Volcker era and later).

My results may explain why Blanchard and Gali (2007) found that the response of the economy to oil shocks had weakened over time. Like BGW, their (block-recursive) identification scheme does not distinguish between oil demand and supply shocks. One view is that, in the 1970s and early 1980s, the oil price was driven mostly by supply shocks whereas, in the later period, by demand shocks. In that case, the decrease in the effect of oil shocks on the economy as documented by BG is not necessarily due to an improved ability of the economy to adjust to oil shocks as they claim (lower wage rigidity, more effective MP, etc.). Instead, a “good luck” explanation – the fact that no major oil supply shocks compared with the 1970s have happen recently – might sound more appealing.

The rest of the paper is organized as follows. Section 2 discusses FAVAR methodology, explains how the structural shocks are identified and describes data. Section 3 presents empirical results. Section 4 concludes.

2 Factor-augmented vector autoregression

The key issue in this study is valid identification of two structural shocks of interest, MP and oil supply. Most of the oil literature are studies based on conventional VARs and exploit a version of block-recursive identification scheme.

One question that naturally arises is how to identify and estimate the MP shock. In principle, it can be done using a structural VAR but the models of this family are notorious for their poor out-of-sample forecasting performance. Furthermore, SVARs are prone to omitted variable bias because of their inability
to operate with a sufficiently large number of variables. A leading example is the “price puzzle”, a counter-intuitive empirical finding that a contractionary MP shock raises price level in the short run. A common interpretation of this empirical pattern is that some important variables from the Fed’s information set are omitted and this results in misidentified MP shock. It is conceivable that the Fed tightens MP in an attempt to reduce a spike in future inflation that they anticipate based on available information. The curse of dimensionality prevents an econometrician from including into VAR all variables from the Fed’s information set. Although adding a commodity price index as a forward-looking proxy of inflation expectations removes the price puzzle, there is still no guarantee that the MP shock is identified correctly since other sources of OVB may be present.

Oil shocks are identified in the literature by assigning oil price a particular Wold order in the context of block recursive identification. For example, oil price is ordered prior to all other included variables in Rotemberg and Woodford (1996), Blanchard and Galí (2007) while it is placed between slow moving series and the federal funds rate in Bernanke, Gertler, and Watson (1997). Kilian (2007) emphasizes that one should distinguish between different kinds of oil shocks, supply, demand, and oil-sector-specific. He identifies these three structural shocks block-recursively within a three-variable VAR. There is no guarantee that Kilian’s approach is not subject to an OVB.

Factor Augmented VAR (FAVAR) provides a more profound and systematic treatment of OVB and therefore yields potentially more reliable identification of structural shocks compared with VARs. They are able to accommodate a large number of variables, potentially, all variables from the Fed’s information set. The key assumption that makes estimation feasible is that all macro time series are driven by a small number of common factors/shocks. All interdependencies among the variables work only through the dependence on common factors. The factors are estimated as principal components of data. Effectively, FAVAR turns curse of dimensionality into a blessing: more data help more precisely estimate the space of common shocks.

Bernanke, Boivin, and Eliasz (2005) estimate FAVAR using a panel of 120 US macro time series and identify structural MP shock. They find no evidence of the price puzzle and conclude that it is an artifact of misspecified VAR.

The rest of this section describes the FAVAR methodology and identification of MP shock in FAVAR. This discussion follows very closely Stock and Watson (2005).

2.1 Factor Augmented Vector Autoregression

Conventional reduced-form VAR can be written as

\[ X_t = C(L)X_{t-1} + e_t \]
where $X_t$ is $n \times 1$ vector of endogenous variables and $C(L)$ is a matrix of lag polynomials. If the VAR order is $p$ then the number of parameters to be estimated is $n((2p+1)n-1)/2$ in $C(L)$ and $n(n-1)/2$ in the covariance matrix of $e_t$. VAR is subject to the “curse of dimensionality” since the number of estimated parameters grows as $n^2$. Adding more variables makes estimates less precise. This is why, in practice, VARs rarely include more than 7-12 variables, which represent a very small subset of information available to economic agents and policy makers.

To make the problem manageable, assume that all included variables depend on a relatively small number of common factors:

$$X_{it} = \tilde{\lambda}_i(L)'f_t + u_{it} \quad (1)$$

where $f_t$ is $q \times 1$ vector of dynamic factors, $\tilde{\lambda}_i$ is $q \times 1$ vector of lag polynomials (dynamic factor loadings) of order $k$, and $u_{it}$ is idiosyncratic term (measurement error). Factors $f_t$ are dynamic in the sense that $X_{it}$ depends on contemporaneous and lagged values of $f_t$. Representation 1 assumes that all cross-dependencies among $X_{it}$’s work only through common factors $f_t$. In general, $u_{it}$ is serially correlated:

$$u_{it} = \delta_i(L)u_{it-1} + v_{it} \quad (2)$$

Combining equations (1) and (2) gives

$$X_{it} = \tilde{\lambda}_i(L)'(I - \delta_i(L)L)f_t + \delta_i(L)X_{it-1} + v_{it}$$

$$= \lambda_i(L)'f_t + \delta_i(L)X_{it-1} + v_{it} \quad (3)$$

The evolution of factors is governed by a VAR process:

$$f_t = \Gamma(L)f_{t-1} + \eta_t \quad (4)$$

where $\eta_t$ is $q \times 1$ vector of exogenous shocks that hit the economy with

$$\mathbb{E}(\eta_t\eta_t') = I_q$$

Shocks and the error term in (3) are assumed uncorrelated

$$\mathbb{E}(\eta_tv_{it}) = 0$$

as well as the error terms for any two different $X$’s

$$\mathbb{E}(v_{it}v_{js}) = 0$$

for all $t, s, i, j, i \neq j$. 

7
Introduce the vector of static factors $F_t$ in such a way that $X_{it}$ would depend only on contemporaneous values of $F_t$. Then equation (3) will take the form

$$X_{it} = \Lambda_i'F_t + \delta_i(L)X_{it-1} + v_{it}$$  \hspace{1cm} (5)$$

or, in vector notation,

$$X_t = \Lambda F_t + D(L)X_{t-1} + v_t$$  \hspace{1cm} (6)$$

where $\Lambda$ is a matrix of static factor loadings

$$\Lambda = \begin{pmatrix} 
\Lambda_1' \\
\Lambda_2' \\
\vdots \\
\Lambda_n' 
\end{pmatrix}$$  \hspace{1cm} (7)$$

and

$$D(L) = \begin{pmatrix} 
\delta_1(L) & 0 & \ldots & 0 \\
0 & \delta_2(L) & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \delta_n(L) 
\end{pmatrix}$$  \hspace{1cm} (8)$$

Comparing dynamic and static factor representations (3) and (6), it is clear that $F_t$ will be an $r \times 1$ vector that contains all elements of $f_t$ and certain elements of lagged $f_t$ with $q \leq r \leq kq$. Equation (4) implies the following law of motion for static factors $F_t$:

$$F_t = \Phi(L)F_{t-1} + G\eta_t$$  \hspace{1cm} (9)$$

Substitute (9) in (6) to obtain

$$X_t = \Lambda \Phi(L)F_{t-1} + D(L)X_{t-1} + \Lambda G\eta_t + v_t$$  \hspace{1cm} (10)$$

Equations (9) and (10) constitute what is called Factor Augmented Vector Autoregression (FAVAR) model. Equation (9) describes evolution of common factors $F_t$ and equation (10) is a VAR for $X_t$ augmented by lags of $F_t$. If factors $F_t$ are known then estimation of (9)-(10). First, estimate the VAR for factors

$$F_t = \Phi(L)F_{t-1} + \varepsilon F_t$$  \hspace{1cm} (11)$$

Second, for each $i$, estimate an OLS regression of $X_{it}$ on own lags and lags of factors:

$$X_{it} = \pi_i(L)'F_{t-1} + \delta_i(L)X_{it-1} + \varepsilon X_{it}$$  \hspace{1cm} (12)$$

In applications, the number of static factors is typically not greater than 10 and estimation of VAR for factors (11) therefore does not create any difficulty. Given $r$, the number of parameters to be estimated in (12) does not grow with $n$. It follows that, unlike VAR, FAVAR (11)-(12) is not prone to the curse of dimensionality. Instead, the curse is turned to a blessing: adding more variables to FAVAR helps estimate factors more precisely.
2.2 Interpretation of factors

2.2.1 Factors as state variables

One way to interpret static factors is to view them as analogs of state variables in the solution of a log-linearized Dynamic Stochastic General Equilibrium Model (DSGE) as suggested by Boivin and Giannoni (2006). The DSGE solution in the state-space form can be written as

\[ Y_t = M Z_t \]  
\[ Z_t = P Z_{t-1} + Q \eta_t \]

where \( Y_t \) is the vector of endogenous variables and \( Z_t \) is the state vector. State variables, components of \( Z_t \), include exogenous processes for shocks (e.g., monetary, fiscal, technological) and predetermined endogenous variables (e.g., capital). The components of \( Y_t \) refer usually to *broad economic concepts* (in BBE words) such as production output and tightness of the credit market, which may be unobserved directly or measured with error. In FAVAR, observed time series \( X_{it} \) serve as proxies for unobserved concepts \( Y_t \). The first FAVAR equation (9) is thus the analog of the state equation (13) while the second one (10) corresponds to the observer equation (14). The measurement errors \( u_{it} \) that are present in proxies \( X_{it} \) are likely to be serially correlated. Adding lags of \( X_{it} \) to the right hand side of the observer equation in FAVAR aims to get rid of that serial correlation.

2.2.2 DFM as parsimonious ARMA-like representation

Another way to interpret factors is to consider the dynamic factor representation of \( X_{it} \)

\[ X_{it} = \lambda_i(L) f_t + \delta_i(L) X_{it-1} + v_{it} \]  
\[ f_t = \Gamma(L) f_{t-1} + \eta_t \]

as a multi-shock analog of ARMA model. Indeed, if \( f_t = \nu_t \) is a scalar i.i.d. process and \( v_{it} \equiv 0 \) then equation (15) becomes

\[ X_{it} = \lambda_i(L) \nu_t + \delta_i(L) X_{it-1} \]

which characterizes a univariate ARMA process for \( X_{it} \).

In general, assume that there are a small number of major economic shocks \( \eta_t \) that hit the economy. For simplicity, set \( v_{it} \equiv 0 \) in (15). In general, each variable \( X_{it} \) can be represented as an infinite distributed lag of shocks:

\[ X_{it} = \alpha_i(L) \eta_t \]
One can think about (18) as the solution to a log-linearized DSGE model. Using (16), express \( \eta_t \) as

\[
\eta_t = (I - \Gamma(L)L)f_t \tag{19}
\]

and substitute it in (18) to obtain

\[
X_{it} = a_i(L)'(I - \Gamma(L)L)f_t \tag{20}
\]

In general, lag polynomial \( a_i(L)'(I - \Gamma(L)L) \) is of infinite order but it can be well approximated by a ratio of two finite-order polynomials:

\[
X_{it} = \frac{\lambda_i(L)'}{c_i(L)}f_t \tag{21}
\]

where, without loss of generality, \( c_i(0) = 1 \) so that \( c_i(L) = 1 - \delta_i(L)L \). A similar trick is used when the infinite MA representation of a univariate time series is approximated by a finite-order ARMA. Multiplying both sides of (21) by \( c_i(L) \) and moving lags of \( X_{it} \) to the right-hand side yields

\[
X_{it} = \lambda_i(L)'f_t + \delta_i(L)LX_{it} \tag{22}
\]

The latter expression coincides with equation (15) up to the idiosyncratic error term \( v_{it} \).

### 2.2.3 Summary

The preceding discussion suggests that factors can be interpreted in several ways. One interpretation relates static factors to a vector of state variable in a DSGE model that describes the economy. Vector \( F_t \) consists of exogenous processes for shocks and predetermined endogenous variables. Another potential interpretation suggests considering DFM/FAVAR representation of \( X_{it} \) as a multi-shock analog of ARMA model and dynamic factors as exogenous processes governing economic shocks. Given that I am interested in estimating the space spanned by shocks \( \eta_t \) it is not very important perhaps which particular specification I prefer. Unlike factors, shocks \( \eta_t \) have clear economic meaning. Strictly speaking, the space of \( \eta_t \) spans the space of structural shocks \( \zeta_t \). By imposing certain identifying restrictions, I can in principle identify some of the structural shocks. It is discussed in more detail below.

### 2.3 Estimation of factors and space of shocks

#### 2.3.1 Static factors

Suppose that the number of factors \( r \) is known. Then static factors can be estimated as first \( r \) principal components of data. One can extract principal components either from original or prefiltered data. In the former case, factors minimize

\[
\frac{1}{nT} \sum_{t=1}^{T} [X_t - \Lambda F_t]'[X_t - \Lambda F_t]
\]
with respect to $F_1, \ldots, F_T$, and $\Lambda$. Technically, factors $F_t$ turn out to be the eigenvectors of sample covariance matrix
\[
\frac{1}{nT} \sum_{t=1}^{T} X_t X_t'\]
that correspond its $r$ largest eigenvalues. This method is typically used when when the main objective of a study is forecasting with factors as predictors. If instead one is interested in a more precise estimation of the space of shocks then applying principal components to prefiltered data might be preferable. In this case, factors minimize
\[
\frac{1}{nT} \sum_{t=1}^{T} [(I_n - D(L)L)X_t - \Lambda F_t]'[(I_n - D(L)L)X_t - \Lambda F_t]
\]
with respect to $F_1, \ldots, F_T, \Lambda, D(L)$. Stock and Watson (2005) suggest the following two-stage iteration procedure that implements this estimation.

1. Pick initial $\hat{D}^{(0)}(L)$, compute $\tilde{X}_t = (I_n - \hat{D}^{(0)}(L)L)X_t$.
2. Estimate $\hat{F}_t^{(1)}$ as first $r$ principal components of $\tilde{X}_t$.
3. Run OLS regression of $X_t$ on $\hat{F}_t^{(1)}$ and own lags, set $\hat{D}^{(1)}(L)$ to estimated coefficients on own lags.
4. Compute $\tilde{X}_t = (I_n - \hat{D}^{(1)}(L)L)X_t$, then estimate $\hat{F}_t^{(2)}$, etc. Iterations will converge soon.

### 2.3.2 Space of dynamic shocks

Factor-based forecasts typically use static factors as predictors. If the purpose is to compute impulse responses and forecast error variance decompositions to an identified structural shock of interest, then one needs to estimate the space of dynamic shocks $\eta_t$, which, in the DFM context, coincides with the space of dynamic factor innovations. The space of $\eta_t$ spans the space of structural shocks $\zeta_t$.

Suppose that the number of dynamic factors $q$ is known. Re-write the factor equation of FAVAR in the moving average form:
\[
F_t = (I - \Phi(L)L)^{-1} \varepsilon_{F_t} = (I - \Phi(L)L)^{-1}G\eta_t
\]
and substitute it in the FAVAR equation for $X_t$ to obtain
\[
X_t = (I - D(L)L)^{-1} \Lambda(I - \Phi(L)L)^{-1}G\eta_t + u_t = A(L)G\eta_t
\]
where
\[
A(L) = (I - D(L)L)^{-1}\Lambda(I - \Phi(L)L)^{-1}
\]
I use a standard normalization:
\[
\mathbb{E}(\eta_t\eta_t') = I_q
\]
As suggested in Stock and Watson (2005), I choose \( r \times q \) matrix \( G \) such that maximizes the trace R squared of \( A(L)G\eta_t \):

\[
tr(\mathbb{E}(A(L)G\eta_tA(L)'G')) = tr(G'(\sum_{j=0}^{\infty} A_j' A_j)G)
\]

It follows that \( G \) that maximizes the trace R squared equals the matrix of eigenvectors of \( \sum_{j=0}^{\infty} A_j' A_j \) that correspond to its \( q \) largest eigenvalues. Matrix \( \sum_{j=0}^{\infty} A_j' A_j \) can be estimated using impulse responses of \( X \)'s to FAVAR innovations of factors \( \varepsilon_F t \).

Given \( G \) and impulse responses to static factors innovations \( A(L) \), I can compute impulse responses to dynamic factor innovations \( \eta_t \) as

\[
B(L) = A(L)G
\]

### 2.3.3 Number of factors

How do we know the number of static and dynamic factors? In both cases, the answer can be obtained using Bai-Ng information criteria \( IC_{p1} \) and \( IC_{p2} \) (Bai and Ng, 2002). These criteria are very similar to well-known Akaike and Schwartz information criteria. Each of them trades off goodness of fit against the number of factors:

\[
IC_{p1}(r) = ln(V(r, \hat{F}^r)) + r\left(\frac{n+T}{nT}\right)ln\left(\frac{nT}{n+T}\right)
\]

\[
IC_{p2}(r) = ln(V(r, \hat{F}^r)) + r\left(\frac{n+T}{nT}\right)ln(C_{nT}^2)
\]

where \( C_{nT}^2 = \min\{n, T\} \), \( r = 1, ..., r_{max} \) and

\[
V(r, \hat{F}^r) = \frac{1}{nT} \sum_{t=1}^{T} [(I - \hat{D}^r(L)L)X_t - \hat{\Lambda}^r \hat{F}^r_t]'[(I - \hat{D}^r(L)L)X_t - \hat{\Lambda}^r \hat{F}^r_t]
\]

The way to proceed is to choose \( r \) that minimizes \( IC_{p1} \) and/or \( IC_{p2} \). Bai and Ng prove that specific forms of the penalty term, which are different from Akaike and Schwartz, deliver consistent estimates of \( r \).

With \( r \) known, one can estimate factors \( F_t \) as first \( r \) principal components of original or prefiltered data and estimate FAVAR as discussed above. Estimated FAVAR yields estimated vector of innovations of data \( \varepsilon_X t \) where each element equals the OLS residual of respective \( X_{it} \) on own lags and lags of factors. The observer equation (10) of FAVAR suggests that

\[
\varepsilon_X t = \Lambda G \eta_t + v_t
\]

It is clear that \( \varepsilon_X t \) has a factor structure given that \( \eta_t \) and \( v_t \) are uncorrelated. Each \( \varepsilon_X_{it} \) is a linear function of \( q \) common factors \( \eta_t \) plus idiosyncratic noise \( v_{it} \). This suggests that the dimensionality of \( \eta_t \) can be determined by applying Bai and Ng information criteria to FAVAR innovations of data \( \varepsilon_X t \) (Amengual and Watson, 2007).
2.4 Identification of structural shocks

For the purposes of my study, I need to identify two structural shocks of interest, MP shock and oil shock. I borrow identification procedure for the MP shock from Stock and Watson (2005). Their approach is a modification of Bernanke, Boivin, and Eliasz’s (2005) “slow/fast” scheme and generalizes the partial identification procedure by Christiano, Eichenbaum, and Evans (1999), which is considered as the state of the art in SVAR literature.

Oil shock is identified in a novel way. Its identification exploits the decomposition of structural shocks onto slow and fast as delivered by the Stock and Watson’s partial identification procedure. From theoretical perspectives, the slow component of the oil price innovation should reflect shifts in oil supply and current demand for oil while the fast component is driven by shifts in expectations regarding future conditions of the oil market.

The distinction between contemporaneous shocks and shocks in expectations both affecting the price of oil is not a common feature of the empirical literature on oil and its effect on macroeconomy. Such a distinction, however, turns out to be crucial. Indeed, as many observers notice, the swings in the oil price in the 1970s were caused mostly by OPEC decisions on the supply side. In the mid-1980s, the massive inflow of non-OPEC supply seriously undermined the market power of OPEC, and demand shocks started to play a more important role in oil price fluctuations. Furthermore, the stochastic properties of oil price time series have changed remarkably, from series of plateaux divided by considerable price changes in the 1970s to very close to a mean-reverting process in the 1980s and later (Shiller’s remark in general discussion following Bernanke, Gertler, and Watson, 1997; Rotemberg, 2007). Any identification scheme that fails to distinguish between oil supply and demand shocks will treat oil price increases caused by OPEC decisions in the 1970s and the price spike in the 2000s, which are believed to be inspired by demand, in the similar fashion. Based on theoretical considerations, adverse oil supply shocks may lead to a recession while oil price increases caused by demand should be followed by booms. As a result, failure to distinguish between the two different kinds of oil shocks will bias the effects of both towards zero. This, in turn, may mistakenly lead to a conclusion that the the effect of oil shocks on the economy has decreased (Blanchard and Galí, 2007).

2.4.1 Monetary policy shock

Once FAVAR is estimated and mapping $G$ from the space of FAVAR innovations of static factors to the space $\eta_t$ of dynamic shocks is found, one can compute impulse responses and forecast error variance decomposi-
tions with respect to $\eta_t$. In general, however, $\eta_t$’s are not true structural shocks but their mixtures:

$$\zeta_t = H \eta_t$$

Rotation matrix $H$ relates reduced-form dynamic factor innovations $\eta_t$ to structural shocks $\zeta_t$. This matrix summarizes our identifying assumptions that are based on a certain economic model. The task is to impose a minimal set of restrictions on elements of $H$ sufficient to identify a row in $H$ that corresponds to a structural shock of interest.

I identify the structural monetary policy shock in the US using Stock and Watson’s (2005) version of the so-called slow/fast scheme suggested by Bernanke, Boivin, and Eliasz (2005). This is a block-recursive identification scheme. All variables are divided onto three groups: slow-moving series (production indices, labor market variables, consumer prices), the federal funds (FF) rate as an indicator of monetary policy, and fast-moving series (producer prices, expectations, financial market indicators). All structural shocks are also split onto three categories: slow shocks $\zeta_t^S$, the MP shock $\zeta_t^R$, and fast shocks $\zeta_t^F$.

The following identifying assumptions are made:

1. Slow-moving series can respond contemporaneously (i.e. within a month) only to slow shocks. They can respond to the MP shocks and fast shocks only with a lag.

2. The FF rate can respond contemporaneously (within a month) only to the slow shocks and the MP shock. It can respond to fast shocks only with a lag.

3. Fast-moving series can respond contemporaneously to all shocks.

One can notice that the identifying assumptions are similar to those in a well-known block-recursive scheme suggested by Christiano, Eichenbaum, and Evans (1999) for conventional structural VAR’s, which is considered as state-of-the-art in this class of identification procedures. Motivation of identifying assumptions 1-3 is therefore similar to those in CCE. For example, a surprise MP tightening or bad news about future prospects of the US economy that immediately affects financial market variables is unlikely to switch production cuts and layoffs within one month. This might be harder to motivate though if the study is done on quarterly data.

BBE in fact impose somewhat different identifying restrictions. First of all, they make the FF rate be the observed MP factor. Also, they allow the MP contemporaneously (i.e. within a month) respond to all shocks. This means, in particular, that the Fed can react to developments on the stock market very fast. One can view the two versions of the slow/fast scheme as two extremes. The BBE version assumes that Fed can process new information very quickly and react immediately. The version by Stock and Watson (2005),
which I use in this study, reflects a more conservative view about information capacity and responsiveness of Fed.

Let $B^*_j$ be $n \times q$ matrices of impulse responses to structural shocks $\zeta_t$:

$$X_t = B^*(L)\zeta_t = \sum_{j=0}^{\infty} B^*_j \zeta_{t-j}$$

Then assumptions 1-3 of the slow/fast scheme can be summarized as:

$$\varepsilon^\text{slow}_{Xt} = B^{*\text{slow},S}_{0} \zeta_{S}^t + v^\text{slow}_t$$  \hspace{1cm} (23)

$$\varepsilon^\text{FF}_{Xt} = B^{*\text{FF},S}_{0} \zeta_{S}^t + B^{*\text{FF},R}_{0} \zeta_{R}^t + v^\text{FF}_t$$  \hspace{1cm} (24)

$$\varepsilon^\text{fast}_{Xt} = B^{*\text{fast},S}_{0} \zeta_{S}^t + B^{*\text{fast},R}_{0} \zeta_{R}^t + B^{*\text{fast},F}_{0} \zeta_{F}^t + v^\text{fast}_t$$  \hspace{1cm} (25)

In practice, one needs to know the number of slow dynamic shocks $q^S$. It can be obtained by using Amengual and Watson (2007) method applied to slow-moving series only. Equation (24) implies that FAVAR innovations of slow-moving series depend on common factors $\zeta^S_t$. Hence, applying Bai-Ng criterion to $\varepsilon^\text{slow}_{Xt}$ will yield $q^S$.

Given $q^S$, one runs a reduced-rank regression of $\varepsilon^\text{slow}_{Xt}$ on the space of factor innovations $\eta_t$ imposing a restriction that the column rank of the coefficient matrix equals $q^S$. Indeed, $\eta_t$ spans the space of $\zeta_t$ and they are related through a rotation matrix $H$. Hence, $\zeta^S_t$ that are common factors for $\varepsilon^\text{slow}_{Xt}$ have to be $q^S$ orthogonal linear combinations of $\eta_t$. This is exactly what RR regression of $\varepsilon^\text{slow}_{Xt}$ on $\eta_t$ delivers:

$$\varepsilon^\text{slow}_{Xt} = B^{*\text{slow},S}_{0} H^S \eta_t + v^\text{slow}_t$$

where $H^S$ is $q^S \times q$ submatrix of $H$:

$$H = \begin{pmatrix} H^S \\ H^R \\ H^F \end{pmatrix}$$

Once the subspace of slow shocks $\zeta^S_t$ is estimated, one can exploit equation (25) to identify the MP shock. It equals the residual from OLS regression where the dependent variable is OLS projection of FAVAR innovation of Federal Funds Rate $\varepsilon^\text{FF}_{Xt}$ on the space of factor innovations $\eta_t$ and regressors are slow shocks $\zeta^S_t$. The knowledge of the space of $\zeta^S_t$ (as opposite to its components) is just enough to identify the MP shock $\zeta^R_t$.

Finally, the space of fast shocks $\zeta^F_t$ is estimated as the subspace of $\eta_t$ orthogonal to $\zeta^S_t$ and $zeta^R_t$.

To summarize, the slow/fast scheme yields three outputs: exactly identified MP shock, the space of slow shocks and the space of fast shocks. The MP shock and delayed overshooting produced by it as reported by a number of SVAR studies is the main motivation of the main idea of this paper to use structural shocks as
predictors. Nevertheless, decomposition of the rest of shocks into two groups, slow and fast, may also be helpful despite the fact that the procedure does not isolate individual structural shocks within each of the two groups. I exploit this decomposition when identifying oil demand and oil supply shocks, which is discussed in detail in the next subsection.

2.4.2 Oil shock

International oil market has two prominent features. First, oil is an exhaustible natural resource. Hence, a producer, i.e. an owner of an oil stock, always has an option to leave oil in the ground until next period if he expects that the price for oil will be higher in future. As a result, the present price of oil is linked to the expected future price of oil (Hotelling rule).

Second, the market is dominated by a group of producers, OPEC, that, if cooperate, can exercise a considerable market power. Among those, Saudi Arabia is considered as a key player. Certain amount of oil (about one third in the 1970s and about one half these days) is supplied by non-OPEC producers. Some models treat the world oil market as an oligopoly (OPEC), which sometimes practices collusive behavior, and competitive fringe (non-OPEC). Others assume that monopoly (Saudi Arabia) and competitive fringe (rest of OPEC and non-OPEC) is a reasonable approximation, especially taking into account that Saudi oil accounts for about one fourth of world supply.

It is instructive to think of oil market in terms of a simple two-period model of exhaustible resource produced by monopoly that faces a residual demand in either period (i.e. net of fringe supply). It follows that the current price of oil set by the monopoly will be affected by shifts in current and expected future demand for oil (e.g., due to productivity shocks in manufacturing) as well as shifts in the fringe supply (e.g., shortages caused by wars or increased supply by non-OPEC). One additional source of oil supply shocks might be a oligopoly regime change, from collusion to competition and vice versa. To summarize, the current market price of oil positively depends on current and expected future demand and adverse shifts in current or expected future supply.

The “slow/fast” scheme that I use to identify the MP shock delivers a partition of the space of structural shocks onto the subspace of slow shocks, the MP shock, and the subspace of fast shocks. If I believe that the oil supply shock, e.g. an abrupt reduction in the oil supply caused by a military conflict (Arab-Israeli war in 1973, invasion of Kuwait by Iraq in 1990) or a switch in the mode of oligopolistic game played by OPEC, is an empirically important economic shock for the US then it should be a part of the space of structural (dynamic) shocks as estimated by FAVAR.

The first step to the identification of the oil supply shock is to obtain FAVAR innovations of oil price. It is worth noting that, in this study, I do not include oil or any other additional series at the stage when factors
are estimated. Basically, I only replicate Stock and Watson’s (2005) exercise: I take their data (132 US series), extract static factors by principal components, estimate the space of structural shocks, and identify MP shock (the respective part of my code is actually a translation of their’s from GAUSS to MATLAB). I intentionally do not add any additional information to make my point clear. The presumption is that 132 US series are a dataset comprehensive enough to estimate all empirically important shocks that hit the US economy including the oil supply shock. Adding the the 133th series, oil price, should not improve remarkably my ability to estimate the space of shocks. But it will help me to identify the oil supply shock.

The FAVAR innovation of oil price \( \varepsilon_t^{oil} \) is obtained in the standard way as prescribed by formula (10). I estimate an OLS regression of oil price \( X_t^{oil} \) on own lags and lags of static factors \( F_t \) and set \( \xi^{oil} \) equal residuals from this regression. Then I project \( \varepsilon_t^{oil} \) on the space of structural shocks \( \zeta_t = (\xi_t^{S'}, \xi_t^R, \xi_t^{F'})' \) by OLS to get \( \xi_t^{oil} = (\xi_t^{oil,S'}, \xi_t^{oil,R}, \xi_t^{oil,F'})' \). The slow component \( \xi_t^{oil,S} \) of this projection is a natural candidate for a proxy for contemporaneous demand and supply shocks, while the fast component \( \xi_t^{oil,F} \) is likely to account for changes caused by shifts in expectations of future market conditions. Indeed, by construction, fast shocks are those that affect fast moving variables such as financial market indicators contemporaneously whereas have a delayed effect on real variables. Given that stock prices and interest rates are essentially forward looking, fast shocks have to be related to shifts in expectations caused by news, for instance.

In what follows I will be looking at impulse responses to both kinds of oil shocks, slow \( \xi_t^{oil,S} \) and fast \( \xi_t^{oil,F} \). This decomposition does not guarantee that each of those “shocks” is an actual oil shock that originates from the oil market. Instead, from theoretical perspective, \( \xi_t^{S} \) is likely to be a mix of contemporaneous demand and supply shocks while \( \xi_t^{F} \) should capture the effect of news. But such “partial” identification turns out to be sufficient for the purposes of my study. I will be more specific when discussing empirical results.

The rest of the formal identification procedure is the following. The “slow” oil shock is identifies as \( \zeta_t^{OS} = \xi_t^{oil,S} \). The non-oil slow shocks \( \zeta_t^{NOS} \) are defined as the subspace of \( \zeta_t^{S} \) orthogonal to \( \zeta_t^{OS} \). “Fast” oil shock \( \zeta_t^{OF} \) and non-oil fast shocks \( \zeta_t^{NOF} \) are identified in a similar fashion.

To summarize, the described identification procedure results in a partition on the space of structural shocks onto “slow” oil shock, a group of non-oil slow shocks, MP shock, “fast” oil shock, and a group of non-oil fast shocks:

\[
\zeta_t = (\zeta_t^{OS'}, \zeta_t^{NOS'}, \zeta_t^R, \zeta_t^{OF'}, \zeta_t^{NOF'})'
\]
2.5 Data

This study uses mostly the same data as Stock and Watson (2005) and which are available from Mark Watson’s webpage http://www.princeton.edu/~mwatson/. The dataset includes 132 US macroeconomic time series. The nominal price of petroleum, which is the average of the prices of Brent, West Texas Instrumental, and Dubai Arab Light, is used to identify oil shocks. The data is at monthly frequencies and covers the period between 1959:1-2003:12. All time series, except oil, come from Hoover Analytics database as indicated in Bernanke, Boivin, and Eliasz (2005) and Stock and Watson (2005). The oil price data is taken from the IMF’s International Financial Statistics database. I also use four IMF’s commodity price indices: metals, agricultural raw materials, food, and beverages.

In accordance with standard practices, the original time series were transformed to obtain stationarity. I apply the transformations of data that are conventional in the literature (Bernanke, Boivin, and Eliasz, 2005; Stock and Watson, 2005). In brief, real quantities (indices) such as industrial production and employment are log-differenced; price levels such as PPI and CPI are twice log-differenced; nominal exchange rates are log-differenced; nominal interest rates are differenced. The nominal oil price is log-differenced, which is also conventional (Blanchard and Galí, 2007). To some extent, the transformation I chose for the oil price, contradicts one that was applied to other nominal price variables, e.g. the index of sensitive materials prices, which is log-differenced twice rather than once. Nevertheless, I decide to be conservative and follow the literature. Re-running the entire analysis with oil being twice log-differenced might be a useful robustness check worth to do.

Following Stock and Watson (2005), I adjust data to outliers at the stage when factors are estimated. The reason is that the estimates of factors are likely to be very sensitive to outliers. An observation is classified as an outlier if its absolute deviation from the sample median exceeds the inter-quarter range multiplied by the factor of 6. Every outlier is replaced by the median of four preceding values.

3 Empirical results

In this section, I report impulse responses to identified structural oil shocks under various identifying assumptions. All identification strategies are based on block-recursive approach, which is described in more detail in the previous section. Standard-error bands for impulse responses are not shown but will be computed using Kilian’s bootstrap method (Kilian, 1998) and added later.

Each graph contains two impulse responses, one actual and the other counterfactual. Counterfactual impulse responses are produced by Sims – Zha (2006) monetary policy experiment. Under such an experiment, the endogenous response of monetary policy to the oil shock is shut off so that, in the aftermath of a shock
the policy variable, federal funds rate, remains unchanged. In practice, this can be achieved, in principle, by generating a series of expansionary monetary surprises. Addressing potential Lucas critique, Bernanke, Gertler, and Watson (2004) assume that such a policy change, although long-lived, is not permanent and that it takes time for economic agents to learn that it happened. As a result, one can expect that the severity of Lucas critique with respect to this policy experiment is moderate. I make a similar assumption that the change in the regime by the Fed is temporary and agents learn about it slowly.

I report impulse responses only for six key variables of interest: total industrial production index, total private non-farm employment, federal funds rate, 10-year Treasury bond rate, PPI for finished goods, and CPI for finished goods.

3.1 Number of factors

Using the methodology borrowed from Stock and Watson (2005) and described in the previous section, I estimated the number of factors that presumably drive the US economy. The number of static factors is 9, the number of dynamic factors is 7, the number of “slow” dynamic factors is 4. All three numbers coincide with those found by Stock and Watson (2005). It is very likely that the original set of 132 time series is large enough to guarantee that the space spanned by all major structural shocks that drive the US economy is estimated so that adding the oil price does not change the picture. Adding oil may be helpful, however, since one can hope to use the oil price for the purpose of identification of the structural oil shock, by which I understand exogenous shifts in the world supply of oil and/or in the demand for oil from the rest of the world.

3.2 Sims-Zha experiment under conventional identification

Figure 1 shows impulse responses obtained under conventional identification of the oil shock (Rotemberg and Woodford, 1996; Blanchard and Gali, 2007). It is assumed that contemporaneously, i.e. within one month, oil price responds only to shocks in oil supply and it reacts to all other shocks with a lag. This identifying assumption corresponds to assigning the oil price the first rank in a recursive scheme applied to a conventional VAR.

The impulse responses presented on Figure 1 suggest that oil shocks have a recessionary effect on the economy but the contribution of the monetary policy response is moderate contrary to the conclusion in Bernanke, Gertler, and Watson (1997).

One problem with the identification scheme when oil price is ordered first is that it fails to separate oil demand from oil supply shocks. As Rotemberg (2007) notices, the application of this procedure can be justified if one works with data prior to mid-1980s for two reasons. First, during that period the market
Figure 1: Impulse responses to oil shock under conventional identification
power of OPEC was higher than in the subsequent time period when the supply from non-OPEC producers had increased substantially compared with the 1970s. Second, before 1983, OPEC practiced setting so-called *posted prices* while they switched to allocating quotas afterwards.

I am not aware of any identification procedure available in the literature that distinguishes between oil demand and oil supply shocks. Kilian (2007) is an exception but his approach is based on a recursive ordering within a three-variable VAR and is subject to a potential omitted variable bias. In the next subsection, I try to fill this gap.

3.3 Sims-Zha experiment under alternative identification

Here, I identify two kinds of oil shocks, “slow” and “fast”. Examination of relevant impulse responses suggests that the slow oil shock can be interpreted as oil supply shock and fast oil shock as oil demand shock. As discussed above, oil supply shock is identified as the OLS projection of FAVAR innovation of the oil price on the space of slow shocks and the oil demand shock as the OLS projection on the space of fast shocks.

Impulse responses on Figures 2 and 3 show quite a different picture from Figure 1. Specifically, both the policy reaction of the Fed and the behavior of the economy are different for the two different shocks. Indeed, Fed cuts interest rates in response to a the positive (i.e. adverse) realization of the slow oil shock. This shock turns the economy into a recession. Holding the interest rates unchanged (as under Sims – Zha experiment) would make the recession even deeper. This finding is consistent with the view that monetary policy in the 1970s, when two major oil supply shocks happened, was accommodating (Blanchard and Galí, 2007).

A positive realization of the fast oil is followed by a boom and the Fed raising interest rates. This looks very much like a shift in the demand for oil that occurs due to current or perceived productivity growth. A positive productivity shock would raise demand for commodities including oil and this will drive up the price of oil and other commodities. Strictly speaking, fast oil shock is not an oil shock in the sense that, unlike the oil supply shock, it is not specific to the oil market but is instead economy-wide.

Figures 4 and 5 give support to the interpretation of slow and fast oil shocks as demand and supply shocks respectively. They show impulse responses of two stock market indicators and metal price index to the two identified oil shocks. All these three variables are essentially forward looking. The slow oil shock drives the stock market and metal prices down while the fast oil shock is associated with higher stock and metal prices.
Figure 2: Impulse responses to “slow” oil shock
Figure 3: Impulse responses to “fast” oil shock
Figure 4: Impulse responses to “slow” oil shock
Figure 5: Impulse responses to “fast” oil shock
4 Conclusion

This paper offers a new procedure for identification of oil shocks that distinguishes between oil demand and oil supply shocks. Using this identification scheme, I revisit the question about the role of systematic monetary policy response in amplification of oil shocks. I reach a conclusion that is quite opposite to the consensus view in the literature. The systematic monetary policy did not contribute to recessions induced by oil shocks. Furthermore, the monetary policy response to oil supply shocks was accommodating rather then anti-inflationary.

References


