Time-Varying Effects of Oil Supply Shocks on the US Economy

By Christiane Baumeister and Gert Peersman

Using time-varying BVARs, we find a substantial decline in the short-run price elasticity of oil demand since the mid-1980s. This finding helps explain why an oil production shortfall of the same magnitude is associated with a stronger response of oil prices and more severe macroeconomic consequences over time, while a similar oil price increase is associated with smaller output effects. Oil supply shocks also account for a smaller fraction of real oil price variability in more recent periods, in contrast to oil demand shocks. The overall effects of oil supply disruptions on the US economy have, however, been modest. (JEL E31, E32, Q41, Q43)

This paper studies the relationship between oil prices and US macroeconomic performance since 1974. There is considerable evidence that this relationship has been unstable over time (see, e.g., Edelstein and Kilian 2009; Herrera and Pesavento 2009; Blanchard and Galí 2010; Ramey and Vine 2011). In particular, several researchers have noted a substantial decline in the macroeconomic consequences of oil price shocks. One of the reasons for this temporal instability is the fact that oil price shocks are merely symptoms of underlying oil supply and oil demand shocks. As the composition of these structural oil supply and oil demand shocks evolves, so does the dynamic correlation between the US economy and the real price of oil (see, e.g., Peersman and Van Robays 2009; Kilian 2009a).

Even if we distinguish between oil demand and oil supply shocks, however, there are additional reasons why the response of the US economy to these shocks may have changed over time. Such time-varying effects may come about, for example, because of variation in the oil intensity of economic activity, because of changes in the regulation of energy markets, or because of changes in the composition of automobile production and the overall importance of the US automobile sector (see, e.g., Edelstein and Kilian 2009; Ramey and Vine 2011). Likewise, changes in...
capacity utilization rates in crude oil production and the transition toward a market-based system of oil trading in the 1980s may account for additional instabilities in the responses to oil supply shocks (see, e.g., Hubbard 1986). Indeed, there is a consensus in the literature that the short-run price elasticity of oil demand has declined since the mid-1980s, although the extent of these changes has proved hard to pin down (see, e.g., Hughes, Knittel, and Sperling 2008; Dargay and Gately 1994, 2010).

Time-varying responses are not allowed for in the existing literature on the effects of oil supply shocks on US macroeconomic aggregates. It is by now well understood that oil price shocks are not the same as oil supply shocks (see Kilian 2008a). As a result, in recent years, considerable work has been done to explicitly identify oil supply shocks and to understand their effect on the real price of oil and on the US economy. Hamilton (1983) first remarked on the coincidence of oil supply disruptions and subsequent oil price surges during 1948–1972. Building on this insight, Hamilton (2003) developed a quantitative dummy measure based on physical oil supply disruptions associated with major political crises in oil-producing countries and investigated its predictive power for changes in the price of oil and for US real GDP. Kilian (2008b,c) proposed an alternative measure of exogenous oil supply shocks obtained by constructing explicit counterfactuals for all major oil producers, and studied the responses of inflation and real output in major industrial economies to this measure of oil supply shocks. Yet another approach developed by Kilian (2009a) has relied on exclusion restrictions in structural vector autoregressive (VAR) models of the global oil market to identify oil supply shocks.

A common feature of all these empirical studies is that they rely on time-invariant regressions. Implicitly, it is assumed that the effect of oil supply shocks on the real price of oil and on macroeconomic aggregates has not changed over time. This observation raises the question of how reliable existing estimates of the effects of oil supply shocks are.

In this paper, we use a time-varying parameter structural VAR model to investigate how the effects of oil supply shocks on the US economy have changed over time. Our analysis incorporates several innovations relative to the previous literature. First, rather than imposing an arbitrary sample split, as some previous studies have done, we let all model coefficients evolve continuously, allowing the data to tell us when and how any changes may have occurred. Second, we identify oil supply shocks not based on contemporaneous exclusion restrictions but, instead, based on the sign restrictions that an oil supply shock moves oil prices and oil production in opposite directions. Ours is the first application of this identification approach in the context of the global oil market. Third, we allow for time-varying heteroskedasticity in the VAR innovations that accounts for changes in the magnitude of structural shocks and their immediate impact. This feature is particularly important in the present setting given the observed increase in oil price volatility and the reduction in macroeconomic volatility during the Great Moderation. Fourth, in

---

1 This quantitative dummy variable takes the value of the amount of oil production lost in the affected countries, expressed as a percentage of world oil production, at the time of an historical episode, and zero otherwise.

2 Sign restrictions for modeling oil market dynamics have subsequently been adopted by Peersman and Van Robays (2009, 2012), Baumeister and Peersman (forthcoming), Kilian and Murphy (2012, forthcoming), and Lippi and Nobili (2012), among others.
addition to shedding light on time variation in the responses of US real GDP and consumer prices to oil supply shocks, our analysis also allows us to obtain estimates of changes over time in the short-run price elasticity of oil demand, complementing the existing literature. By construction, this demand elasticity corresponds to the ratio of the impact response of world oil production to the real price of oil triggered by an unanticipated shift in the world oil supply curve along the contemporaneous oil demand curve.

Our analysis yields several intriguing results. First, we find that even though the effect on economic aggregates of an oil supply disruption associated with a 1 percent decrease in world oil production has increased over time, the effect on US real GDP of an oil supply disruption associated with a 10 percent increase in the real price of oil has declined over time. We show that this evidence cannot be explained merely by changes in the variance of oil supply shocks. Rather, it requires a decline in the price elasticity of oil demand over time, such that a given decrease in the quantity of oil supplied on impact is associated with a larger increase in the real price of oil on impact, and/or suitable changes over time in the dynamic responses of all model variables to an oil supply shock. Second, the contribution of oil supply shocks to the variability of the real price of oil has moderately declined over time, indicating a larger role for oil demand shocks. It is reassuring that accounting for time variation does not overturn this important insight from the recent literature. Third, although the contribution of oil supply shocks to the variance of macroeconomic aggregates is nonnegligible, they explain only a small part of the recessions since the 1970s and of the “Great Inflation.”

The remainder of the paper is organized as follows. Section I outlines the econometric methodology and describes the identification strategy in more detail. Section II presents the main empirical results and evaluates the robustness of our findings. In Section III, we discuss some potential explanations for a less price-elastic oil demand curve since the second half of the 1980s. Section IV contains the concluding remarks.

I. Empirical Methodology

The apparent instabilities in the oil-macro relationship suggest that modeling the transmission of oil supply shocks adequately requires an empirical framework that can account for changes over time. The anecdotal evidence presented in the introduction points toward a gradual evolution of the interaction between the oil market and the US economy. The idea of a slow-moving but continuous adjustment is also in line with the adaptive behavior of economic agents that results from an ongoing learning process. For example, Primiceri (2005) makes the case that the aggregation among agents’ assessments tends to be reflected in smooth changes since not all agents would be expected to update their beliefs at the same time. This line of reasoning suggests that the appropriate modeling approach is a time-varying parameter model (TVP-VAR) featuring smoothly evolving coefficients and heteroskedasticity in the innovations.

Although the possibility of abrupt breaks cannot be excluded a priori, it can be shown that a time-varying parameter model is capable of capturing such discrete
shifts should they occur. As the Monte Carlo study in section 1 of the online Appendix illustrates, drifting coefficient models in practice can successfully track processes subject to structural breaks or regime shifts. In contrast, models of discrete shifts cannot accommodate smooth structural change.

A. A VAR with Time-Varying Parameters and Stochastic Volatility

We model the joint behavior of global oil production, the real US refiners’ acquisition cost of imported crude oil (obtained by deflating the nominal price by US CPI), US real GDP, and US consumer prices as a VAR($p$) with time-varying parameters and stochastic volatility:

\[
y_t = c_t + B_{1,t}y_{t-1} + \cdots + B_{p,t}y_{t-p} + u_t = X_t\theta_t + u_t,
\]

where $y_t \equiv [\Delta q_{oil}, \Delta p_{oil}, \Delta gd p_t, \Delta cp i_t]'$ and $\Delta$ denotes the first difference operator.\(^3\) All variables were transformed to nonannualized quarter-on-quarter rates of growth by taking the first difference of their natural logarithm. The time-varying intercepts $c_t$ and the matrices of time-varying coefficients $B_{1,t}...B_{p,t}$ are stacked in $\theta_t$, and $X_t$ is a matrix including lags of $y_t$ and a constant to obtain the state-space representation of the model. The data frequency is quarterly. The overall sample covers the period 1947:I–2011:I, but the first 25 years of data are used as a training sample to generate the priors for estimation over the actual sample period, which starts in 1974:I. This is the earliest possible starting date given that before 1974 the oil price was regulated, which impairs the use of standard time-series models of the oil market even when time variation is allowed for (see, e.g., Alquist, Kilian, and Vigfusson 2013). The lag length is set to $p = 4$ to allow for sufficient dynamics in the system and to capture lags in the transmission of oil shocks (see, e.g., Hamilton and Herrera 2004). The $u_t$ in the observation equation is an unconditionally heteroskedastic disturbance term that is normally distributed with a zero mean and a time-varying covariance matrix $\Omega_t$ that can be decomposed in the following way:

\[
\Omega_t = A_t^{-1} H_t (A_t^{-1})'.
\]

$A_t$ is a lower triangular matrix that models the contemporaneous interactions among the endogenous variables and $H_t$ is a diagonal matrix that contains the stochastic volatilities

\[
A_t = \begin{bmatrix}
1 & 0 & 0 & 0 \\
\alpha_{21,t} & 1 & 0 & 0 \\
\alpha_{31,t} & \alpha_{32,t} & 1 & 0 \\
\alpha_{41,t} & \alpha_{42,t} & \alpha_{43,t} & 1 \\
\end{bmatrix}
\quad \quad \quad
H_t = \begin{bmatrix}
h_{1,t} & 0 & 0 & 0 \\
0 & h_{2,t} & 0 & 0 \\
0 & 0 & h_{3,t} & 0 \\
0 & 0 & 0 & h_{4,t} \\
\end{bmatrix}.
\]

\(^3\) A detailed description of the data used in this paper can be found in Appendix A.
The drifting coefficients are meant to capture possible nonlinearities or time variation in the lag structure of the model. The multivariate time-varying variance covariance matrix allows for heteroskedasticity of the shocks and time variation in the simultaneous relationships between the variables in the system. Allowing for time variation in both the coefficients and the variance covariance matrix leaves it up to the data to determine whether the time variation of the linear structure comes from changes in the size of the shock and its contemporaneous impact or from changes in the propagation mechanism. Let $\alpha_t = [\alpha_{21,t}, \alpha_{31,t}, \alpha_{32,t}, \alpha_{41,t}, \alpha_{42,t}, \alpha_{43,t}]'$ be the vector of nonzero and nonone elements of the matrix $A_t$, and $h_t = [h_{1,t}, h_{2,t}, h_{3,t}, h_{4,t}]'$ be the vector containing the diagonal elements of $H_t$. Following Primiceri (2005), the three driving processes of the system are postulated to evolve as follows:

$$
\theta_t = \theta_{t-1} + \nu_t, \quad \nu_t \sim N(0, Q)
$$

$$
\alpha_t = \alpha_{t-1} + \zeta_t, \quad \zeta_t \sim N(0, S)
$$

$$
\ln h_{i,t} = \ln h_{i,t-1} + \sigma_i \eta_{i,t}, \quad \eta_{i,t} \sim N(0, 1).
$$

The time-varying parameters $\theta_t$ and $\alpha_t$ are modeled as driftless random walks. We impose a stability constraint on the evolution of the time-varying parameters $\theta_t$ to enforce stationarity of the VAR system. Specifically, we include an indicator function that selects only stable draws i.e., the indicator function $I(\theta_t) = 0$ if the roots of the associated VAR polynomial are inside the unit circle, as in e.g., Cogley and Sargent (2005). The elements of the vector of volatilities $h_t$ are assumed to evolve as geometric random walks, independent of each other. The error terms of the three transition equations are independent of each other and of the innovations of the observation equation. In addition, we postulate a block-diagonal structure for $S$ of the following form:

$$
S = \text{Var}(\zeta_t) = 
\begin{bmatrix}
S_1 & 0_{1x2} & 0_{1x3} \\
0_{2x1} & S_2 & 0_{2x3} \\
0_{3x1} & 0_{3x2} & S_3
\end{bmatrix},
$$

with $S_1 \equiv \text{Var}(\zeta_{21,t})$, $S_2 \equiv \text{Var}(\zeta_{31,t}, \zeta_{32,t})$, and $S_3 \equiv \text{Var}(\zeta_{41,t}, \zeta_{42,t}, \zeta_{43,t})$, which implies that the nonzero and nonone elements of $A_t$ belonging to different rows evolve independently. Thus, the shocks to the coefficients of the contemporaneous relations among the endogenous variables are assumed to be correlated within equations, but uncorrelated across equations. As pointed out by Primiceri (2005), this block-diagonal structure has the desirable property of simplifying inference and increasing the efficiency of the estimation algorithm in an already highly-parameterized model since it permits to estimate the covariance states equation by equation. We estimate this model using Bayesian methods described in Kim and Nelson (1999). An overview of the prior specifications and the estimation strategy (Markov Chain Monte Carlo algorithm) is provided in Appendix B.
B. Identification of Oil Supply Shocks

It is now widely accepted that oil prices are not only determined by supply-side factors but also driven by demand conditions (see Barsky and Kilian 2002, 2004; Peersman 2005; Kilian 2008a; Hamilton 2009a,b; Kilian 2009a,b). Innovations in the oil price equation of a VAR model are not an adequate measure of exogenous variation in oil supply because they also capture shifts in the demand for crude oil. The resulting estimates of macroeconomic effects only represent the consequences of an average oil price shock determined by a combination of supply and demand factors. Blanchard and Galí (2010), among others, make the case that this distinction does not matter because an oil price shock triggered by increased demand for oil in one country can still be experienced as a supply shock by the remaining countries. This presumption is, however, very stringent in light of the results of Peersman and Van Robays (2009) and Kilian (2009a) who show that there exist important differences in the responses of macroeconomic aggregates depending on the underlying source of the oil price shock. Intuitively it is clear that an increase in the real price of oil induced by favorable global economic conditions exerts a different influence on the macroeconomic performance than one due to oil supply disruptions resulting from a war (see Kilian 2009a; Rotemberg 2010).

Kilian (2009a) disentangles oil supply from oil demand shocks based on contemporaneous exclusion restrictions in a monthly vector autoregression that includes world oil production and the real price of crude oil. An oil supply shock is identified as the sole disturbance that has an immediate influence on the level of oil production. Accordingly, global oil production does not respond instantaneously to oil demand shocks, which implies that the short-run oil supply curve is vertical. This assumption, while arguably tenable at the monthly frequency, is, however, less appropriate when quarterly data are used, as in our study.4

Therefore, we propose a new approach to identifying structural oil supply shocks that involves sign restrictions on the estimated time-varying impulse responses. This approach builds on other applications of sign restrictions including the work of Faust (1998), Davis and Haltiwanger (1999), Canova and De Nicoló (2002), Uhlig (2005), and Peersman (2005), among others. Related work on time-varying parameter VAR models with sign restrictions includes Benati and Mumtaz (2007); Canova and Gambetti (2009); and Hofmann, Peersman, and Straub (2012). Our identifying restrictions are based on insights derived from a basic supply and demand model for the global oil market, which is represented by world oil production and the real price of crude oil. An oil supply shock is identified as a disturbance that shifts the upward-sloping oil supply curve along the downward-sloping oil demand curve and hence, results in an opposite movement in oil production and in the real price of oil. The identifying assumptions are that after a negative oil supply shock, world oil production declines and the real price of oil increases. No constraints are imposed

4 An alternative approach in the literature has been to quantify oil supply shocks directly through suitable counterfactual thought experiments. Hamilton (2003) and Kilian (2008b,c) have developed measures of oil supply shocks based on physical oil supply disruptions in the wake of major exogenous political events in oil-producing countries. The advantage of defining oil supply shocks within a structural VAR model is that it also captures endogenous production responses among oil producers to exogenous oil supply disruptions.
on the responses of US real output and of consumer prices. The reactions of these variables will be determined by the data. The sign restrictions are imposed to hold for four quarters following the shock, consistent with widely held beliefs about oil price dynamics in the literature (Hamilton 2003; Kilian 2008c). As a consequence of these inequality constraints, our identification scheme does not deliver exact identification (see, e.g., Fry and Pagan 2011).

Also note that the model is only partially identified in the sense that only the oil supply shock is explicitly identified and the conglomerate of residual oil demand shocks has no structural interpretation, which is in line with the focus of this paper, as outlined in the introduction. In Baumeister and Peersman (forthcoming), we extend the present identification strategy to explore, in more detail, changes in global oil price dynamics. Notably, we include a measure of global real economic activity, which enables us to differentiate between oil demand shocks driven by the global business cycle and oil-specific demand shocks. While Baumeister and Peersman (forthcoming) study the time-varying dynamics of the real price of oil as a function of all three types of oil market shocks, here we are concerned with understanding the transmission of oil supply shocks to the US economy.

II. Results

We begin our analysis with a discussion of the impulse response estimates in Section IIA. The computation of the time-varying generalized impulse response functions and the implementation of the sign restrictions are described in Appendix C. In Section IIB, we assess the quantitative importance of oil supply shocks and in Section IIC, we conduct several sensitivity analyses.

A. Responses to an Oil Supply Shock

In structural VAR models, it is conventional to report the responses of the endogenous variables to one standard deviation shocks. The problem in time-varying structural VAR models is that a one standard deviation shock corresponds to a different-sized shock at each point in time, which in turn may affect the scale of the impulse response functions. This fact complicates the analysis of the dynamic effects of an oil supply shock on the macroeconomy as we move across time. In order to compare the economic consequences across episodes, it is necessary to establish a benchmark scenario against which the time-varying responses can be assessed. Since oil supply shocks are characterized by the joint response of oil production and the real price of oil, both variables are suitable candidates for such a benchmark scenario. Granting that in a time-varying parameter model one cannot make the nature of an oil supply shock identical over time, one can consider the less ambitious thought experiment of making these shocks comparable along one of those dimensions at a time.

Normalization on Oil Quantity.—For example, the oil supply shock can be normalized on the quantity of oil supplied, which helps relate our analysis to the previous literature on oil supply shocks that has focused on physical disruptions in the
production of crude oil (see Hamilton 2003; Kilian 2008b,c). The dynamic effects of exogenous oil supply shocks, normalized so that they correspond to a 1 percent decrease in global oil production on impact at each point in time, are shown in panel A of Figure 1 for the median responses together with the sixteenth and eighty-fourth percentiles of the posterior distribution. We plot the reaction of the real price of oil for the quarter in which the shock occurs. For the macroeconomic variables instead, we depict the responses four quarters after the shock. This convention is adopted throughout the paper. The reason for this choice is that the greatest effect on real GDP is expected to occur with a delay of about one year (see Hamilton 2008). With regard to consumer prices, our objective is not so much to capture the direct effects of oil supply shocks, but to gauge the pass-through over time. Peersman and Van Robays (2009), Herrera and Pesavento (2009), and Clark and Terry (2010) show that the greatest effect of an oil price shock occurs after about three to four quarters. Notice, however, that our conclusions are robust for alternative horizons. The estimated responses have been accumulated and are shown in levels. The impact response of the real price of crude oil with respect to a 1 percent shortfall in world oil production increases substantially over time, from an average value of 3 percent in the 1970s and 1980s to 8 percent in the 1990s, up to 15 percent in the 2000s, with a spike of 28 percent in 2008. The oil price increases triggered by a given reduction in oil production on impact are in turn more disruptive to the economy in the second part of the sample. The accumulated loss in real GDP growth is about twice as big in the 1990s and almost three times as big in the 2000s as it is in the 1970s. The responses of consumer prices get more pronounced from the mid-1990s onward, and climb considerably in the 2000s. This evidence underscores the importance of allowing for time variation in studying the dynamic effects of physical oil production shortfalls.

Normalization on Oil Price.—Given that the focus of much previous research has been on the effects of an unanticipated increase in the price of oil, we now consider alternatively the effect of an oil supply shock normalized so that it raises the real price of oil by 10 percent on impact at each point in time. The latter thought experiment is used by Blanchard and Galí (2010) as a benchmark for their intertemporal comparison. The normalized time-varying impulse responses are shown in panel B of Figure 1. For this scenario, we find a more muted reaction of economic activity in the latter part of the sample. This finding complies well with existing empirical evidence on the time-varying effects of oil price shocks (e.g., Edelstein and Kilian 2007; Herrera and Pesavento 2009; Blanchard and Galí 2010; Ramey and Vine 2011). This experiment shows that a 10 percent rise in the real price of oil is currently accompanied by an oil production shortfall of 0.5 to 1.5 percent on impact. To elicit the same oil price move in the 1970s, a decline in the physical supply of crude oil of up to 12 percent is required.\(^5\)

\(^5\)To gain a better idea about the extent of time variation over the sample period, we also examined the joint posterior distribution of impulse responses for all models across selected pairs of representative dates following the approach proposed by Cogley, Primiceri, and Sargent (2010). The results of these bilateral diagnostics, which are relegated to section 2.1 of the online Appendix, largely confirm the pattern of time variation described for both normalizations.
How can we interpret the decrease in the response of global oil production to a given rise in the real price of oil over time? The simple supply-and-demand diagram of the oil market displayed in Figure 2 illustrates the implications of this finding. A shift of the oil supply curve along a given demand curve implies that the ratio of the quantity response over the price response is invariant to the extent to which the oil supply curve is shifted exogenously, i.e., the ratio $\Delta Q/\Delta P$ is the same for moving

\[ \text{Real price of crude oil} \]

\[ \text{World oil production} \]

\[ \text{Real GDP} \]

\[ \text{Consumer prices} \]

\[ \text{Notes: Panel A: Oil supply shock normalized to a 1 percent decrease in world oil production. Panel B: Oil supply shock normalized to a 10 percent increase in the real price of crude oil.} \]
from $S_1$ to $S_2$ and from $S_1$ to $S_2'$, as shown in panel A. This means that a change in the size of oil supply shocks alone cannot explain the fact that the impact responses of world oil production decrease from $Q_2$ to $Q_2'$ over time for a given increase in the real price of crude oil. This result can only be explained by a steepening of the oil demand curve as illustrated in panel B. To obtain an equilibrium outcome characterized by the intersection of the oil supply and demand curves at $P_2$ and $Q_2'$, it is not enough for the supply curve to shift by less, but we need the slope of the demand curve to steepen. In particular, the intersection of $S_2'$ with $D$ is not consistent with the equilibrium solution.

While we cannot exclude the possibility that the variability of the underlying oil supply shocks has changed over time, only a steepening of the oil demand curve (which is equivalent to a decline in the short-run price elasticity of oil demand over time) can reconcile the fact that the impact quantity response is smaller for a given impact price increase. This is indeed what we find in the data. While the average value of the price elasticity is around $-0.6$ in the early part of the sample with the exception of the 1979–1980 episode, that elasticity declines considerably starting in the mid-1980s, and reaches a low of $-0.1$ toward the end of the sample. In other words, it now takes a smaller exogenous reduction in world oil production to push up the price of oil by the same amount. Note that by construction variations in the expected time path of future oil production are reflected in the impact price and quantity responses in our model, which underlie the short-run price elasticity of oil demand.

In sum, we can reproduce the findings of other researchers that a given supply-driven oil price increase is associated with a decreasing effect on real economic activity later in the sample, and we show that this finding is consistent with the result that a given reduction in oil production leads to a greater decline in real GDP over time, once we take the steepening of the oil demand curve into account and allow for time-varying dynamics in the propagation of the oil supply shock.
B. Quantitative Importance of Oil Supply Shocks

Impulse responses are only informative of the transmission of a one-time oil supply shock but do not tell us how important oil supply shocks have been on average, nor do they tell us how much of the historical variation in oil market and macroeconomic variables is due to oil supply shocks. To shed some light on these questions, we now examine the forecast error variance decompositions and the historical decompositions based on our TVP-VAR model. As in the case of the impulse responses, we allow the parameters used in computing these decompositions to evolve over time to account for changes in the structure of the economy. In other words, these parameters do not remain fixed at their value in the impact period but evolve with the horizon according to their law of motion.

Variance Decomposition. Figure 3 displays the evolution of the median of the contribution of oil supply shocks to the forecast error variance of all four endogenous variables after 20 quarters with sixteenth and eighty-fourth percentiles of the posterior distribution. The contribution of oil supply disturbances to the variance of US real GDP growth and CPI inflation consistently ranges between 15 and 20 percent. The share of output volatility attributable to oil supply shocks oscillates moderately over time, whereas the fraction of movements in consumer price inflation induced by oil supply shocks gradually increased since the early 2000s. The latter finding is not surprising given...
that the volatility of consumer price inflation has decreased over time, while the impact of oil supply shocks on inflation has increased slightly. We conclude that oil supply shocks are still relevant for macroeconomic fluctuations.

The fraction of the variance of global crude oil production growth explained by oil supply shocks fluctuates between 25 and 35 percent in the early part of the sample, but has stabilized since the early 1990s at around 30 percent with the exception of a sharp drop in 2008. Figure 3 also shows that the contribution of oil supply shocks to the variability of changes in the real price of oil declines from around 30 to 35 percent in the first half of the sample to around 20 to 25 percent in the second half.

These estimates indicate that oil supply shocks have become a less important source of oil price movements in recent years. Although we do not identify specific oil demand shocks, we can view all shocks other than the oil supply shock collectively as an oil demand shock. While the contribution of these oil demand shocks to the variability in world oil production has remained relatively stable over time, they are responsible for a substantial share of the volatility in the real oil price, which has increased notably since the early 1990s. This evidence is consistent with the empirical results of Kilian (2008a, 2009a), reflected in the conventional wisdom that “demand increases rather than supply reductions have been the primary factor driving oil prices over the last several years” (Hamilton 2008, 175).

**Historical Decomposition.**—The historical contribution of oil supply disturbances to the four endogenous variables is presented in Figure 4. The dashed line shows the actual time series relative to its baseline forecast and the solid line shows the cumulative effect of the estimated oil supply shocks on the evolution of each variable. The difference between the actual data and the cumulative contribution of oil supply shocks can be interpreted as the cumulative contribution of the composite of oil demand shocks.\(^6\)

The overall cumulative effect of oil supply disruptions on the US economy has been modest. Oil supply shocks contributed to the 1991 recession and slowed the economic boom of 1999–2000 when OPEC and non-OPEC countries jointly decided to cut oil production, but they appear not to be crucial for other US recessions according to the historical decomposition of real GDP growth. With regard to inflation, Figure 4 reveals that oil supply shocks also explain little of the Great Inflation. Despite the fact that there is some contribution, the bulk of the inflation in the 1970s and early 1980s is explained by other shocks. While this insight is apparently in contrast with popular perception, it is consistent with related evidence in

---

\(^6\)Note that historical decompositions are computed on the basis of historical shocks and estimates of the time-varying impulse response coefficients. In line with the existing literature on time-varying parameters models (e.g., Benati and Mumtaz 2007; Canova and Gambetti 2009), we condition on the entire history to date, as we move through time, when computing impulse responses to an oil supply shock. In particular, the impulse response functions are defined as the difference between two conditional expectations, where both conditioning information sets contain everything that has happened up to that point in time. The impulse responses thus trace out the future path of the endogenous variables conditional upon that history which reflects the effects of all previous shocks. Accordingly, the dynamic effects of oil supply shocks could potentially also be affected by earlier demand shocks, whereas the effects of oil demand shocks could be affected by earlier supply shocks. Such indirect effects of oil supply shocks on the evolution of the variables over time are hence not captured by the solid lines in Figure 4.
the literature. In particular, Barsky and Kilian (2002, 2004) showed that shifts in monetary policy regimes associated with demand shifts in the oil market were the source of the stagflationary experience of the 1970s.

The contribution of the estimated oil supply shocks to the evolution of the real price of crude oil changes from episode to episode. We find an important role for oil supply shocks to explain the oil price increases in 1987, 1989, and particularly the oil price spike after Iraq’s invasion of Kuwait in August 1990. Oil supply shocks also played a big role in 1999. On the other hand, we find a dominant role for shocks at the demand side of the oil market during the events of 1979, which is consistent with the conjecture in Barsky and Kilian (2002), who argue that oil supply shocks were never the sole driving force behind the fluctuations in the real price of crude oil in 1979. The latter also matches the empirical evidence in Kilian (2009a). Similarly, oil demand shocks made a considerable contribution to the oil price collapse in...
1986 and to oil price movements since 2002. This implies that the oil price surge in 2007–2008 and the subsequent dramatic collapse were predominantly driven by oil demand shocks, a finding which is consistent with Hamilton (2009b) and Kilian (2009a). It is noteworthy that notwithstanding the striking differences in impulse response dynamics over time, the historical decomposition of the real price of crude oil is broadly similar to results from a time-invariant model in Kilian (2009b).

C. Sensitivity Analysis

Alternative Models of Time Variation.—In the benchmark model we postulated a drifting coefficient model. We now illustrate the importance of allowing for smooth time variation as opposed to a one-time break in 1986. Figure 5, panel A displays the median responses of the four endogenous variables, after an oil supply shock identified with sign restrictions and normalized to a 1 percent decrease in oil production in a fixed-coefficient VAR model estimated over the two subperiods 1974:I–1985:IV and 1986:I–2011:I, together with the sixteenth and eighty-fourth percentiles. It turns out that the impact response of the real price of oil is much larger in the second subsample, reaffirming our substantive finding of a less price-elastic oil demand curve. In contrast to our baseline results, however, there is no compelling evidence for time variation in the impulse response functions of real GDP and consumer prices across subsamples given that the posterior intervals overlap. This analysis highlights that one would not have been able to uncover the same changes in the macroeconomic consequences of oil supply shocks by imposing a one-time break in the oil-macro relationship.

Alternative Identifying Assumptions.—Our second methodological contribution was the identification of oil supply shocks based on sign restrictions. An alternative assumption would have been to impose a vertical short-run oil supply curve that implies three exclusion restrictions on the first row of the structural impact multiplier matrix. This corresponds to a simplified version of the monthly oil market VAR model proposed by Kilian (2009a). Although these identifying restrictions were not intended for quarterly data, it is useful to compare the results to our baseline model. Unlike Kilian (2009a), we implement this procedure allowing for time variation in the parameters. Figure 5, panel B presents the time-varying median responses of the real price of oil and of the macroeconomic aggregates to an oil supply shock identified by exclusion restrictions to a 1 percent oil production shortfall, together with the sixteenth and eighty-fourth percentiles. Although there is some evidence of time variation in the response of the real price of oil, the posterior intervals are so wide to leave open the possibility that the responses remained unchanged. In contrast, the time-varying estimates in our baseline model cannot be explained by estimation uncertainty only. Moreover, a puzzling finding in Figure 5, panel B is that the median real GDP response to a negative oil supply shock is positive, especially toward the end of the sample. This counterintuitive result suggests that the recursive identification scheme is inappropriate at the quarterly frequency and highlights the potential benefits of using sign restrictions.
Figure 5. Posterior Median Responses after an Oil Supply Shock Normalized to a 1 Percent Decrease in Oil Production, Where Shaded Areas Indicate 68 percent Posterior Credible Sets

Other Model Specifications.  We also investigated the sensitivity of our results to changes in the variables included in the benchmark model. When US unemployment instead of real GDP is used as the indicator of real economic activity, we find a substantial rise in unemployment following a negative oil supply shock again normalized on oil production. The strength of this response increases notably in the most recent past. When we replace the consumer price index with the implicit GDP deflator as a measure of inflation, we observe a somewhat more subdued increase in the price level after a negative oil supply shock that corresponds to a 1 percent oil production shortfall, but a similar pattern of time variation emerges. Likewise, using different oil price measures, such as the real refiners’ acquisition cost of composite crude oil or the West Texas Intermediate spot oil price, does not affect our conclusions. Finally, augmenting our model by the federal funds rate, as is common in the literature on monetary policy, does not change our findings about the dynamic response of the US economy to oil supply shocks and the structural change in the crude oil market.

Timing of the Sign Restrictions.  Since in the early part of our sample the nominal oil price was constrained by long-term bilateral agreements that were subject to revision only periodically, an obvious concern arises with regard to the timing of the sign restrictions imposed in the baseline model. When nominal oil prices are sluggish, a positive demand shock in the US economy, which raises world oil production and the consumer price level, causes a fall in the real price of oil, unless nominal contracts are renegotiated timely to reflect the new macroeconomic conditions. The resulting opposite movement in world oil production and in the real price of oil would imply that this demand shock is erroneously identified as a positive oil supply shock. One way of addressing this problem is to impose that the sign restrictions are only binding from the fourth quarter after the shock onward, such that the immediate reaction of oil production and of the real price of oil is unconstrained. It can be shown that our findings are not sensitive to this change in the timing of the identification restrictions. For further details the reader is referred to section 2.2 of the online Appendix.

III. Reasons for the Steepening of the Oil Demand Curve

In this section, we consider developments in the economy and in the crude oil market that might explain the substantial reduction of the short-run price elasticity of oil demand we documented earlier. These developments include a decline in the oil intensity of US real output, fuel switching, shifts in the composition of total oil demand, and variation in the global spare capacity of oil production. We do not preclude that other factors contributed to the steepening of the oil demand curve.

Energy Efficiency and Sectoral Shifts.  Following the oil price surges of the 1970s, the extent to which the US economy and other industrialized countries rely on oil has changed substantially since the mid-1980s. Industries gradually switched away from oil to alternative sources of energy, developed more energy-efficient technologies and improved energy conservation. These efforts were supported by
government policies aimed at reducing oil usage and increasing energy awareness. This transition together with service-biased growth in the United States caused oil intensity in aggregate economic activity to fall steadily, reflected in a reduction in the use of oil input per unit of output as shown in Figure 6, panel A. This development helps account for the decline of the short-run price elasticity of oil demand (see, e.g., Ryan and Plourde 2002; Hughes, Knittel, and Sperling 2008; Dargay and Gately 2010).\footnote{Dargay and Gately (1994) attribute this phenomenon to the irreversibility of technical knowledge, the durability of efficiency-improving investments and the nonabrogation of laws regarding energy-cost labeling and energy-efficiency standards.}

In addition, the composition of total oil demand has changed over time with oil consumption now being concentrated in sectors such as transportation, which traditionally were characterized by a low own-price elasticity of demand due to the lack of substitutes for transportation fuels. The increased share of transportation in total oil demand has also contributed to a steepening of the oil demand curve (see Dargay and Gately 2010; Ramey and Vine 2011). Finally, industrialized countries have increasingly outsourced their industrial production to emerging economies. Emerging economies typically rely heavily on oil as an input factor. Their oil demand tends to be less sensitive to changes in oil prices than advanced economies. In addition, it has been suggested that governments in developing economies have used fuel subsidies to shield consumers from the impact of rising global oil prices thereby stimulating oil consumption and making consumer demand unresponsive to international price signals.\footnote{"Both wholesale and retail prices of oil products in the domestic market are lower than they are in the global market" as exemplified for China by Hang and Tu (2007, 2978). An estimate by Morgan Stanley shows that almost a quarter of the world’s petrol is sold at less than the market price (The Economist 2008).} Given the larger share of emerging economies in global oil demand over time, this observation is consistent with the view that the elasticity of oil demand in global markets has declined.

\textit{Capacity Constraints in Crude Oil Production}.—Yet another possible explanation is that in the presence of capacity constraints in crude oil production, the composition of oil demand is likely to change in favor of less elastic speculative or precautionary buying. At times of low spare capacity, even small supply disruptions can lead to large price increases because market participants anticipate that an unexpected loss in oil production cannot be replaced by other oil producers. In that sense, increasing capacity utilization signals some tightness in the market, which increases the willingness of oil consumers to pay a higher price for a barrel of oil at the margin that provides insurance against potential scarcity, i.e., they pay an insurance premium (see Alquist and Kilian 2010).\footnote{As has been noted by Gately (1984, 1103), "aggravating the market tightness was an extended period of aggressive stockbuilding by the importing countries for much of 1979 and 1980. Such a stockbuilding "scramble" during a disruption was certainly perverse. It undoubtedly drove the price higher than it would have gone otherwise." This aggressive hoarding behavior could hint at the increased importance of less elastic precautionary demand in total oil demand in a tightening market. In fact, Adelman (2002, 179) states that "when buyers fear damage from sudden dearth, there is also a precautionary motive; which may be joined to a speculative motive, to profit by buying sooner."} Kilian (2008b) has documented that annual global capacity utilization rates in crude oil production have been steadily increasing from the mid-1980s to the early 1990s and stayed at levels close to full capacity ever
since (see Figure 6, panel B). This observation is in line with the gradual decline in the short-run price elasticity of oil demand uncovered in our empirical analysis.

IV. Conclusions

In this paper, we have analyzed the time-varying effects of oil supply shocks on the US economy and the oil market from 1974 onward. There are several a priori reasons to expect this relationship to have evolved over time. For example, the transition from a regime of administered oil prices to a market-based system, changing capacity utilization rates in crude oil production, and changes in the energy dependence of the US economy all have implications for the effects of oil supply shocks on oil market variables and US macroeconomic aggregates.
Our analysis combines a novel identification strategy for oil supply shocks based on inequality constraints with the estimation of a time-varying parameter VAR model. The first generation of structural oil market models has relied on exclusion restrictions on the impact multiplier matrix that can be interpreted as a vertical short-run oil supply curve. Instead, we propose to identify oil supply shocks based on sign restrictions derived from a simple supply and demand model of the crude oil market. Specifically, we identify an oil supply shock as a disturbance in the global oil market that moves oil production and the real price of oil in opposite directions. This approach is particularly appealing when quarterly data are used, since traditional delay restrictions are credible only at the monthly frequency.

Until now, time variation in oil markets has been analyzed by splitting the sample or by estimating bivariate VARs on rolling windows. The split-sample approach is not appealing when dealing with smooth structural change. The rolling-window approach relies on short samples that would cause severe degrees-of-freedom problems when estimating the reduced form of our high-dimensional VAR model. An additional concern would be the ability of such a VAR model to identify the oil supply shocks on the basis of very short samples. We dealt with these two challenges by estimating a multivariate structural VAR with time-varying parameters and stochastic volatility in the spirit of Cogley and Sargent (2005), Primiceri (2005), and Benati and Mumtaz (2007), among others.

We showed that when an exogenous oil supply shock in the time-varying parameter model is normalized to hold constant across time the implied change in the real price of oil on impact, the response of real GDP declines over time, which is consistent with other recent evidence. When normalizing the shock to hold constant across time the implied change in global oil production on impact, however, the response of US output and inflation have become larger in magnitude in more recent years. We showed that these two findings can be reconciled by a combination of a decline in the short-run price elasticity of oil demand (such that a given shortfall in oil production is associated with a greater price response on impact) and of changes over time in the response functions of all model variables. In particular, we documented that the oil demand curve is currently much steeper relative to the 1970s and early 1980s.

We further showed that the share of the volatility of the real price of oil explained by oil supply shocks has moderately decreased over time, indicating that oil supply shocks are not the primary driver of oil price movements in more recent periods. The contribution of oil supply shocks to the variability of real activity and inflation is economically relevant, ranging fairly steadily between 15 and 20 percent with the exception of 2008. Our analysis also reaffirmed that even allowing for time variation, the Great Inflation of the 1970s and early 1980s cannot be accounted for by negative oil supply shocks. Nor have the recessionary effects of oil supply shocks been very pronounced. We found that oil supply disruptions mattered for real economic activity mainly during two episodes. They contributed to the 1991 recession and they slowed the ongoing boom at the end of the millennium. There is little or no evidence that oil supply shocks mattered much for the recessions of the early 1980s and for the downturns in 2001 and in 2008.

Our analysis adds to a growing literature on endogenous oil prices and their implications for the macroeconomy, including Baumeister and Peersman (forthcoming).
One key difference between our analysis in this paper and that in the other paper is that here we focused on studying and understanding the responses of US macroeconomic aggregates, whereas Baumeister and Peersman (forthcoming) focus on explaining oil price dynamics. A second key difference is that in this paper we were solely concerned with the effects of oil supply shocks without taking a stand on the difficult problem of disentangling different oil demand shocks. Indeed, trying to address both of these problems at the same time in a TVP-VAR framework would cause a degrees-of-freedom problem. Thus, the analysis in these two papers is complementary.

APPENDIX A: DATA DESCRIPTION

Monthly world oil production data measured in thousands of barrels of oil per day were obtained from the US Energy Information Administration’s (EIA) *Monthly Energy Review* starting in January 1973. Monthly data for global production of crude oil for the period 1953:4 to 1972:12 were taken from the weekly *Oil & Gas Journal* (issue of the first week of each month). For the period 1947:1 to 1953:3, monthly data were constructed by interpolation of yearly world oil production data by means of the Litterman (1983) methodology using US monthly oil production data from the EIA as an indicator variable. Annual oil production data were obtained from *World Petroleum* (1947–1954), the *Oil & Gas Journal* (end-of-year issues, 1954–1960), and the EIA’s *Annual Energy Review* (1960–2010). Consistency between these different data sources was checked at each of the overlapping periods. Quarterly data are averages of monthly observations.

The nominal US refiners’ acquisition cost of imported crude oil was taken from the *Monthly Energy Review*. Since this series is only available from January 1974 onward, it was backcast until 1947:1 with the quarterly growth rate of the producer price index (PPI) for crude oil retrieved from the Bureau of Labor Statistics (BLS) database (WPU0561). Data were converted to quarterly frequency before backcasting by averaging over months. For the robustness checks with regard to the choice of the oil price variable, we use the quarterly average of the West Texas Intermediate (WTI) spot oil price obtained from the Federal Reserve Economic Data (FRED) database maintained by the St. Louis FED (OILPRICE) and of the nominal US refiners’ acquisition cost of composite crude oil from the *Monthly Energy Review*. The latter was adjusted for price controls on domestic oil production for the period 1971:III to 1974:I as described in Mork (1989) and reconstructed backward to 1947:1 in the same way as the imported refiners’ acquisition cost series.

Quarterly seasonally adjusted series for US real GDP (GDPC96: real gross domestic product, billions of chained 2005 dollars) and for the US GDP deflator (GDPDEF: gross domestic product implicit price deflator) were obtained from

---

10 Since this part of the data is only needed for the training sample to calibrate the priors based on the estimation of a fixed-coefficient VAR, the use of interpolated data as opposed to actual ones is of minor consequence.

11 The refiners’ acquisition cost of crude oil imports (IRAC) is a volume-weighted average price of all kinds of crude oil imported into the US over a specified period. Since the US imports more types of crude oil than any other country, it may represent the best proxy for a true “world oil price” among all published crude oil prices. The IRAC is also similar to the OPEC basket price.

12 The entitlement system in force during the 1970s in the United States required buyers to purchase foreign and domestic oil in fixed proportions so that the aggregate price was a weighted average of these two kinds of oil.
the FRED database. Monthly seasonally adjusted data for US industrial production (INDPRO: industrial production index, index 2007 = 100), for the US consumer price index (CPIAUCSL: consumer price index for all urban consumers: all items, index 1982–1984 = 100), for the civilian unemployment rate, 16 years and older, seasonally adjusted (UNRATE), and for the effective federal funds rate (FEDFUNDS) were taken from the FRED database, where the latter three were converted to quarterly frequency by taking averages.

Data on annual energy consumption by sector measured in billions of British Thermal Units (BTU) were obtained from the Annual Energy Review (Tables 2.1c–2.1f) for the period 1974–2010. Primary and secondary uses of petroleum in the industrial, commercial and part of the transportation sector (buses and heavy trucks) were aggregated to obtain a measure of total petroleum consumption in the production process. Data on highway transportation energy consumption by mode (1970–2009) were taken from the Transportation Energy Data Book (Table 2.7).

Annual data on global spare capacity of oil production for the period 1974–2010 were taken from the IMF World Economic Outlook and updated with the EIA's Short-Term Energy Outlook. Spare capacity refers to production capacity that can be brought online within 30 days and sustained for 90 days. Global capacity utilization rates are calculated as a percentage of total potential annual world oil production, which is the sum of available spare capacity and actual oil production taken from the Annual Energy Review.

APPENDIX B: BAYESIAN ESTIMATION OF A VAR WITH TIME-VARYING PARAMETERS AND STOCHASTIC VOLATILITY

Consider the time-varying vector autoregression model with stochastic volatility described by equations (1) to (7) in the main text.

A. Prior Distributions and Starting Values

The priors for the initial states of the drifting coefficients, the covariances and the log volatilities, \( p(\theta_0), p(\alpha_0) \) and \( p(\ln h_0) \), respectively, are assumed to be normally distributed, independent of each other and independent of the hyperparameters, which are the elements of \( Q \), \( S \) and the \( \sigma^2_i \) for \( i = 1, \ldots, 4 \). The priors are calibrated on the point estimates of a constant-coefficient VAR (4) estimated over the period 1947:II–1972:II.

The unrestricted prior for the VAR coefficients is set to

\[
\theta_0 \sim N[\hat{\theta}_{OLS}, \hat{P}_{OLS}],
\]

where \( \hat{\theta}_{OLS} \) corresponds to the OLS point estimates of the training sample and \( \hat{P}_{OLS} \) to four times the covariance matrix \( \hat{V}(\hat{\theta}_{OLS}) \).

With regard to the prior specification of \( \alpha_0 \) and \( h_0 \), we follow Primiceri (2005) and Benati and Mumtaz (2007). Let \( P = AD^{1/2} \) be the Choleski factor of the time-invariant variance covariance matrix \( \Sigma_{OLS} \) of the reduced-form innovations from the estimation of the fixed-coefficient VAR (4), where \( A \) is a lower triangular
matrix with ones on the diagonal, and $D^{1/2}$ denotes a diagonal matrix whose elements are the standard deviations of the residuals. The prior for the log volatilities is as follows:

$$\ln h_0 \sim N(\ln \mu_0, 10 \times I_4),$$

where $\mu_0$ is a vector that contains the diagonal elements of $D^{1/2}$ squared, and the variance covariance matrix is arbitrarily set to ten times the identity matrix to make the prior only weakly informative. The prior for the contemporaneous interrelations is set to

$$\alpha_0 \sim N[\tilde{\alpha}_0, \tilde{V}(\tilde{\alpha}_0)],$$

where the prior mean for $\alpha_0$ is obtained by taking the inverse of $A$ and stacking the elements below the diagonal row by row in a vector in the following way:

$$\tilde{\alpha}_0 = [\tilde{\alpha}_{0,21}, \tilde{\alpha}_{0,31}, \tilde{\alpha}_{0,32}, \tilde{\alpha}_{0,41}, \tilde{\alpha}_{0,42}, \tilde{\alpha}_{0,43}]'.$$

The covariance matrix, $\tilde{V}(\tilde{\alpha}_0)$, is assumed to be diagonal, and each diagonal element is set to ten times the absolute value of the corresponding element in $\tilde{\alpha}_0$. While this scaling is obviously arbitrary, it accounts for the relative magnitude of the elements in $\tilde{\alpha}_0$ as pointed out by Benati and Mumtaz (2007).

With regard to the hyperparameters, we make the following assumptions along the lines of Benati and Mumtaz (2007). We postulate that $Q$ follows an inverse-Wishart distribution:

$$Q \sim IW(Q^{-1}, T_0),$$

where $T_0$ is the prior degrees of freedom, which is set equal to the length of the training sample, which is sufficiently long (25 years of quarterly data) to guarantee a proper prior. Following Primiceri (2005), we adopt a relatively conservative prior for the time variation in the parameters in setting the scale matrix to $Q = (0.01)^2 \cdot \hat{V}(\theta_{OLS})$ multiplied by the prior degrees of freedom. This is a weakly informative prior, and the particular choice for its starting value is not expected to influence the results substantially since the prior is dominated by the sample information as time progresses. We experimented with different initial conditions inducing a different amount of time variation in the coefficients to test whether our results were sensitive to the choice of the prior specification. We follow Cogley and Sargent (2005) in setting the prior degrees of freedom alternatively to the minimum value allowed for the prior to be proper, $T_0 = \dim(\theta_i) + 1$, together with a smaller value of the scale matrix, $Q = 3.5e^{-4} \cdot \hat{V}(\theta_{OLS})$, which together put little weight on our prior belief about the drift in $\theta_i$. Our results are not materially affected by the different choices for this prior.

The three blocks of $S$ are assumed to follow inverse-Wishart distributions, with the prior degrees of freedom set equal to the minimum value required for the prior to be proper:

$$S_i \sim IW(S_i^{-1}, i + 1),$$
where \(i=1,2,3\) indexes the blocks of \(S\). The scale matrices are calibrated on the absolute values of the respective elements in \(\tilde{\alpha}_0\), as in Benati and Mumtaz (2007). Specifically, \(\tilde{S}_i\) is a diagonal matrix with the relevant elements of \(\tilde{\alpha}_0\) multiplied by \(10^{-3}\).

Given the univariate feature of the laws of motion of the stochastic volatilities, the variances of the innovations to the univariate stochastic volatility equations are drawn from an inverse-Gamma distribution as in Cogley and Sargent (2005):

\[
\sigma_i^2 \sim IG\left(\frac{10^{-4}}{2}, \frac{1}{2}\right).
\]

This distribution is proper and has fat tails.

B. Markov Chain Monte Carlo Algorithm for Simulating the Posterior Distribution

Since sampling from the joint posterior is complicated, we simulate the posterior distribution by sequentially drawing from the conditional posterior of the four blocks of parameters: the coefficients \(\theta^T\); the simultaneous relations \(A^T\); the variances \(H^T\), where the superscript \(T\) refers to the whole sample; and the hyperparameters—the elements of \(Q\), \(S\) and the \(\sigma_i^2\) for \(i = 1,\ldots, 4\)—collectively referred to as \(M\). Posteriors for each block of the Gibbs sampler are conditional on the observed data \(Y^T\) and the rest of the parameters drawn at previous steps.

**Step 1: Drawing coefficient states**

Conditional on \(A^T\), \(H^T\), \(M\), and \(Y^T\), the measurement equation is linear and has Gaussian innovations with known variance. Therefore, the conditional posterior is a product of Gaussian densities, and \(\theta^T\) can be drawn using a standard simulation smoother (see Carter and Kohn 1994; Cogley and Sargent 2002), which produces a trajectory of parameters:

\[
p(\theta^T|Y^T, A^T, H^T) = p(\theta_T|Y^T, A^T, H^T) \prod_{t=1}^{T-1} p(\theta_t|\theta_{t+1}, Y^T, A^T, H^T).
\]

From the terminal state of the forward Kalman filter, the backward recursions produce the required smoothed draws that take the information of the whole sample into account. More specifically, the last iteration of the filter provides the conditional mean \(\theta_{T|T}\) and conditional variance \(P_{T|T}\) of the posterior distribution. A draw from this distribution provides the input for the backward recursion at \(T-1\), \(T-2\) and so on until the beginning of the sample according to

\[
\theta_{t|t+1} = \theta_{t|t} + P_{t|t} P_{t+1|t}^{-1}(\theta_{t+1} - \theta_t)
\]

\[
P_{t|t+1} = P_{t|t} - P_{t|t} P_{t+1|t}^{-1} P_{t|t}.
\]

**Step 2: Drawing covariance states**

Similarly, the posterior of \(A^T\) conditional on \(\theta^T\), \(H^T\), and \(Y^T\) is a product of normal densities and can be calculated by applying the same algorithm as in step 1 as a consequence of the block diagonal structure of the variance covariance matrix \(S\). More specifically, a system of unrelated regressions based on the relation
\( \mathbf{A}, \mathbf{u}_t = \varepsilon_t \), where \( \varepsilon_t \) are orthogonalized innovations with known time-varying variance \( \mathbf{H}_t \) and \( \mathbf{u}_t = \mathbf{y}_t - \mathbf{X}_t\theta_t \) are observable residuals, can be estimated to recover \( \mathbf{A}_T \) according to the following transformed equations, where the residuals are independent standard normal:

\[
\begin{align*}
(17) & \quad u_{1,t} = \varepsilon_{1,t} \\
(18) & \quad (h_{2,t}^{-\frac{1}{2}}u_{2,t}) = -\alpha_{2,1}(h_{2,t}^{-\frac{1}{2}}u_{1,t}) + (h_{2,t}^{-\frac{1}{2}}\varepsilon_{2,t}) \\
(19) & \quad (h_{3,t}^{-\frac{1}{2}}u_{3,t}) = -\alpha_{3,1}(h_{3,t}^{-\frac{1}{2}}u_{1,t}) - \alpha_{3,2}(h_{3,t}^{-\frac{1}{2}}u_{2,t}) + (h_{3,t}^{-\frac{1}{2}}\varepsilon_{3,t}) \\
(20) & \quad (h_{4,t}^{-\frac{1}{2}}u_{4,t}) = -\alpha_{4,1}(h_{4,t}^{-\frac{1}{2}}u_{1,t}) - \alpha_{4,2}(h_{4,t}^{-\frac{1}{2}}u_{2,t}) - \alpha_{4,3}(h_{4,t}^{-\frac{1}{2}}u_{3,t}) + (h_{4,t}^{-\frac{1}{2}}\varepsilon_{4,t}).
\end{align*}
\]

**Step 3: Drawing volatility states**

Conditional on \( \theta_T, \mathbf{A}_T, \) and \( \mathbf{Y}_T \), the orthogonalized innovations \( \varepsilon_t = \mathbf{A}_t(\mathbf{y}_t - \mathbf{X}_t\theta_t) \) with \( \text{Var}(\varepsilon_t) = \mathbf{H}_t \) are observable. However, drawing from the conditional posterior of \( \mathbf{H}_T \) is more involved because the conditional state-space representation for \( \ln h_{i,t} \) is not Gaussian. The log-normal prior on the volatility parameters is common in the stochastic volatility literature, but such a prior is not conjugate. Following Cogley and Sargent (2005, Appendix B.2.5) and Benati and Mumtaz (2007), we apply the univariate algorithm by Jacquier, Polson, and Rossi (1994) that draws the volatility states \( h_{i,t} \) one at a time.

**Step 4: Drawing hyperparameters**

The hyperparameters \( \mathbf{M} \) of the model can be drawn directly from their respective posterior distributions since the disturbance terms of the transition equations are observable given \( \theta_T, \mathbf{A}_T, \mathbf{H}_T, \) and \( \mathbf{Y}_T \).

We perform 100,000 iterations of the Gibbs sampler and discard the first 50,000 draws as “burn-in.” The remaining sequence of draws from the conditional posteriors of the four blocks form a sample from the joint posterior distribution \( p(\theta_T, \mathbf{A}_T, \mathbf{H}_T, \mathbf{M} | \mathbf{Y}_T) \). We keep only every tenth draw in order to mitigate the autocorrelation among the draws. Following Primiceri (2005) and Benati and Mumtaz (2007), we ascertain that our Markov chain has converged to the ergodic distribution by computing the draws’ inefficiency factors, which are the inverse of the relative numerical efficiency measure (RNE) introduced by Geweke (1992),

\[
RNE = (2\pi)^{-1} \frac{1}{S(0)} \int_{-\pi}^{\pi} S(\omega) d\omega,
\]

where \( S(\omega) \) is the spectral density of the retained draws from the Gibbs sampling replications for each set of parameters at frequency \( \omega \).\(^{13}\) Figure 1A in the online

\(^{13}\) See Benati and Mumtaz (2007) for details on the implementation.
Appendix displays the inefficiency factors for the states and the hyperparameters of the model, which are all far below the value of 20 designated as an upper bound by Primiceri (2005). Thus, the autocorrelation across draws is modest for all elements, which provides evidence of convergence to the ergodic distribution. In total, we collect 5,000 simulated values from the Gibbs chain on which we base our structural analysis.

**APPENDIX C: GENERALIZED IMPULSE RESPONSES AND SIGN RESTRICTIONS**

Here we describe the Monte Carlo integration procedure that we use to compute the nonlinear impulse response functions to a structural oil supply shock. In the spirit of Koop, Pesaran, and Potter (1996), we compute the generalized impulse responses as the difference between the conditional expectations with and without the exogenous shock:

\[
IRF_{t+k} = E[y_{t+k} | \varepsilon_t, \omega_t] - E[y_{t+k} | \omega_t],
\]

where \(y_{t+k}\) contains the forecasts of the endogenous variables at horizon \(k\); \(\omega_t\) represents the current information set that captures the entire history up to that point in time, and \(\varepsilon_t\) is the current disturbance term. At each point in time, the information set upon which we condition the forecasts contains the actual values of the lagged endogenous variables and a random draw of the model parameters and hyperparameters. More specifically, in order to calculate the conditional expectations, we simulate the model in the following way. We randomly draw from the Gibbs sampler output one possible state of the economy at time \(t\) represented by the time-varying lagged coefficients and the elements of the variance covariance matrix. Starting from this random draw from the joint posterior that includes the hyperparameters, we employ the transition laws and stochastically simulate the future paths of the coefficient vector and the components of the variance covariance matrix for up to 20 quarters into the future. By projecting the evolution of the system in this way, we account for all the potential sources of uncertainty deriving from the additive innovations, variations in the lagged coefficients and changes in the contemporaneous relations among the variables in the system.

To obtain the time-varying structural impact matrix \(B_{0,t}\), we implement the procedure proposed by Rubio-Ramírez, Waggoner, and Zha (2010). Given the current state of the economy, let \(\Omega_t = P_t D_t P_t'\) be the eigenvalue-eigenvector decomposition of the VAR’s time-varying variance covariance matrix \(\Omega_t\) at time \(t\). We draw an \(N \times N\) matrix, \(K\), from the \(N(0, 1)\) distribution, take the QR decomposition of \(K\), where \(R\) is a diagonal matrix whose elements are normalized to be positive, and \(Q\) is a rotation matrix the columns of which are orthogonal to each other, and compute the time-varying structural impact matrix as \(B_{0,t} = P_t D_t R Q'\). Given this contemporaneous impact matrix, we compute the reduced-form innovations based on the relationship \(u_t = B_{0,t} \varepsilon_t\), where \(\varepsilon_t\) contains four structural shocks drawn from a standard normal distribution. Impulse responses are then computed by comparing the effects of a shock on the evolution of the endogenous variables to the benchmark case without a shock, where in the former case the shock is set to \(\varepsilon_{i,t} + 1\),
while in the latter we only consider $\varepsilon_{i,t}$. The reason for this is to allow the system to be impacted by other disturbances during the propagation of the shock of interest. From the set of impulse responses derived in this way, we select only those impulse responses that at horizons $t + k, k = 0, 1, \ldots, 4$, satisfy the sign restrictions, i.e., the responses of the endogenous variables are consistent with the structural shock we wish to identify; all others are discarded.

We repeat this procedure until 100 iterations have fulfilled the sign restrictions for a given simulated future path of the economy, and then calculate the mean responses of the four endogenous variables over the accepted rotations. For each point in time, we randomly draw 500 current states of the economy that provide the distribution of impulse responses taking into account possible developments of the structure of the economy. The representative impulse response function for each variable at each date is the median of this distribution.

REFERENCES


