**Article:**
Shu, Haicheng and Spencer, Peter orcid.org/0000-0002-5595-5360 (2023) Oil Prices in the Real Economy. Journal of Applied Econometrics. pp. 878-897. ISSN 0883-7252

https://doi.org/10.1002/jae.2986

# Oil prices in the real economy*

Haicheng Shu† Peter Spencer‡

March 18, 2023

**Abstract**

This paper presents a macro-finance model of the US economy and the spot and futures markets for oil. The performance of the model is greatly enhanced by using the Kalman filter to model latent variables representing the inflation asymptote, the real price of oil and the slope of the futures curve. We find that these are dominated by innovations in observed futures prices, reflecting the importance of market expectations. Using the Kalman filter to capture inflationary shocks helps solve the notorious price puzzle, the tendency for increases in interest rates to anticipate such developments and apparently cause inflation. Futures prices also depend upon risk premiums, which we find are dominated by the latent variable representing the real oil price rather than macro variables like inflation and interest rates.

**Keywords**: affine term structure model, macro finance model, oil price, oil futures contracts, spanned macro factor risk.

**JEL code**: G12, G13, Q41, Q43

## 1 Introduction

This paper develops an affine macro-finance model of the macroeconomy and the oil futures market, which we use to study the links between the US macro economy, monetary policy, and the spot and futures markets for oil.

*We are grateful to Carol Alexander, Pierre Collin-Dufresne, Laura Coroneo, Jean-Sebastiane Fontaine, Adam Golinski, Leonardo Iania, Anh Le, Peter Smith, Mike Wickens and three referees of this journal for helpful comments.

†Corresponding author, School of Economics, Southwestern University of Finance and Economics, Chengdu, China. *haicheng.shu@swufe.edu.cn*

‡Peter Spencer is Emeritus Professor of Economics and Finance in the Department of Economics and Related Studies, University of York, YO105DD, UK. *peter.spencer@york.ac.uk*

As the name suggests, the macro-finance approach allows oil futures prices to reflect macroeconomic variables as well as the latent variables normally used to model oil market factors in futures markets. It is based on the 'central bank model' (CBM) developed by Svensson (1999), Rudebusch (2002), Smets (2002), Kozicki and Tinsley (2005) and others, which represents the behavior of the macroeconomy in terms of the output gap ($g_t$), inflation ($\pi_t$), and the short term interest rate ($r_t$). We follow Dewachter and Lyrio (2006), Ireland (2007) and others and allow for shifts in the long run inflation asymptote ($\pi_t^*$), using the Kalman filter to represent this empirically.

The oil price also affects this model of the economy. This side of the model is informed by the extensive macroeconometric literature, pioneered by Hamilton (1983), which studies the effect of oil prices on the economy. Hamilton argues that most post-war recessions in the US have been caused by oil price shocks. He and his colleagues have followed this paper up with many studies documenting the adverse effect of oil price changes on real output and inflation over the last few decades (Hamilton (1985), Hamilton (2003), Herrera and Hamilton (2001), Hamilton (2008), Hamilton and Wu (2014)). Many other authors have studied this effect (Raymond and Rich 1997, Finn 2000, Hooker 1996 and Hooker 2002). However, Kilian (2008), (2009) and Barsky and Kilian (2004) attribute a much greater role to demand side pressures, arguing that the effect of supply-side shocks depends critically upon the tightness of the oil market in the run up to the shock.

Whereas these authors focus on the effect of the spot oil price on the economy, the novelty of this paper lies in our exploitation of the information to be found in oil futures prices. These authors assume that the spot price is observed without error and use it as a regressor in the macro model, after adjusting it for inflation to model the real price. However, as Dunn and Holloway (2012) note, because there is very little physical trade in oil, this assumption is unrealistic. In practice, the spot price reported by Price Reporting Agencies like Platts is assessed in terms of near-term forward and futures prices. Our model allows the reported spot and futures prices to depend upon a latent variable ($\rho$) representing the 'true' real oil price. The futures curve, which shows how prices vary with maturity, also depends upon a second latent variable ($\delta$) that determines its slope.

We use the Kalman filter to represent these two latent variables (together with $\pi_t^*$); assuming that spot and futures prices are all measured with error. The filter updates these variables optimally in terms of surprises in the observed macroeconomic variables and the reported spot and futures prices. The reported price ($\rho^o$) is then set equal to $\rho$ plus a measurement error. If, instead, we suppress these errors and mimic the conventional macro model approach by setting $\rho$ equal to the reported price, there is a very large deterioration in the likelihood, suggesting that this approach is inappropriate.

Futures prices play an important role in this model because they reflect market expectations. Specifically, the futures price for any future date represents the risk-neutral expectation of the price at that date. The market's expectation equals the futures price plus a risk premium. Our model allows these three items to be distinguished because it captures the time series dynamics of the data. This model can then be used to calculate the time series expectation of the price at any future date, which is used as a proxy for the market expectation. Subtracting the futures price from this expectation gives an estimate of the risk premium, the expected dollar payoff to an investor holding the future. Central banks routinely use this method to decompose a government bond yield into an expected interest rate and a risk premium.[1]

Although oil futures prices have only become available since 1984, the ability of the Kalman filter to deal with missing data allows us to combine these data with a larger macro dataset (1964Q1 to 2022Q1) and study the period of the oil price shocks of the 1970s. We evaluate two types of model. In the first type, macro variables do not have a contemporaneous effect on futures prices, which, as in previous studies, only depend upon latent variables. However, in this type of model, macro variables affect the time series dynamics and thus the future evolution of prices. In the second type, they also have a contemporaneous effect on futures prices. Our model selection criteria leads us to adopt the second approach.

Our empirical results shed new light on the nature of the risk premium. The conventional view is that oil producers sell futures to hedge against future price falls, depressing the futures price and increasing the risk premium until arbitrageurs are prepared take the other side of the market. Hamilton and Wu (2014) argue that since 2004, buying pressure from commodity index funds has had the opposite effect, turning the premium negative in 2008 and 2014. We checked this by following them in allowing for a structural break in 2005Q1. This exercise suggested that the risk-neutral dynamics did indeed shift then, possibly because of increased participation in the futures market by financial institutions. It also revealed an increase in macroeconomic volatility, consistent with the ending of the Great Moderation. Nevertheless, our results suggest that it was the weakness of oil prices that caused the premium to turn negative in 2008 and 2014. We also checked to see if the exploitation of shale hydrocarbons or other developments had shifted the time-series dynamics, but found no significant evidence of this. Thus, our preferred model (reported as M5 in Section 3) allows for a shift in volatility and the risk-neutral (but not the time- series) dynamics in 2005Q1.

[1]See for example the daily analysis published by the Federal Reserve at: https://www.federalreserve.gov/pubs/feds/2005/200533/200533abs.html.

Our use of the Kalman filter to capture the effect of potential inflationary developments helps to solve the notorious price puzzle - the tendency (noted originally by Sims 1992) for increases in policy interest rates to anticipate such developments and thus apparently cause inflation. We find that the latent variables are almost entirely driven by surprises in futures prices once these become available in 1984.[2] Introducing these variables into the macro model allows them to pick up inflationary surprises that could otherwise be wrongly attributed to innovations in the interest rate.

Thus, we find that the properties of our preferred model are nicely in line with economic priors. An increase in the oil price has the effect of increasing inflation and reducing economic activity initially. Our model suggests that the Federal Reserve responds to this by reducing interest rates, meaning that the reduction in activity is short-lived. It raises interest rates in response to increases in economic activity and inflation, consistent with the Taylor rule. In turn, an increase in interest rates has the effect of slowing activity, which reduces inflation. Our model suggests that US interest rates have surprisingly strong effects on the oil market, reinforcing this disinflationary effect. The results also align with those of Barsky and Kilian (2004) in suggesting that the persistence of oil price shocks depends upon the strength of the US economy.

The paper is organized as follows. The next section sets out the macro-finance dynamic term structure model framework, which specifies the state dynamics under the real-world and risk-neutral probability measures. Section 3 sets out the empirical methodology and econometric model, describes the data that we use, and discusses the empirical findings and results. Section 4 concludes.

# 2 The model framework

Our research strategy is to follow the macro-finance literature, which models the dynamics under the risk-neutral and the state space measures using both latent variables and observable variables like the interest rate, assuming that these are measured without error. As in the existing literature on the oil futures market, we represent the oil price and convenience yield by latent variables, but re-specify the model in terms of the real oil price to align it with a conventional macroeconometric

[2] The Kalman gain matrix for any model and any period shows how surprises in the observed variables are mapped into revisions in the latent variables, reflecting how informative they are. These matrices are not reported, but are available from the authors upon request. We find that the observed (real) spot price ($\rho^o$) has practically no effect on $\rho$ or other latent variables after 1984, when futures were first trade. This suggests that measurement error makes it inappropriate to use this as a regressor in a macro model. Instead, we would suggest using futures prices as instrumental variables for the observed spot price.

structure. We first describe three possible ways of modeling the oil market and then set out the model of the macroeconomy, which is common to all our models..

## 2.1 The spot oil market

The relationship between the spot ($S_t$) and one-period future ($F_{1,t}$) oil price depends upon the cost of carrying inventory $c_t = (r_t - d_t)$, which can be decomposed into the convenience yield from holding physical oil inventories $d_t$ and the spot interest rate $r_t$:

$$F_{1,t} = S_t e^{c_t} = S_t e^{(r_t - d_t)} \tag{1}$$

This shows that if the cost of carry is negative the futures curve is downward sloping. The market is then said to be "in backwardation", as it was in 2012 for example. If the cost of carry is positive, the forward curve is upward sloping and is said to be "in contango", as it was in 2015.

This relationship naturally leads to a two-factor model of the futures curve, in which the level of the curve is dictated by the spot price and its slope by the cost of carry. Many papers, like Heath (2019), treat the cost of carry as a single variable, while others like Cassasus and Dufresne (2005), distinguish the convenience yield and the spot interest rate. The theory of storage (see Working 1933, Kaldor 1939, Working 1949, Brennan 1958, Weymar 1968) suggests that the convenience yield is closely related to the level of the commodity stored in the inventory. This literature states that when inventories are tight, the convenience yield will be high, the cost of carry negative, and the futures curve in backwardation. On the other hand, when oil inventories are abundant, as they were in 2015, the convenience yield will be negative, adding to the interest cost of carry and pushing the futures curve into contango. Similarly, economic theory suggests that the interest rate depends upon inflation, the output gap, and other factors influencing monetary policy.

## 2.2 The risk-neutral dynamics

Crucially, a futures contract, unlike a spot inventory, does not yield convenience or other benefits, which means that the price $F_{n,t}$ of a contract to deliver oil at any future date $(t + n)$ equals the risk-neutral expectation of the spot oil price $S_{t+n}$ at that date, as explained in online appendix (1). Assuming that prices are lognormally distributed, the risk-neutral dynamics are determined by an identity that rules out arbitrage:

$$s_{t+1} = s_t + c_t - \frac{1}{2}\sigma_s^2 + \epsilon^Q_{s,t+1} \qquad \epsilon^Q_{s,t} \sim N(0, \sigma_s^2) \tag{2}$$

where: $s_t$ is the natural logarithm spot oil price at time $t$ and $\sigma_s^2$ its one period ahead conditional variance. In this paper, parameter values and error terms with a $Q$ superscript, as in the equations of this section denote values under the risk-neutral measure $Q$, while those without superscripts as in the equations of the next section) denote values under the time-series or physical probability measure $P$. Heath (2018) develops a two-factor latent factor model that complements this equation with another for $c_t$ :

$$\begin{pmatrix} s_{t+1} \\ c_{t+1} \end{pmatrix} = \begin{pmatrix} k_s^Q \\ k_c^Q \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ \phi_{c,s}^Q & \phi_{c,c}^Q \end{pmatrix} \begin{pmatrix} s_t \\ c_t \end{pmatrix} + \begin{pmatrix} \epsilon_{s,t+1}^Q \\ \epsilon_{c,t+1}^Q \end{pmatrix} \tag{3}$$

where $k_s^Q = -\frac{1}{2}\sigma_s^2$ and $\phi_{c,s}^Q, \phi_{c,c}^Q$ are parameters to be estimated. This is used to find the risk-neutral expectations for future spot prices and hence the futures prices.

This offers a simple stand-alone model of the futures market, similar to Schwartz (1997) and Cassasus and Collin-Dufresne (2005), in which the futures prices only depend upon the latent variables. The simplicity of this type of specification stems from the ability of a few latent variables to explain a very high percentage of the variance of the cross-section of nominal futures prices.

Subtracting the log GDP deflator $p_{t+1} = p_t + \pi_{t+1}$, where $\pi_{t+1} = p_{t+1} - p_t$ is the rate of inflation, from both sides of equation (2) gives a similar real arbitrage identity. This replaces $s_{t+1}$ by the *real* oil price $\rho_{t+1} = s_t - \pi_{t+1}$ and $c_t$ by the real cost of carry $\chi_t = c_t - E_t^Q(\pi_{t+1})$. This change of variable gives a stand-alone two factor model of real futures prices:

$$\begin{pmatrix} \rho_{t+1} \\ \chi_{t+1} \end{pmatrix} = \begin{pmatrix} k_\rho^Q \\ k_c^Q \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ \phi_{c,\rho}^Q & \phi_{c,c}^Q \end{pmatrix} \begin{pmatrix} \rho_t \\ \chi_t \end{pmatrix} + \begin{pmatrix} \epsilon_{\rho,t+1}^Q \\ \epsilon_{\chi,t+1}^Q \end{pmatrix}' \tag{4}$$

We use this in preference to equation (3) because it explains real rather than nominal oil prices and thus aligns better with the models used by macroeconomists. This is used to fit the futures prices in our benchmark model M1. Because macro variables do not appear in this system, they have no contemporaneous effect on oil prices in M1.

A further modification is to split the real cost of carry $\chi_t = r_t - d_t - E_t^Q(\pi_{t+1})$ into the explicitly-observed interest rate $r_t$ and a latent variable $\delta_t$ representing the residual cost of carry (inflation and the convenience yield, which both reduce the real cost of carry): $\delta_t = d_t + E_t^Q(\pi_{t+1})$. The risk-neutral dynamics can then be specified as:

$$\begin{pmatrix} \rho_{t+1} \\ \delta_{t+1} \end{pmatrix} = \begin{pmatrix} k_\rho^Q \\ k_\delta^Q \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ \phi_{\delta,\rho}^Q & \phi_{\delta,\delta}^Q \end{pmatrix} \begin{pmatrix} \rho_t \\ \delta_t \end{pmatrix} + \begin{pmatrix} r_t \\ 0 \end{pmatrix} + \begin{pmatrix} \epsilon_{\rho,t+1}^Q \\ \epsilon_{\delta,t+1}^Q \end{pmatrix} \tag{5}$$

Notably, although this specification still only involves two latent factors, it is fundamentally different from model M1(4) because it allows the macro variables to have a contemporaneous effect on the cross-section of futures prices via the interest rate term. This is how we specify the futures prices in model M2.

Model M2 requires more parameters than model M1, because equation (5) has to be supported by a macro model determining the interest rate under the risk-neutral measure, which is set out in the next section. It is assumed that the algebraic structure of this model is similar to that of the risk-neutral structure, although we allow their parameter values and error terms to differ and relax the arbitrage restriction. Differences in the estimates of the two structures determine the risk premia, as explained in online appendix (4).

Importantly, equation (5) has the same structure as equation (4), except that the interest rate appears in the arbitrage equation alongside the latent variable with a weight of unity. Both models can thus be encompassed by one in which the interest rate appears with a freely estimated coefficient alongside a latent variable that takes the place of $\delta_t$ in M2. We call this encompassing specification model M3.

## 2.3 The time series dynamics

The dynamics of the latent variables and hence futures prices over time are determined by a time series model. Conventionally, this is specified as a VAR for the nominal spot price and cost of carry that is similar in structure to the risk-neutral specification in (3). Heath (2019) develops a model that combines this nominal risk-neutral structure with a time series VAR, which also contains measures of the growth rate and oil inventory. These are called “unspanned macro risk factors” because they do not have a contemporaneous effect on the cross section, but do affect its future evolution. Similarly, the real oil price and cost of carry can affect the macro variables in the VAR.

We modify this approach in several ways. Because macroeconometric models are more naturally specified using the real rather than the nominal oil price, we first use (4) rather than (3) to model the risk-neutral dynamics. Second, to model the time series dynamics, we follow macro-finance studies of the term structure of interest rates (such as Dewachter and Lyrio 2006) in using the Central Bank Model, which sees the spot interest rate $r_t$ being determined by monetary policy, jointly with the output gap $g_{t-1}$ and inflation $\pi_{t-1}$ (and in our case the real oil price). Estimating this model jointly with equation (4) allows these macro variables to act as “unspanned factors”. We call this model M1.

Finally, to get model M2, we use (5) rather than (4) to model the risk-neutral dynamics. Recall that this means the macro variables have a contemporaneous effect on the futures prices. This is called a “spanned macro factor risk” model because the futures prices are determined (i.e. 'spanned') by both macro and latent variables.

Our macroeconomic structure follows the approach of recent macroeconometric research, such as Ireland (2007) and others, in allowing for shifts in the Fed's implicit inflation target by modelling this as a non-stationary latent variable that acts as the inflation asymptote $\pi^*_{t-1}$:

$$\pi^*_t = \kappa_{\pi^*} + \pi^*_{t-1} + \epsilon_{\pi^*,t} \tag{6}$$

The model design ensures that in the absence of shocks, the inflation rate equals this asymptote in steady state. The latent variables are collected in the vector: $z_t = (\pi^*_t, \delta_t, \rho_t)'$.[3] Allowing for the lagged effect of the macro variables $m_t = (g_t, \pi_t, r_t)$ under this measure, the dynamics describing the oil market potentially generalizes to:

$$\delta_t = k_\delta + \theta_{\delta,\pi^*}\pi^*_t + \phi_{\delta,\rho}\rho_{t-1} + \phi_{\delta,\delta}\delta_{t-1} + \phi_{\delta,m}m_{t-1} + \epsilon_{\delta,t} \tag{7}$$

$$\rho_t = k_\rho + \theta_{\rho,\pi^*}\pi^*_t + \phi_{\rho,\rho}\rho_{t-1} + \phi_{\rho,\delta}\delta_{t-1} + \phi_{\rho,m}m_{t-1} + \epsilon_{\rho,t} \tag{8}$$

Stacking equation (6) to (8) gives:

$$z_t = K_z + \Theta_{z,z}z_t + \Phi_{z,z}z_{t-1} + \Phi_{z,m}m_{t-1} + L_z D_z \epsilon_{z,t} \qquad \epsilon_{z,t} \sim N(0, I_z) \tag{9}$$

The full specification of this equation system is presented in online appendix (2).

## 2.4 The observable state variables

The oil market is only assumed to affect the macroeconomy through the real spot oil price. The macro variables depend upon each other (with a lag) as well as the lagged real oil price $\rho_{t-1}$ and the asymptote $\pi^*_t$:

$$g_t = \kappa_g + \theta_{g,\pi^*}\pi^*_t + \phi_{g,\rho}\rho_{t-1} + \phi_{g,g}g_{t-1} + \phi_{g,\pi}\pi_{t-1} + \phi_{g,r}r_{t-1} + \epsilon_{g,t} \tag{10}$$

$$\pi_t = \kappa_\pi + \theta_{\pi,\pi^*}\pi^*_t + \phi_{\pi,\rho}\rho_{t-1} + \phi_{\pi,g}g_{t-1} + \phi_{\pi,\pi}\pi_{t-1} + \phi_{\pi,r}r_{t-1} + \epsilon_{\pi,t} \tag{11}$$

$$r_t = \kappa_r + \theta_{r,\pi^*}\pi^*_t + \phi_{r,\rho}\rho_{t-1} + \phi_{r,g}g_{t-1} + \phi_{r,\pi}\pi_{t-1} + \phi_{r,r}r_{t-1} + \epsilon_{r,t} \tag{12}$$

Stacking equation (10) to (12) gives:

$$m_t = K_m + \Theta_{m,z}z_t + \Phi_{m,z}z_{t-1} + \Phi_{m,m}m_{t-1} + L_m D_m \epsilon_{m,t} \qquad \epsilon_{m,t} \sim N(0, I_m). \tag{13}$$

The full specification of this equation is presented in online appendix (2).

[3] This section describes the time series dynamics in model M2. The time series model M1 is identical except that the latent variable $\delta_t$ is replaced by $\chi_t$, so that for example the latent vector in model M1 is: $z_t = (\pi^*_t, \chi_t, \rho_t)'$.

## 2.5 The time series dynamics

The state variables are contained in the state vector $X_t = (z_t, m_t)' = (\pi_t^*, \delta_t, \rho_t, g_t, \pi_t, r_t)'$, Stacking equations (9) and (13) gives the equation for this under the measure $P$:

$$\begin{pmatrix} z_t \\ m_t \end{pmatrix} = \begin{pmatrix} K_z \\ K_m \end{pmatrix} + \begin{pmatrix} \Theta_{z,z} & 0 \\ \Theta_{m,z} & 0 \end{pmatrix} \begin{pmatrix} z_t \\ m_t \end{pmatrix} + \begin{pmatrix} \Phi_{z,z} & \Phi_{z,m} \\ \Phi_{m,z} & \Phi_{m,m} \end{pmatrix} \begin{pmatrix} z_{t-1} \\ m_{t-1} \end{pmatrix} + \begin{pmatrix} L_z D_z & 0 \\ L_{z,m} & L_m D_m \end{pmatrix} \begin{pmatrix} u_{z,t} \\ u_{m,t} \end{pmatrix} \tag{14}$$

In matrix form, we have:

$$X_t = K + \Theta X_t + \Phi X_{t-1} + LD\epsilon_t \qquad \epsilon_t \sim N(0, I) \tag{15}$$

The full specification of this equation system is presented in online appendix (2).

## 2.6 The identification scheme

Although $\rho_t$ and $\delta_t$ are determined by the model of the futures prices, we need to restrict this system to align the inflation asymptote $\pi_t^*$ with the observed rate $\pi_t$ the steady state. To do this, note that $\Theta = [\theta \ 0_{6,5}]$ where $\theta = [0 \ \theta_{\delta,\pi^*} \ \theta_{\rho,\pi^*} \ \theta_{g,\pi^*} \ \theta_{\pi,\pi^*} \theta_{r,\pi^*}]'$ and write (15) as:

$$X_t = K + \theta\pi_t^* + \Phi X_{t-1} + \epsilon_t \tag{16}$$

Dropping time subscripts and error terms, this has the steady state:

$$X = \varphi + R\pi_t^* \tag{17}$$

where:

$$\varphi = (I - \Phi)^{-1} K, \qquad R = (I - \Phi)^{-1}\theta. \tag{18}$$

Thus, setting:

$$K = (I - \Phi)\varphi, \qquad \theta = (I - \Phi)R. \tag{19}$$

and $\varphi = [0 \ 0 \ \varphi_\rho \ 0 \ 0 \ \varphi_r \ \ ]'$ and $R = [1 \ 0 \ 0 \ 0 \ 1 \ 1]'$ ensures that: $\delta = g = 0$ and that $\rho = \varphi_\rho$ and $\pi = \pi^*$ in the steady state. The last of these restrictions also ensures that the interest rate asymptote also moves in line with the inflation asymptote $r = \varphi_r + \pi^*$.

## 2.7 The companion form

Rearranging equation (15) yields the companion form of the state equation under the measure $P$:

$$X_t = A + BX_{t-1} + CDu_t \qquad u_t \sim N(0, I). \tag{20}$$

$$= A + BX_{t-1} + W_t \qquad W_t \sim N(0, \Sigma). \tag{21}$$

where: $K$ and $\Theta$ are specified in equation (19). $A = (I-\Theta)^{-1}K$ and $B = (I-\Theta)^{-1}\Phi$, $C = (I - \Theta)^{-1}L$ and are further specified in section 2 of the online appendix.

## 2.8 The risk-neutral state dynamics in model M2

Recall that futures prices are modelled by equation (4) in model M1. In model M2 the oil market equation (5) is embedded in a model of the macroeconomy under measure $Q$, which generates the spot interest rate and is congruent with equation (20).

$$X_t = A^Q + B^Q X_{t-1} + W_t^Q \qquad W_t^Q \sim N(0, \Sigma)$$

where $A^Q$, $B^Q$ and $\Sigma$ are specified in section 2 of the online appendix.

## 2.9 The term structure of futures prices

The state dynamics under the risk-neutral measure $Q$ determine the cross-sectional loadings. We model real futures prices $h_{\tau,t} = f_{\tau,t} - p_t$, using the Affine Term Structure Model:

$$h_{\tau,t} = \alpha_\tau + \Psi_\tau X_t \tag{22}$$

The initial condition is implied by the special case when $\tau = 0$, in which $h_{0,t} = \rho_t$, giving the starting values for state variables as:

$$\psi_{\rho,0} = 1 \qquad \psi_{\pi^*,0} = \psi_{\delta,0} = \psi_{g,0} = \psi_{\pi,0} = \psi_{r,0} = 0 \tag{23}$$

Online appendix (1) shows $\alpha_\tau$ and $\Psi_\tau$ in equation (22) have the following recursive close-form solution:

$$\Psi_\tau = \Psi_{\tau-1} B^Q + B_\pi^Q, \tag{24}$$

$$\alpha_\tau = \alpha_{\tau-1} + \Psi_{\tau-1} A^Q + \frac{1}{2}\Psi_{\tau-1}\Sigma\Psi'_{\tau-1} \tag{25}$$

## 2.10 The measurement equation

The measurement equation in the state space representation represents the observed values $y_t = \begin{pmatrix} h_t^o & \rho_t^o & g_t^o & \pi_t^o & r_t^o \end{pmatrix}$, in terms of the state vector $X_t$ by:

$$y_t^o = J + HX_t + e_t \qquad e_t \sim N(0, Q) \tag{26}$$

where: $e_t = \begin{pmatrix} h_t & \rho_t & g_t & \pi_t & r_t \end{pmatrix}$. This assumes that, the commodity futures $h_t = (h_t, \ldots, h_\tau)'$ and spot price $\rho_t$ data are observed with error, and macro data $m_t = (g_t, \pi_t, r_t)'$ are observed without error. Thus, we define the measurement equation (26) as:

$$\begin{pmatrix} h_t^o \\ \rho_t^o \\ g_t^o \\ \pi_t^o \\ r_t^o \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \Psi_{\pi^*} & \Psi_\delta & \Psi_\rho & \Psi_g & \Psi_\pi & \Psi_r \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \pi_t^* \\ \delta_t \\ \rho_t \\ g_t \\ \pi_t \\ r_t \end{pmatrix} + e_t \qquad e_t \sim N(0, Q) \tag{27}$$

where $Q = diag(q_1^2, \ldots, q_\tau^2, q_\rho^2, 0, 0, 0)$. As noted in the introduction, the observed real oil price ($\rho_t^o$) depends upon the latent price ($\rho_t$) plus a measurement error.

## 2.11 The Kalman filter and the likelihood function

To complete the dynamic term structure model, we now outline the maximum likelihood approach used to estimate the Kalman filter and the model parameters. Recall that the filter uses surprises in forecasting the vector of observable variables to update the vector of latent variables. As section 5 of the online appendix explains, the Kalman gain, $\Gamma$, depends upon the matrix $\Theta_{m,z}$ in equation (14), which shows the contemporaneous effect of the factors on the observable variables. So, before 1984, when there are no futures prices, the revisions just depend upon revisions in the macro variables, but after 1984 they also depend upon surprises in the futures prices. We use the Kalman filter rather than the principal components used by many other term structure researchers primarily because it deals in this way with the absence of futures data over the pre-1984 period (and the increase in the number of maturities traded subsequently). Online appendix (5) also derives the log-likelihood function for this model.

# 3 The empirical model

To recap, the empirical model consists of a heteroscedastic VAR describing the three latent variables and three macroeconomic variables, as in equation (20), and the auxiliary equations describing the representative futures prices, as in equation (26). This is estimated by maximum likelihood and the Kalman filter, which gives optimal estimates of the latent variables in this situation. This section describes the data and the empirical results.

## 3.1 Data sources and description

The model is estimated using quarterly time series of the macro variables and crude oil futures. All data are downloaded from the Thompson Reuters Datastream. Summary statistics are presented in Table 1. Figure 1 shows the West Texas Intermediate (WTI) oil futures prices. We also use data for US output, US inflation, and US Federal Funds rate, from 1964Q1 to 2022Q1. This allows the effect of the oil shocks of the 1970s to be analyzed. The Fed Funds rate is specified as a quarterly decimal fraction (the annual rate in percent divided by 400). We generate the US output gap by applying the HP filter to log US GDP, then subtracting this measure of potential output from log GDP. US inflation is the annual log difference of the US implicit GDP price deflator.

The spot oil price is a composite series. The WTI spot price, which matches the futures data, is available from 1984Q1, while the Brent price, which gives the price of a similar grade, is available from 1970Q1[4]. Since the crude oil price was fixed close to \$2.25 per barrel between 1964Q1 and 1970Q1[5], this value is used until then; the Brent price from 1970Q1 to 1982Q4 and WTI thereafter. To represent the term structure of oil futures, the prices of WTI light crude oil futures traded on New York Mercantile Exchange are used, beginning in 1984, when these oil futures contracts started trading. Oil futures contracts with 1, 2, 3, 6, 9, 12, 18 and 24 month maturities are studied. The series for the prices of oil futures with 1, 2, 3, 6, 9 and 12 month maturities are available from 1984Q1, the 18 month contract from 1989Q3; and the 24 month contract from 1995Q1.

[4]Brent and WTI spot oil price series only diverge significantly in recent years, when the latter went to a discount because of export controls and the development of the US shale hydrocarbon industry.

[5]Before the 1970s, the oil market was monopolized by the major Western oil companies, and the oil price at that time was described by the phrase : "take the price used by Exxon, add it to that used by Shell and divide the sum by two" (Carollo 2012).

## 3.2 Unspanned versus spanned macro factors

Recall that models M1, M2, and M3 differ in the weight given to the explicitly observed interest rate in the arbitrage equation. This is zero in the benchmark model M1, unity in M2, and a freely estimated parameter in the encompassing model M3. The upper panel of Table 2 reports the likelihood statistics and the Bayesian Information Criteria (BIC) for these models. The likelihood of M2 is practically identical to that of M3, indicating that this is an acceptable simplification. The estimate of the rate parameter in M3 is 0.9933, with a standard error of 0.0063. The value of zero in M1 is decisively rejected: the M3 parameter estimate is 158.2 standard deviations away from zero. Tests based on log likelihood ratios and t-statistics with classical statistical considerations are known for their tendency to over-reject such restrictions in large data samples. The BIC criterion allows for this by applying a penalty to the large number of parameters in model M2, but still shows that this model is preferred to M1.

## 3.3 Structural change

Macroeconomic volatility has clearly increased since the global financial crisis of 2008 that marked the end of the Great Moderation in volatility seen since the mid-1980s. Ironically, many of the factors that helped to explain this earlier moderation, such as deregulated financial markets and the move to just-in-time inventory control, seem to have contributed to the subsequent increase in volatility. This increase could also reflect an increase in the size and frequency of macroeconomic shocks.

There has also been a marked increase in the volatility of oil prices over this period, partly reflecting the effect of shocks such as the Covid pandemic. We would expect this to shift the volatility parameters in $\Sigma$. Hamilton and Wu (2014) and others have suggested that this may also reflect the financialization of commodity markets (which we call the financialization hypothesis). This could have affected the risk-neutral parameters in $A^Q$ and $B^Q$. On the other hand, the development of the US shale hydrocarbon sector, which now acts as a swing producer, may have reduced the exposure of the US economy to oil prices. These effects would be most likely to affect the time series parameters in $A$ and $B$.

In order to test these various hypotheses, we conducted a series of structural stability tests on model M2. The results are shown in the lower panel of Table 2. We followed Hamilton and Wu (2014) and split the sample into two periods: 1964Q1-2004Q4 and 2005Q1-2022Q1. As a portmanteau test, we first used different sets of perimeter values to optimize the likelihood in each period, thus allowing all of the model parameters to change. Combining the results for these two sub-periods gives the criteria for model M4 shown in the Table 2. M4 naturally has a higher likelihood than M2, but has twice the number of parameters, and taking this into account, the BIC criterion suggests that model M2 is preferable to M4.

Nevertheless, analysis of the differences between the parameter values for the two periods indicated that there was very little change in the time series parameters, but that the volatility and risk-neutral parameters had indeed shifted. Table 2 presents the model selection statistics for two more models that examine these differences. The first, model M5, allows the volatility ($\Sigma$) and risk-neutral ($A^Q$ and $B^Q$) parameters to shift between the two periods, while using a single set of time series parameters ($A$ and $B$) to optimize the likelihood of the full sample. M5 has a higher BIC value than both M2 and M4. Finally, to try to distinguish the effects of volatility and financialization, model M6 allows only the risk-neutral parameters to shift. In terms of the BIC statistic, this model is preferred to M2 and M4 but not M5. Our preferred model is thus M5. This is consistent with both the increase in macro volatility and the financialization hypothesis, but does not suggest that the exploitation of shale hydrocarbons had a significant effect on the structure of the real economy.

## 3.4 The empirical results

The rest of the paper compares the empirical properties of models M1 (with unspanned macro factor risk) and M5 (spanned) in some detail. Table 3 shows the root mean squared error and prediction errors of the futures prices in these models. Table 5 reports the M5 parameters, while Table 1 in the online appendix reports those for M1. Results for the other models are available upon request from the authors.

## 3.5 The factor loadings

The difference between M1 and M5 is most evident in the behavior of the futures market, which is indicated by the factor loadings (Section 2.9). Empirically, both models have a single unit root under $Q$ that is associated with the asymptotic inflation rate and means that the loadings of the futures on the factors ($\Psi_\tau$) increase with maturity ($\tau$). Dividing these loadings by maturity gives the factor loadings for the annualized cost of carry ($\Psi_\tau/\tau$). These loadings are depicted in Figure 2, as a function of maturity (expressed in quarters). The first panel shows the loadings on the three latent variables and the second those on the observed variables.

Futures prices respond in a similar way to the latent variables, but the effect of the observed variables is absent from M1. This is the key difference between the spanned and unspanned factor risk models. The effect of oil price shocks decays slowly with maturity in both models, while the effect of shocks on the cost of carry in M1 and the convenience yield in M5 peaks in the 1-2 year area, before decaying. As we would expect, in M5 the interest rate and inflation shocks have the effect of pushing up futures prices, an effect that mirrors that of shocks to $\delta_t$. This means that portfolio managers could in principle use oil futures to help hedge interest rate and inflation shocks as well as shocks to the oil market.

### 3.6 The behavior of the macro and spot oil market variables

Figures 3 show the estimates of the latent and observed state variables. In the latter case, these estimates are shown alongside their observed values. The long term inflation asymptote ($\pi^*$) in Figure 4 captures the secular trends in inflation. This resembles the inflation target identified by Ireland (2007) and shown in his Figure 4. It accommodated the inflationary oil price hikes in the 1970s, resulting in the peak inflation rate of 10%. However, it fell back after the Volker deflation in the early 1980s, when we saw the peak interest rate of 15.9%.

Figure 5 shows how the oil market and the output gap interact in model M5. The vertical bars show the NBER recession periods. The output gap was very high before both of the oil shocks of the 1970s, as indicated by the left-hand sides of the first two bars, reflecting the strength of the US economy. This helped tighten the oil market. Oil inventory and cost of carry were very low in the run up to the first oil shock, but moved to a higher level afterwards, as shown in Figure 6. Reflecting this tightness, as Kilian (2008) has argued, these oil price increases were persistent, provoking a sharp fall in the output gap as the economy moved into recession. In contrast, the economy was not as strong prior to the oil price spike seen at the time of the first Gulf war in 1991, when the oil price displayed a spike rather than a step increase, which was followed by a mild recession. The US economy was also strong when the oil price peaked in 2008, but the ensuing recession was due to the financial crisis as well as the high oil price, which fell back sharply as the recession took hold. These shocks are also reflected in $\delta_t$, estimated using the Kalman filter. This estimate reflects the tightening of the market in the 1970s, as well as the weakness seen in 2014 and 2020.

Figure 6 also shows the relationships between the US oil inventory (excluding the US Strategic Petroleum Reserve)[6] and the estimated cost of carry from M1 and convenience yield ($d_t = \delta_t - \pi_t$) from M5. Although the inventory is not part of the model, this panel shows that short-run swings in $d_t$ and the oil inventory are inversely related, as the theory of storage would predict. However, there are some notable spikes in $d_t$ that are not reflected in the inventory. For example, there is a sharp spike in 1974Q2, which arguably reflects rationing and other effects designed

[6]Source: U.S. Energy Information Administration at: https://www.eia.gov/

to conserve oil stocks and help shield the economy from the Arab oil embargo.[7]

The lower panel of Figure 6 shows a scatter plot of our convenience yield against the log inventory series. The correlation coefficient for these series is 0.58. We take the log of the inventory series because in principle Pindyck (1994), and Jin (2019) suggest that the relationship between $d_t$ and inventory holdings should be non-linear: to prevent stock-outs, the yield should become extremely large as the inventory approaches zero. However, inventories are maintained at a high level in our sample so the risk of stockout is small.

Table 4 reports the estimates of the parameters obtained from the Kalman-VAR for M5 under the measure $P$. (See Table 1 in the online appendix for M1.) The estimates of the key parameters conform to economic priors and in the main are statistically significant. As we would expect, the estimates of $\phi_{\pi,\rho}$ and $\phi_{g,\rho}$ indicate that the real oil price has a significant short run impact on inflation (and in M5) activity. Consistent with the Taylor rule, which suggests that the central bank adjusts the policy interest rate in order to maintain a stable rate of inflation, increases in inflation and activity lead to increases in interest rates ( $\phi_{r,\pi}$ and $\phi_{r,g}$). In turn, an increase in interest rates has the effect of slowing activity ($\phi_{g,r}$), which reduces inflation ($\phi_{\pi,g}$). These two effects are very significant in M5. This model suggests that an increase in US interest rates also tends to reduce the oil price (presumably reflecting their global significance), reinforcing this disinflationary effect. The longer-run effects, revealed by the impulse response functions (IRFs) reported in the next section, which include the indirect effect of interest rates working through the US economy, are very significant.

### 3.6.1 Impulse response functions

Figures 7 and 8 present the IRFs for the state variables, which show the dynamic effects of innovations in the macroeconomic variables, together with 95% confidence bounds.[8] (See Figures 1 and 2 in the online appendix for M1.) Because these innovations are correlated empirically, orthogonalized innovations using the triangular factorization defined in section 2.7 are applied here. The orthogonalized impulse responses show the effect on the macroeconomic system of increasing each of these innovations by one percentage point for just one-period using the Wald representation of the system. Each column shows the effect of a unit shock upon a macro

[7]For example, the US Congress passed the Emergency Highway Energy Conservation Act to impose a national maximum speed limit of 55 mph in 1974, with similar restrictions imposed in European countries. In the UK, petrol coupons were issued in preparation for petrol rationing, although this was not actually implemented.

[8]The confidence bound for the IRFs in this paper uses the methodology reported in section 3 in Lütkepohl (2000) and Appendix B in Coroneo (2016). Note that because the IRFs (and the ANOVA results below) only depend upon the time series parameters, which do not shift in model M5, these are the same for both pre-and Post 2004 parameterisation

variable, while the rows show their effects.

The relationships between the output gap, inflation, and interest rates are in line with economic priors and similar to those seen in previous macro-finance models. The final row of Figure 8 shows that the US Fed changes interest rates in response to inflation and economic activity, consistent with the Taylor rule. (See Figure 2 in the online appendix for M1.) The final column shows that output and inflation in turn fall in response to the higher interest rate. Thus, the inclusion of the oil price within the CBM and use of Kalman filters to pick up the effect of unobservable expectational influences help to solve the notorious price puzzle - the tendency (noted originally by Sims 1992) for increases in policy interest rates to anticipate inflationary developments, and thus apparent to cause inflation. The oil markets appear to be surprisingly interest sensitive. A rise in interest rates gives an incentive to reduce the inventory. This reduces the oil price ($\phi_{\rho,r}$) and increases $\delta_t$ ($\phi_{\delta,\rho}$). The US interest rate is relevant because oil is priced in US dollars and this influences interest rates globally.

The top two panels on the right-hand side show that interest rates give the Federal Reserve a surprisingly sharp and strong leverage over the oil market. A one point increase in interest rates reduces the real oil price by 9.65% after one year in the model M5.

### 3.6.2 Analysis of variance

The real-world dynamics are also reflected in Figure 9. (See Figure 3 in the online appendix for M1.) These report the results of the Analysis of Variance exercise and show the share of the total variance attributable to the innovations at different lag lengths. These are also obtained using the Wald representation of the system, as described in Cochrane and Piazzesi (2009). They indicate the contribution each innovation would make to the volatility of each model variable if the error process was suddenly started, having been dormant previously. As such, they reflect both the impulse responses and the variance of the shocks.

The first column of Figure 9 shows the effects of oil market shocks in M5, while the second shows the effects of macro shocks. (See Figure 3 in the online appendix for M1.) The first two rows show that the variances of $\rho_t$ and $\delta_t$ are dominated by oil market shocks. However, macro shocks account for approximately 15% of the variance in $\delta_t$ and 30% of the variance in the oil price after 20 quarters. The remaining rows show oil market shocks affect the variance of inflation, accounting for 10% of the variance initially, then decaying gradually. The effects of oil shocks on the variance of output and interest rates are much smaller. The longer run variances of inflation and interest rates are naturally dominated by shocks to the inflation asymptote.

### 3.7 The risk premium

Econometric models frequently produce oil price forecasts that differ significantly from the comparable futures price. These differences are usually interpreted as risk premiums.[9] They can be expressed either in dollars or as percentage returns. The first is the difference between the time series expectation of a futures price in the next period $E(F_{\tau-1,t+1})$ and the current price of the future $F_{\tau,t}$, modelled by the risk-neutral expectation (see online appendix, section 4). This is the measure used by Baumeister and Kilian (2016), and gives the expected profit in dollars for holding a contract of this maturity for one quarter. The first panel of Figure 10 shows our estimates of the dollar premium in 3, 12, and 24 month contracts. The second panel shows this as an annualized percentage return, following Hamilton and Wu (2014) and Heath (2019), for example.

Producers want to sell futures to hedge against future price falls. When they are dominant, as they have been historically, this depresses the futures price relative to the expected spot rate until arbitrageurs are prepared take the other side of the market, so the premium is normally positive. However, Hamilton and Wu (2014) argue that buying pressure from commodity index funds has recently had the opposite effect, turning the premium negative in 2008 and 2014. They employ a similar methodology to ours, using two latent variables (level and slope factors) to fit the futures and model the premium. However, we also introduce macroeconomic variables into the model and can therefore offer a macroeconomic explanation.

The remaining two panels of Figure 10 decompose the premium into the effect upon the oil $(\delta_t, \rho_t)$ and macro $(\pi_t^*, g_t, \pi_t, r_t)$ factors. This shows that oil market variables play an important role here, with the oil price having a strong positive effect on the premium. The correlation between the 12-month dollar premium and the oil price is 0.80 pre-2005, increasing to 0.87% afterwards. This correlation is similar for the other maturities. If oil producers sell futures to lock in high profits on future production when oil prices are high, that could explain this correlation. That would push up the premium by pushing down futures prices. The increase post-2004 could be due to the effect of financialization, which made the futures market more convenient for hedging and speculation. We also find that the risk premium has more than doubled since the Great Moderation, in line with the findings of Hamilton and Wu (2014). The increased volatility of the oil market factors explains most of this effect. However, in contrast to Hamilton and Wu (2014), our model suggests

[9] However, Leduc et al (2021) suggest that these differences could also be due to mistakes in the market's ability to distinguish between transitory and permanent shocks when setting futures prices.

that it was the low level of oil prices in 2008 and 2014, rather than financialization, that caused the premium to turn negative in those years.

# 4 Conclusion

This paper presents a macro-finance model that includes the oil price and its cost of carry and makes crude oil futures exponential-affine in the state variables. As expected, we find significant links between oil prices and the macroeconomy. The model also throws light on the notorious 'price puzzle', indicating the importance of modelling the links between US monetary policy, commodity prices, and inflation on a comprehensive basis, using latent variables to capture the effect of inflationary developments reflected in futures prices. The macro-finance framework would seem to offer practitioners and academic researchers an important tool for understanding the effects of monetary policy on the commodity markets, and the economy.

# References


Barsky, R. B., & Kilian, L. (2004). Oil and the macroeconomy since the 1970s. *The Journal of Economic Perspectives*, *18*(4), 115–134.

Baumeister, C., & Kilian, L. (2016). A general approach to recovering market expectations from futures prices with an application to crude oil. *CESIFO Working Paper*(NO. 5782).

Brennan, M. J. (1958). The supply of storage. *The American Economic Review*, *48*(06), 50-72.

Carollo, S. (2011). *Evolution of the price of crude oil from the 1960s up to 1999.* John Wiley & Sons, Ltd.

Casassus, J., & Collin-Dufresne, P. (2005). Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance*, *60*(5), 2283–2331.

Cochrane, J. H., & Piazzesi, M. (2009). Decomposing the yield curve. *SSRN Working Paper Series*.

Dewachter, H., & Lyrio, M. (2006). Macro factors and the term structure of interest rates. *Journal of Money, Credit and Banking*, 119–140.

Dunn, S., & Holloway, J. (2012). The pricing of crude oil. *RBA Bulletin (Print copy discontinued)*, 65-74.

Finn, M. G. (2000). Perfect competition and the effects of price increases on economic activity. *Journal of Money, Credit and Banking*, 400–416.

Hamilton, J. D. (1983). Oil and the macroeconomy since world war ii. *The Journal of Political Economy*, 228–248.

Hamilton, J. D. (1985). Historical causes of postwar oil shocks and recessions. *The Energy Journal*, *6*(1), 97–116.

Hamilton, J. D. (2003). What is an oil shock? *Journal of Econometrics*, *113*(2), 363–398.

Hamilton, J. D. (2008). *Understanding crude oil prices* (Tech. Rep.). National Bureau of Economic Research.

Hamilton, J. D., & Wu, J. C. (2014). Risk premia in crude oil futures prices. *Journal of International Money and Finance*, *42*, 9–37.

Herrera, A. M., & Hamilton, J. D. (2001). Oil shocks and aggregate macroeconomic behavior: the role of monetary policy. *Department of Economics, UCSD*.

Hooker, M. A. (1996). What happened to the oil price-macroeconomy relationship? *Journal of Monetary Economics*, *38*(2), 195–213.

Hooker, M. A. (2002). Are oil shocks inflationary? asymmetric and nonlinear specifications versus changes in regime. *Journal of Money, Credit, and Banking*, *34*(2), 540–561.

Ireland, P. N. (2007). Changes in the federal reserve's inflation target: causes and consequences. *Journal of Money, Credit and Banking*, *39*(8), 1851–1882.

Jin, X. (2019). The role of market expectations in commodity price dynamics: evidence from oil data. *Journal of International Money and Finance*, *90*, 1-18.

Kaldor, N. (1939). Speculation and economic stability. *The Review of Economic Studies*, *7*(1), 1–27.

Kilian, L. (2008). Exogenous oil supply shocks: how big are they and how much do they matter for the us economy? *The Review of Economics and Statistics*, *90*(2), 216–240.

Kilian, L. (2009). Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. *The American Economic Review*, *99*(3), 1053–1069.

Kozicki, S., & Tinsley, P. A. (2005). Permanent and transitory policy shocks in an empirical macro model with asymmetric information. *Journal of Economic Dynamics and Control*, *29*(11), 1985–2015.

Leduc, S., Moran, K., & Vigfusson, R. J. (2021, 07). Learning in the oil futures markets: evidence and macroeconomic implications. *The Review of Economics and Statistics*, 1-45.

Lütkepohl, H. (2000). Bootstrapping impulse responses in var analyses. In *Compstat* (pp. 109–119).

Pindyck, R. S. (1994, Spring). Inventories and the short-run dynamics of commodity prices. *RAND Journal of Economics*, *25*(1), 141-159.

Raymond, J. E., & Rich, R. W. (1997). Oil and the macroeconomy: A markov state-switching approach. *Journal of Money, Credit, and Banking*, 193–213.

Rudebusch, G. D. (2002). Assessing nominal income rules for monetary policy with model and data uncertainty. *The Economic Journal*, *112*(479), 402–432.

Schwartz, E. S. (1997). The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance*, *52*(3), 923-73.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review*, *36*(5), 975–1000.

Smets, F. (2002). Output gap uncertainty: does it matter for the taylor rule? *Empirical Economics*, *27*(1), 113–129.

Svensson, L. E. (1999). Inflation targeting as a monetary policy rule. *Journal of Monetary Economics*, *43*(3), 607–654.

Weymar, F. H. (1968). The dynamics of the world cocoa market. *The dynamics of the world cocoa market.*.

Working, H. (1933). Price relations between july and september wheat futures at chicago since 1885. *Wheat Studies*, *9*(06).

Working, H. (1949). The theory of price of storage. *The American Economic Review*, *39*(06).

# Figures and tables

Table 1: Summary Statistics

| WTI futures prices | | | | | | |
|---|---|---|---|---|---|---|
| $h_\tau$ | Mean | S.D | Skew. | Kurt. | ADF (p-value) | Obs |
| $h_3$ | -0.857 | 0.490 | 0.276 | -0.982 | 0.209 | 153 |
| $h_6$ | -0.862 | 0.491 | 0.318 | -1.046 | 0.266 | 153 |
| $h_9$ | -0.869 | 0.492 | 0.344 | -1.087 | 0.305 | 153 |
| $h_{12}$ | -0.875 | 0.493 | 0.362 | -1.114 | 0.334 | 153 |
| $h_{18}$ | -0.842 | 0.505 | 0.258 | -1.286 | 0.225 | 131 |
| $h_{24}$ | -0.752 | 0.507 | -0.096 | -1.280 | 0.113 | 107 |
| Observable macro variables | | | | | | |
| $n^o$ | Mean | S.D | Skew. | Kurt. | ADF (p-value) | Obs |
| $\rho_t^o$ | -1.032 | 0.652 | -0.337 | -0.714 | 0.073 | 233 |
| $g_t^o$ | 0.000 | 0.016 | -1.135 | 5.014 | 0.001 | 233 |
| $\pi_t^o$ | 0.008 | 0.005 | 1.297 | 1.028 | 0.673 | 233 |
| $r_t^o$ | 0.012 | 0.009 | 0.778 | 0.714 | 0.224 | 233 |

These data were downloaded by Datasteam and are discussed in the text. $h_\tau$ denotes the $\tau$-month maturity log futures price less the log GDP deflator as in equation (20), where $h_3$, $h_6$, $h_9$, and $h_{12}$ are available from from 1984Q1 to 2022Q1; $h_{18}$ from 1989Q3 to 2022Q1; $h_{24}$ from 1995Q1 to 2022Q1. Observable macro variables (as in the vector $n^o$), namely $\rho^o$, $g^o$, $\pi^o$, and $r^o$ are available from 1964Q1 to 2022Q1. Mean denotes the sample arithmetic mean, S.D. standard deviation; Skew. and Kurt. report skewness and excess kurtosis, standard measures of the third and fourth moments. Obs. reports the number of observations. ADF shows the Augmented Dickey Fuller test statistic under the null hypothesis of non-stationarity.

Table 2: Model Selection

| Models | k | log likelihood | BICs |
|---|---|---|---|
| Basic models: unspanned v.s. spanned macro factors | | | |
| M1 | 46 | 7448.14 | -14551.95 |
| M2 | 65 | 7663.10 | -14839.64 |
| M3 | 66 | 7663.19 | -14832.34 |
| Extended models: structural stability tests | | | |
| M4 | 128 | 7895.39 | -14832.64 |
| M5 | 108 | 7840.46 | -14872.49 |
| M6 | 93 | 7780.76 | -14865.37 |

This table shows the model selection criteria for the various specifications shown in sections 3.2 and 3.3, $k$ is the number of parameters. The number of observations $n$ is 1782 for each model.

Table 3: Root Mean Squared Errors (RMSE) and Root Mean Squared Prediction Errors (RMSPE)

| $n_t^o$ | M1 | M5 | $h_{\tau,t}$ | M1 | M5 |
|---|---|---|---|---|---|
| | | Root Mean Squared Errors | | | |
| $\rho_t^o$ | 0.040 | 0.031 | 3m | 0.023 | $2.5\times10^{-13}$ |
| $g_t^o$ | $1.4\times10^{-15}$ | $1.1\times10^{-14}$ | 6m | 0.007 | 0.003 |
| $\pi_t^o$ | $1.2\times10^{-16}$ | $5.0\times10^{-16}$ | 9m | $1.1\times10^{-13}$ | $2.1\times10^{-13}$ |
| $r_t^o$ | $1.0\times10^{-16}$ | $4.1\times10^{-16}$ | 12m | 0.003 | 0.002 |
| | | | 18m | $1.1\times10^{-13}$ | $1.9\times10^{-13}$ |
| | | | 24m | 0.006 | 0.004 |
| | | Root Mean Squared Prediction Errors | | | |
| $\rho_t^o$ | 0.179 | 0.180 | 3m | 0.177 | 0.178 |
| $g_t^o$ | 0.010 | 0.010 | 6m | 0.157 | 0.156 |
| $\pi_t^o$ | 0.001 | 0.001 | 9m | 0.143 | 0.142 |
| $r_t^o$ | 0.002 | 0.002 | 12m | 0.133 | 0.132 |
| | | | 18m | 0.114 | 0.113 |
| | | | 24m | 0.109 | 0.108 |

The left-hand panel of this table shows the RMSEs and RMSPEs for the observed macro variables for the period 1964-2022. The right-hand panel shows these for the futures price (less the log GDP deflator as in Equation 22) for the periods that they are available (see notes to Figure 1).

Table 4: Parameter Estimates for M5

| Parameters | Estimates | t-stats | Parameters | Estimates | t-stats |
|---|---|---|---|---|---|
| | | Under the Physical Dynamics | | | |
| $\varphi_\rho$ | 0.040 | 0.377 | $\phi_{g,g}$ | 0.825 | 23.359 |
| $\varphi_r$ | -0.006 | -2.189 | $\phi_{g,\pi}$ | 0.088 | 1.879 |
| $\phi_{\delta,\delta}$ | 0.819 | 16.960 | $\phi_{g,r}$ | -0.127 | -1.968 |
| $\phi_{\delta,\rho}$ | -0.013 | -5.297 | $\phi_{\pi,\rho}$ | $5.7\times10^{-4}$ | 4.897 |
| $\phi_{\delta,r}$ | $-1.5\times10^{-39}$ | $-1.1\times10^{-38}$ | $\phi_{\pi,g}$ | 0.019 | 4.301 |
| $\phi_{\rho,\delta}$ | -0.367 | -0.953 | $\phi_{\pi,\pi}$ | 0.920 | 46.198 |
| $\phi_{\rho,\rho}$ | 0.889 | 18.354 | $\phi_{\pi,r}$ | $-5.8\times10^{-28}$ | $-1.0\times10^{-25}$ |
| $\phi_{\rho,g}$ | 0.402 | 0.504 | $\phi_{r,\rho}$ | $-8.2\times10^{-4}$ | -3.741 |
| $\phi_{\rho,\pi}$ | 9.104 | 1.587 | $\phi_{r,g}$ | 0.009 | 1.309 |
| $\phi_{\rho,r}$ | -4.340 | -2.329 | $\phi_{r,\pi}$ | 0.099 | 3.324 |
| $\phi_{g,\rho}$ | -0.001 | -1.317 | $\phi_{r,r}$ | 0.931 | 45.933 |

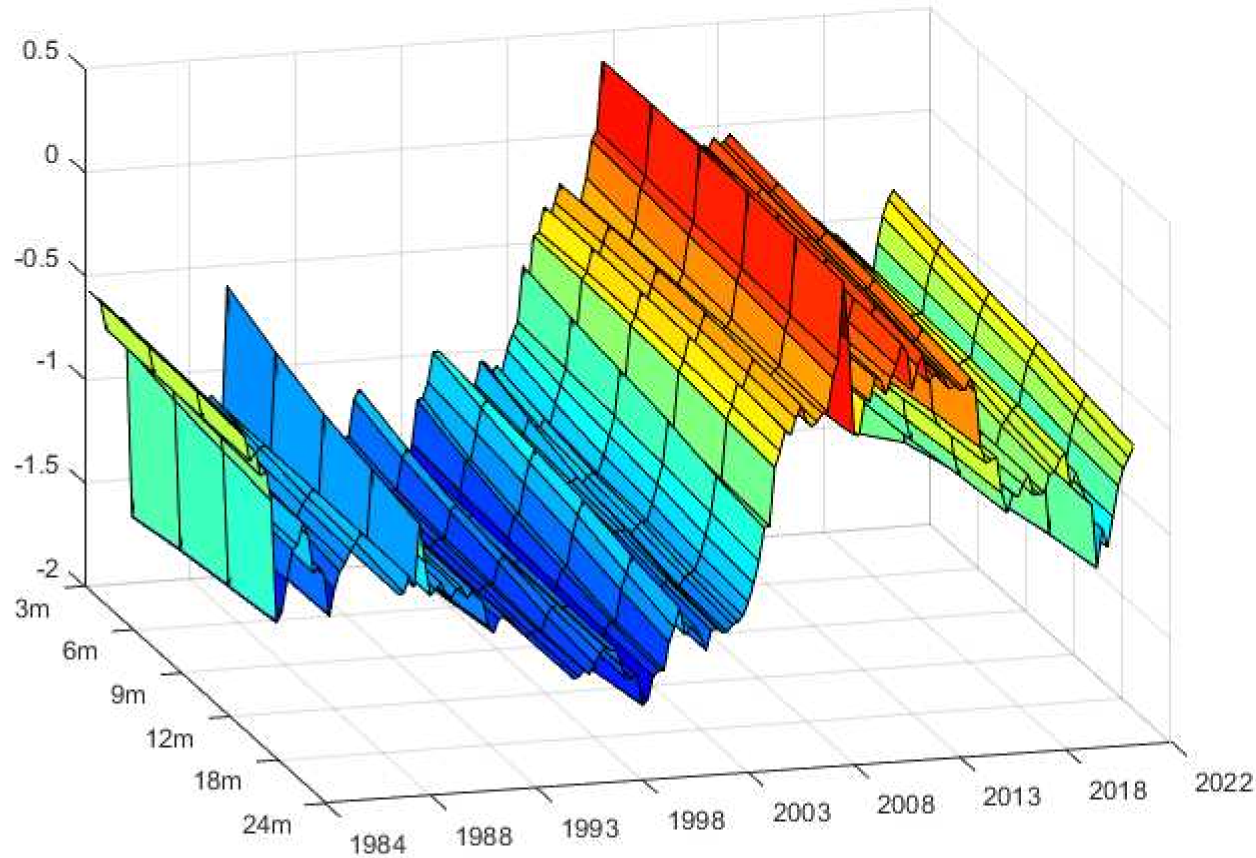
This table presents parameter estimates and their t-statistics for model M5.

Table 5: Parameter Estimates for M5 (cont)

| Parameters | Estimates | t-stats | Parameters | Estimates | t-stats |
|---|---|---|---|---|---|
| Under the Risk-neutral Dynamics (1964Q1-2004Q4 sub-sample) | | | | | |
| $k^Q_\delta$ | 0.077 | 120.551 | $\phi^Q_{g,r}$ | -1.904 | -2.288 |
| $k^Q_\rho$ | 0.042 | 47.861 | $\phi^Q_{\pi,\pi^*}$ | 7.950 | 5.656 |
| $k^Q_g$ | 0.728 | 10.970 | $\phi^Q_{\pi,\rho}$ | 0.145 | 11.599 |
| $k^Q_\pi$ | $8.7\times10^{-30}$ | $1.2\times10^{-27}$ | $\phi^Q_{\pi,g}$ | -0.018 | -16.673 |
| $k^Q_r$ | 0.061 | 127.000 | $\phi^Q_{\pi,\pi}$ | 0.193 | 4.001 |
| $\phi^Q_{\pi^*,\pi^*}$ | 1.000 | 31.583 | $\phi^Q_{\pi,r}$ | 0.724 | 16.101 |
| $\phi^Q_{\delta,\delta}$ | 0.815 | 51.561 | $\phi^Q_{r,\pi^*}$ | -2.911 | -4.185 |
| $\phi^Q_{\delta,\rho}$ | 0.121 | 20.318 | $\phi^Q_{r,\rho}$ | 0.082 | 33.735 |
| $\phi^Q_{g,\pi^*}$ | -84.984 | -4.408 | $\phi^Q_{r,g}$ | -0.014 | -677.744 |
| $\phi^Q_{g,\rho}$ | 0.831 | 9.525 | $\phi^Q_{r,\pi}$ | 0.372 | 31.207 |
| $\phi^Q_{g,g}$ | 0.362 | 7.500 | $\phi^Q_{r,r}$ | 0.750 | 59.540 |
| $\phi^Q_{g,\pi}$ | 12.020 | 35.509 | | | |
| Under the Risk-neutral Dynamics (2005Q1-2022Q1 sub-sample) | | | | | |
| $k^Q_\delta$ | 0.083 | 59.106 | $\phi^Q_{g,r}$ | -2.209 | -1.217 |
| $k^Q_\rho$ | 0.027 | 5.711 | $\phi^Q_{\pi,\pi^*}$ | 6.190 | 33.129 |
| $k^Q_g$ | 0.465 | 36.109 | $\phi^Q_{\pi,\rho}$ | 0.313 | 10.813 |
| $k^Q_\pi$ | 0.086 | 13.108 | $\phi^Q_{\pi,g}$ | -0.012 | -4.731 |
| $k^Q_r$ | 0.057 | 52.636 | $\phi^Q_{\pi,\pi}$ | 0.195 | 4.765 |
| $\phi^Q_{\pi^*,\pi^*}$ | 1.000 | 43.552 | $\phi^Q_{\pi,r}$ | 0.743 | 11.432 |
| $\phi^Q_{\delta,\delta}$ | 0.936 | 94.825 | $\phi^Q_{r,\pi^*}$ | -3.128 | -10.083 |
| $\phi^Q_{\delta,\rho}$ | 0.182 | 42.983 | $\phi^Q_{r,\rho}$ | 0.067 | 36.714 |
| $\phi^Q_{g,\pi^*}$ | -70.244 | -10.553 | $\phi^Q_{r,g}$ | -0.014 | -43.905 |
| $\phi^Q_{g,\rho}$ | 0.000 | 0.000 | $\phi^Q_{r,\pi}$ | 0.413 | 22.156 |
| $\phi^Q_{g,g}$ | 0.377 | 11.745 | $\phi^Q_{r,r}$ | 0.744 | 23.041 |
| $\phi^Q_{g,\pi}$ | 11.414 | 16.591 | | | |
| Volatility Parameters (1964Q1-2004Q4 sub-sample) | | | | | |
| $l_{\delta,s}$ | 6.108 | 4.832 | $d_{\pi^*}$ | 0.004 | 3.540 |
| $l_{g,\rho}$ | $7.7\times10^{-47}$ | $7.4\times10^{-45}$ | $d_\delta$ | 0.024 | 7.328 |
| $l_{\pi,\rho}$ | 0.003 | 2.426 | $d_\rho$ | 0.094 | 7.303 |
| $l_{\pi,g}$ | 0.012 | 1.576 | $d_g$ | 0.008 | 17.171 |
| $l_{r,s}$ | $-2.7\times10^{-46}$ | $-6.4\times10^{-44}$ | $d_\pi$ | $-6.5\times10^{-4}$ | -7.732 |
| $l_{r,g}$ | 0.139 | 5.647 | $d_r$ | 0.002 | 14.922 |
| $l_{r,\pi}$ | 0.863 | 2.498 | | | |
| Volatility Parameters (2005Q1-2022Q1 sub-sample) | | | | | |
| $l_{\delta,s}$ | 3.383 | 0.919 | $d_{\pi^*}$ | 0.006 | 31.888 |
| $l_{g,\rho}$ | 0.040 | 4.887 | $d_\delta$ | 0.008 | 8.699 |
| $l_{\pi,\rho}$ | 0.002 | 3.480 | $d_\rho$ | 0.193 | 10.790 |
| $l_{\pi,g}$ | 0.013 | 0.995 | $d_g$ | 0.011 | 9.201 |
| $l_{r,s}$ | $-2.3\times10^{-25}$ | $-3.3\times10^{-22}$ | $d_\pi$ | -0.001 | -8.982 |
| $l_{r,g}$ | 0.047 | 4.708 | $d_r$ | 0.001 | 7.441 |
| $l_{r,\pi}$ | 0.022 | 0.208 | | | |

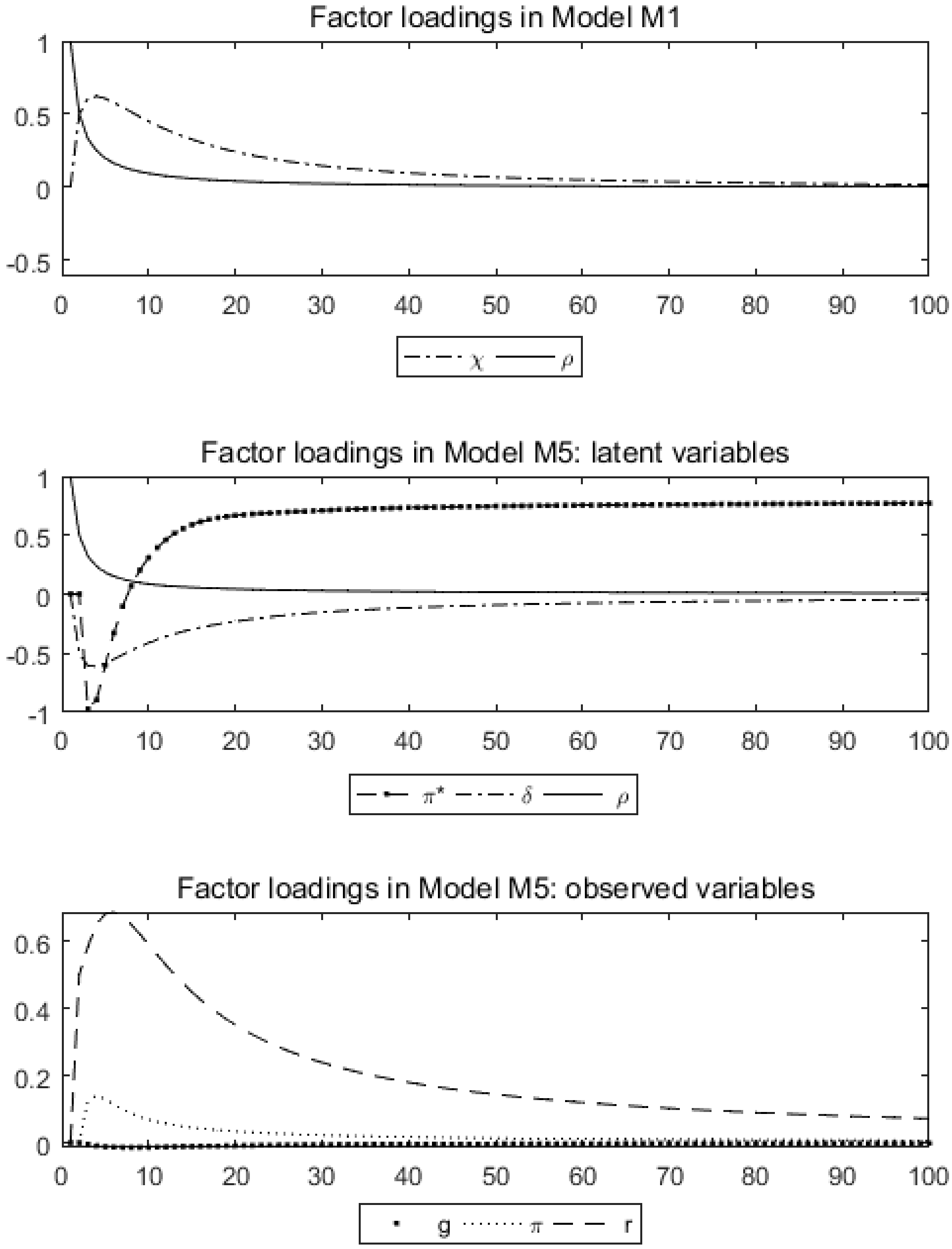
This table presents parameter estimates and their t-statistics for model M5.

Figure 1: The Term Structure of Log Real WTI Oil Futures Contracts




This figure shows the term structure of (log) WTI light crude oil futures deflated by the (log) implicit GDP deflator, $h_{\tau,t}$, as defined by equation (22). Value of data increase from deep blue to dark orange. WTI light crude oil futures started trading on NYMEX in 1984. The data are from Thomson Reuters Datastream. The 1, 2, 3, 6, 9, and 12 months maturities are available from Q1 1984, the 18 month contract from 1989Q3; and the 24 month contract from 1995Q1.

Figure 2: Factor loadings showing the effect of the state variables on the term structure of the cost of carry in models M1 and M5




The behaviour of the futures curve is dictated by the factor loadings. This figure shows factor loadings, $\Psi_\tau/\tau$, (expressed in quarters), in M1 and M5, which depend upon parameters of the risk-neutral factor dynamics (Section 2.10). The spot oil price has a unit effect at the beginning of the futures curve in both models, but its influence fades with maturity. The other variables only affect the slope of the futures curve. The first panel shows the loadings on the two latent oil market variables in M1. The second and third panels show the loadings on the latent and observed variables respectively in M5.

Figure 3: Variables Representing the Macroeconomy and Oil Markets in model M5

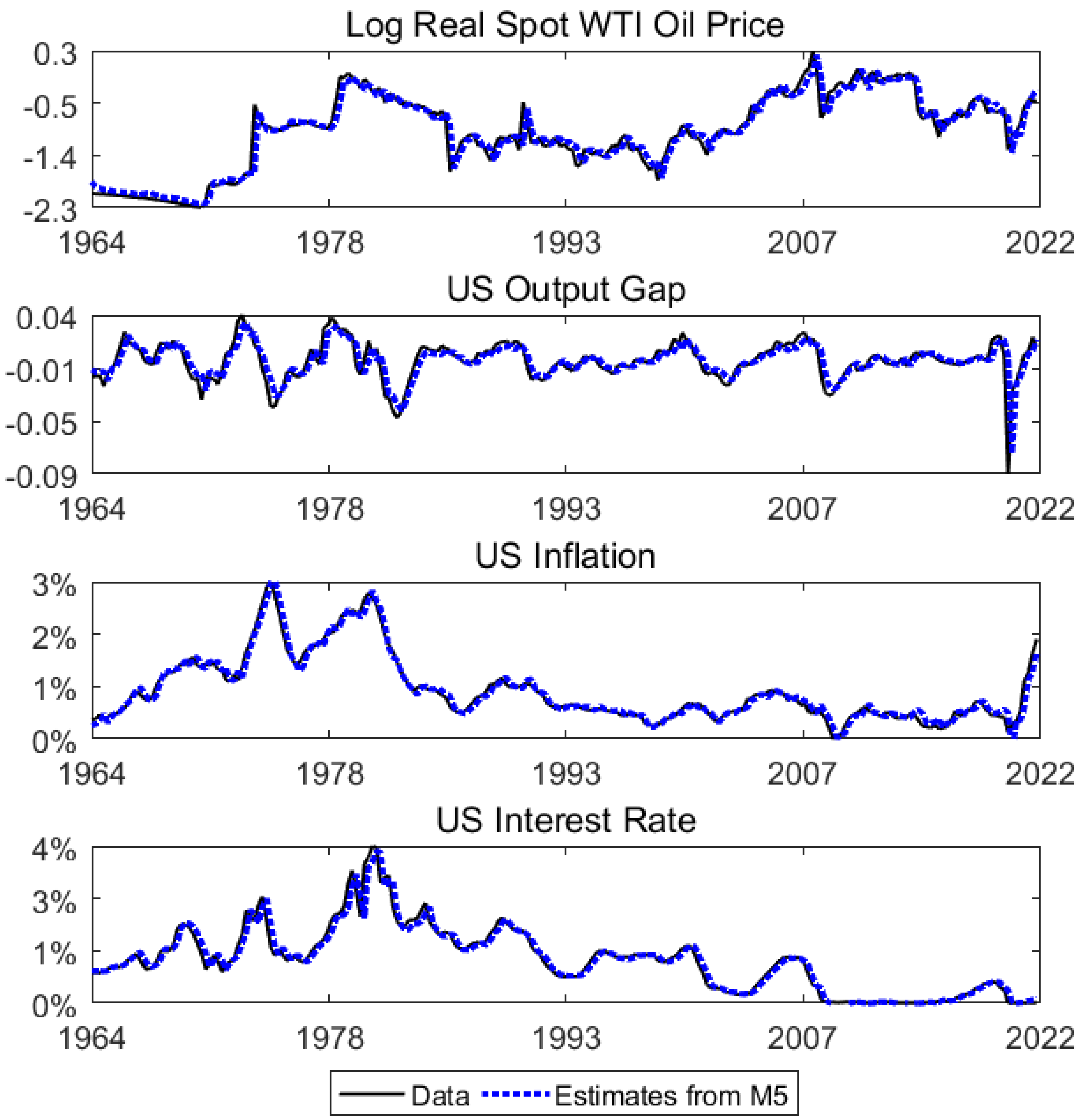
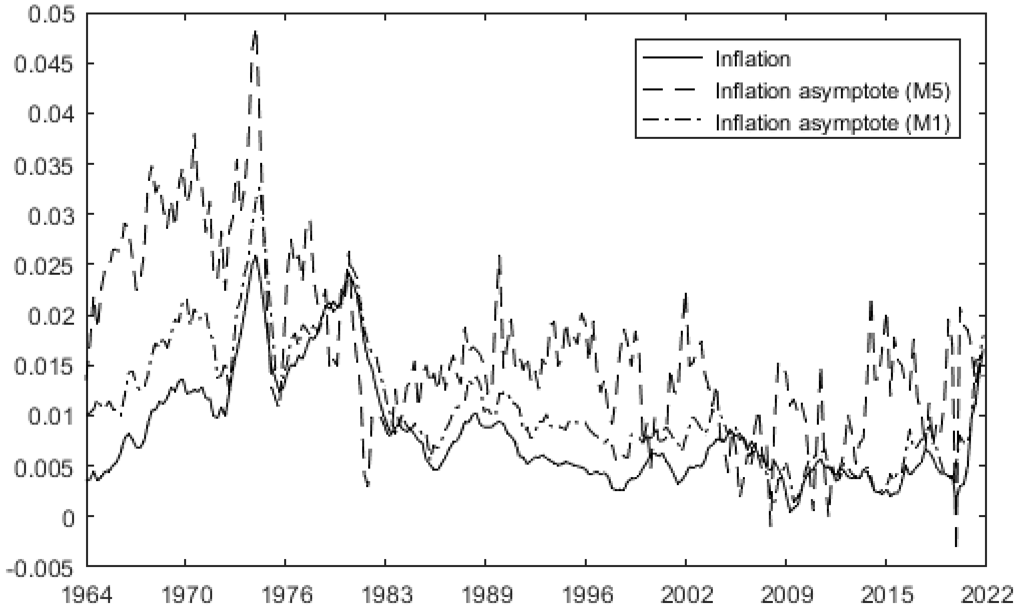
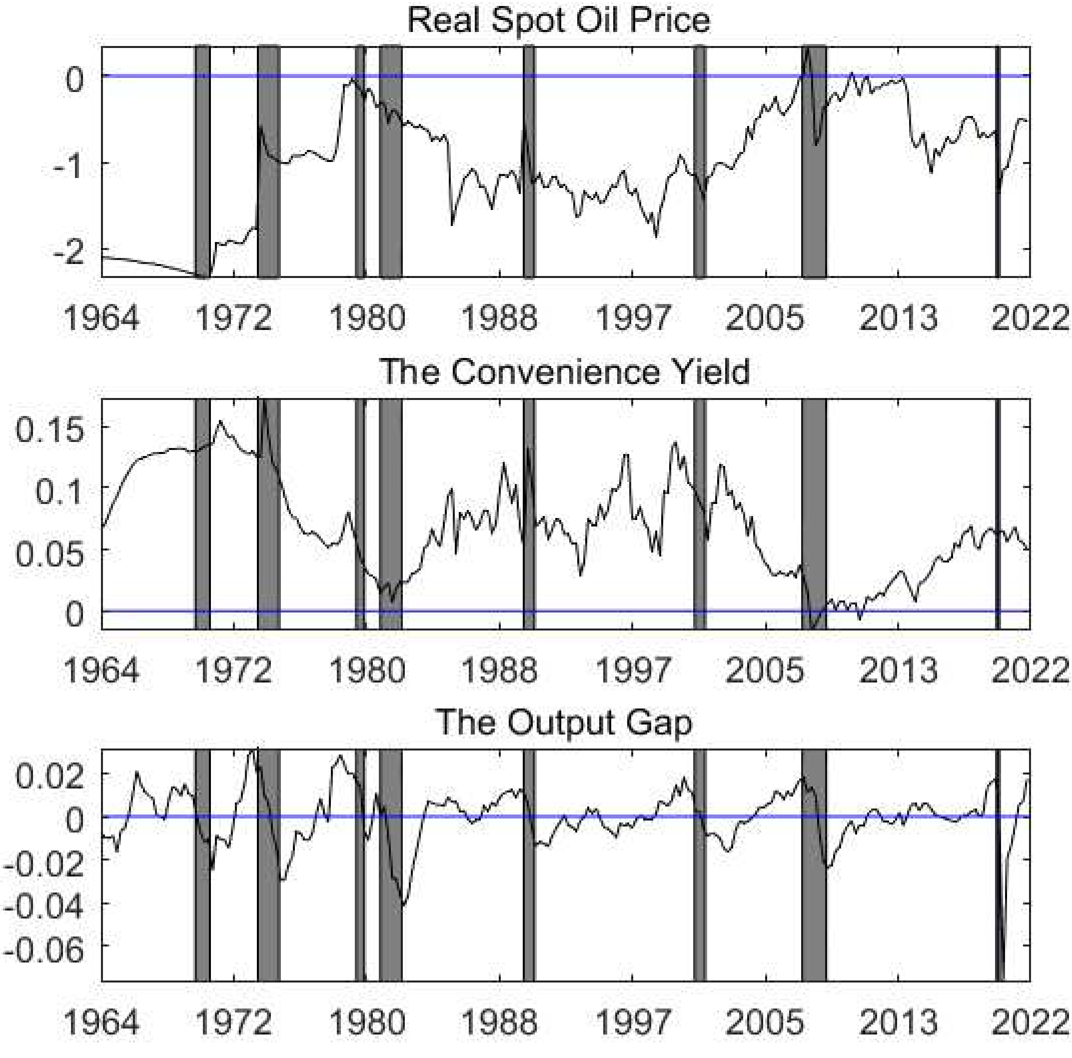


Figure 4: Inflation and the Inflation Asymptote



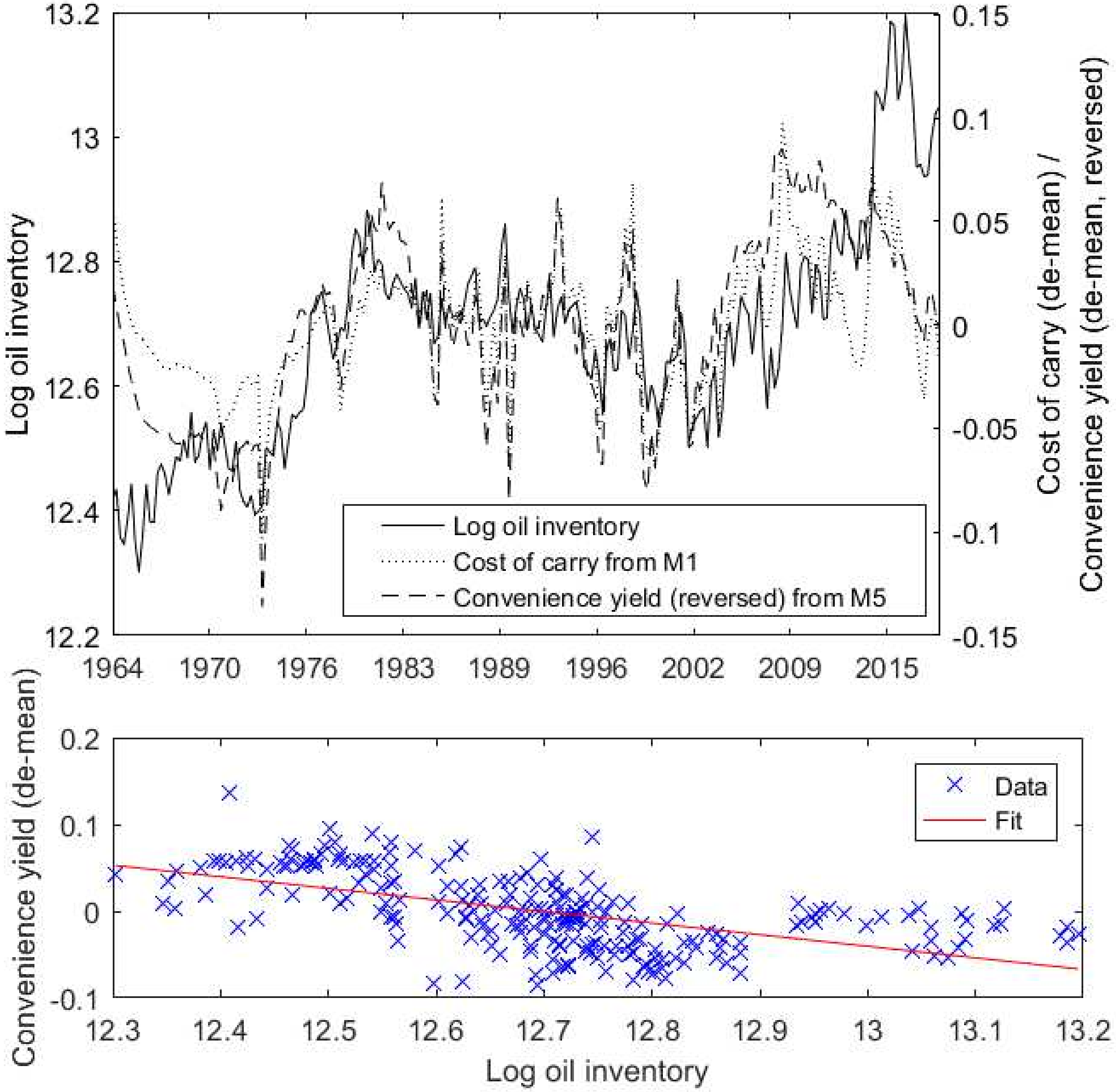

This figure presents the estimates of long term inflation asymptotes of ($\pi^*$) in models M1 and M5 alongside the data for US inflation.

Figure 5: The oil market and the output gap in model M5



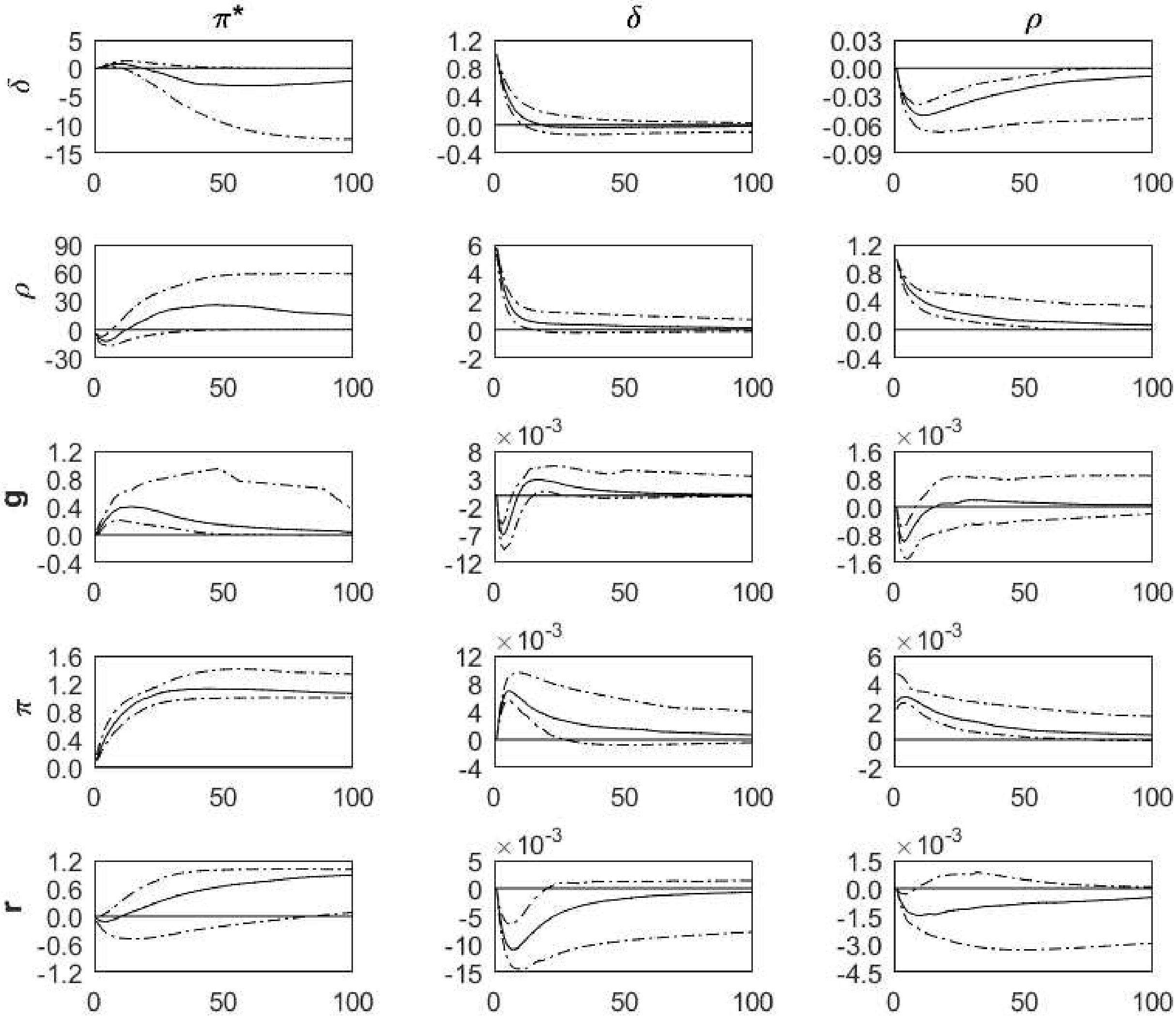

The top panel of this figure plots the real spot price ($\rho$). The middle panel of this figure plots the convenience yield ($d = \delta - \pi$) estimated by model M5. The output gap ($g$) is plotted in the lower panel, and reflects the strength of economic activity. Shadow areas stand for NBER recession periods.

Figure 6: The Convenience Yield and US Oil Inventories



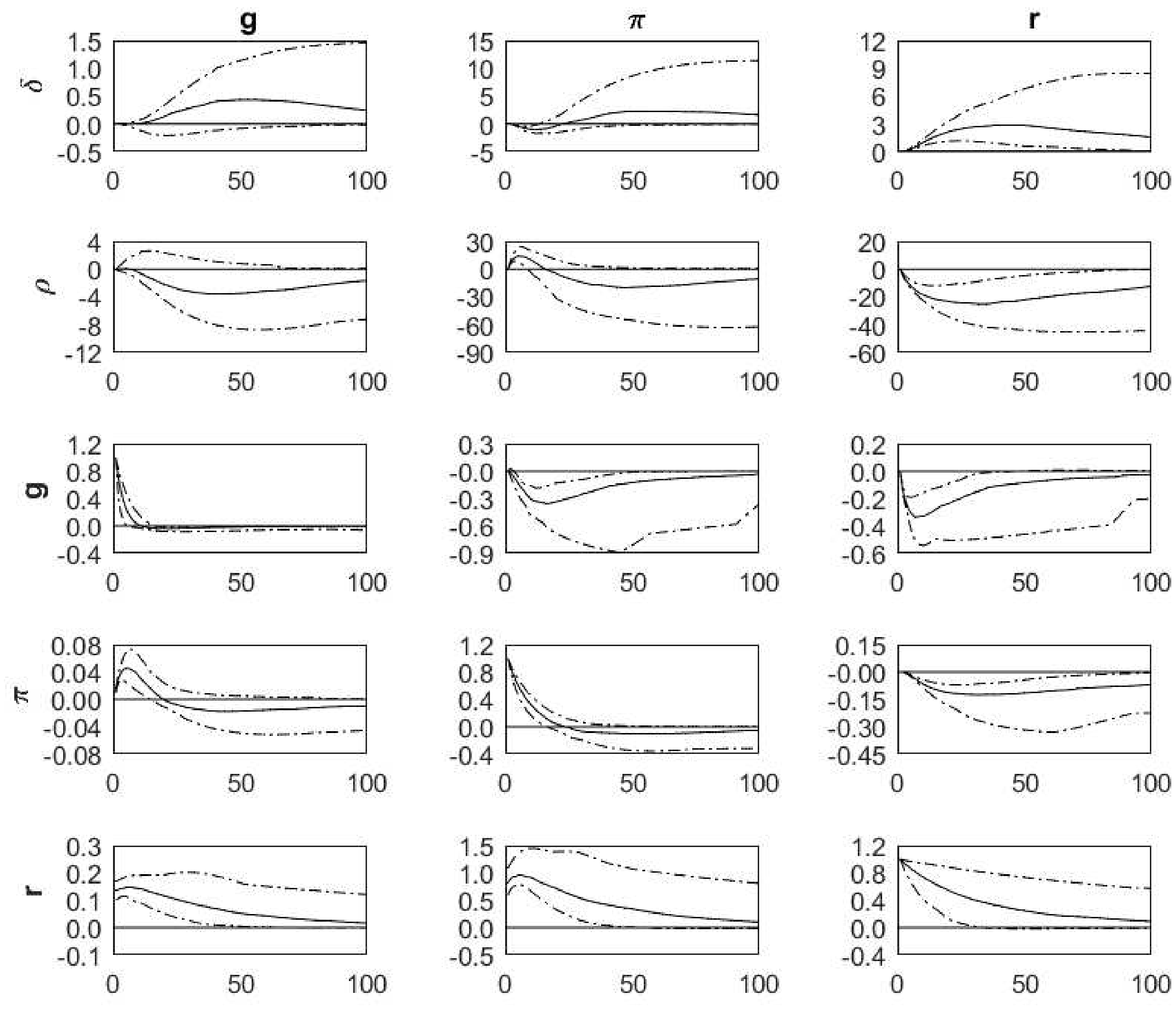

This figure presents estimates of the cost of carry in model M1 ($\chi_t$) the convenience yield ($d = \delta - \pi$) in model M5 alongside the data for log US oil inventories. It shows that the oil inventory is positively correlated with the cost of carry and negatively correlated with the convenience yield.

Figure 7: Responses to latent variable shocks in model M5



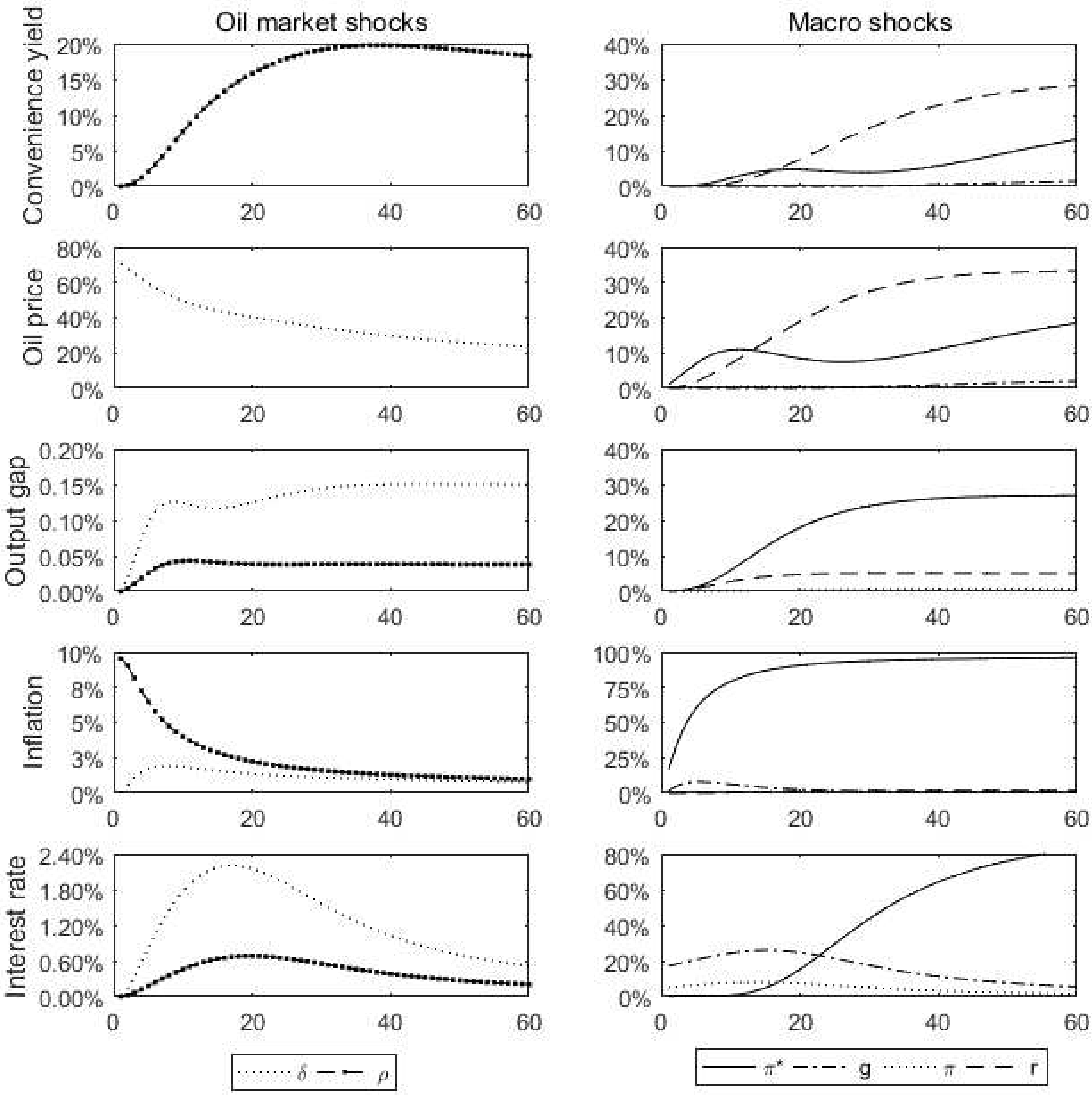

This figure shows how the variables in each row respond to shocks to the latent variables in the model with spanned macro factor risk.

Figure 8: Responses to macro shocks in model M5



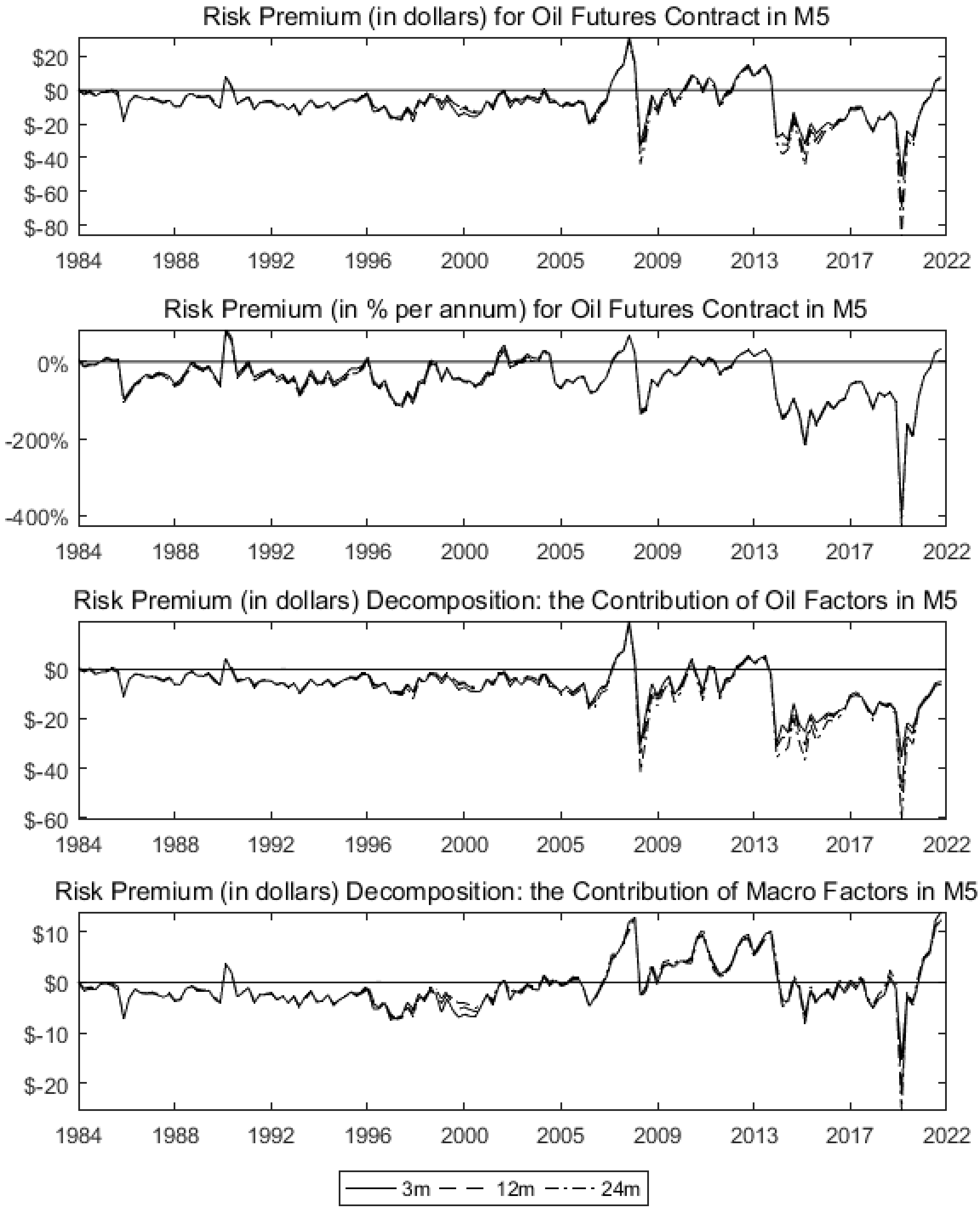

This figure shows how the variables in each row respond to shocks to the macro variables in the model with spanned macro factor risk.

Figure 9: Variance Decomposition of model M5 for the State Variables




This figure reports the results of the Analysis of Variance (ANOVA) exercise for M5. The inflation asymptote ($\pi^*$) is important in explaining the variations of all these variables. Together with shocks to the oil price itself (not shown), shocks to the interest rate and the convenience yield ($\delta$) explain most of the variation in the real oil price. The final row of this figure shows that the variance of the interest rate is strongly influenced by shocks to the inflation target ($\pi^*$). The variance of inflation is affected by real oil price shocks.

Figure 10: Risk Premiums in Futures Contracts




Commodity futures incorporate risk premiums, which reflect the difference between the real world and risk-neutral expectations of the future spot price (see online appendix 4). Oil market variables play an important role here, with the underlying price having a strong positive effect on the premium. The correlation between the contribution of the oil factors ($\rho$, $\delta$) and the 6-month risk premium is 0.8038. However, macro variables are also influential.