Journal of Banking and Finance 95 (2018) 128–147

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Equilibrium commodity prices with irreversible investment and non-linear technologies[☆]

Jaime Casassus[a,1,*], Pierre Collin-Dufresne[b,c], Bryan R. Routledge[d]

[a] Instituto de Economia, Pontificia Universidad Catolica de Chile, Chile
[b] Ecole Polytechnique Federale de Lausanne, Switzerland
[c] NBER, USA
[d] Tepper School of Business, Carnegie Mellon University, USA

ARTICLE INFO

Article history:
Received 18 September 2016
Accepted 1 April 2018
Available online 7 April 2018

JEL classification:
C0
G12
G13
D51
D81
E2,

Keywords:
Commodity prices
Futures prices
Convenience yield
Risk premium
Scarcity
Investment
Irreversibility
General equilibrium
Simulated Method of Moments (SMM)
Regime-switching model

ABSTRACT

We model oil price dynamics in a general equilibrium production economy with two goods: a consumption good and oil. Production of the consumption good requires drawing from oil reserves at a fixed rate. Investment necessary to replenish oil reserves is costly and irreversible. We solve for the optimal consumption, production and oil reserves policy for a representative agent. We analyze the equilibrium price of oil, as well as the term structure of oil futures prices. Because investment in oil reserves is irreversible and costly, the optimal investment in new oil reserves is periodic and lumpy. Investment occurs when the crude oil is relatively scarce in the economy. This generates an equilibrium oil price process that has distinct behavior across two regions (characterized by the abundance/scarcity of oil). We undertake three empirical tests suggested by our model. First, we estimate key parameters using SMM to match moments of oil price futures as well as other macroeconomic properties of the data. Second, we estimate an affine regime switching model of the oil price, which captures the main features of our equilibrium model and preserves the tractability of reduced-form models. Lastly, we compare the risk premium in short-maturity oil futures implied by our model to the data.

1. Introduction

Oil prices, and energy costs in general, play an important role in the economy. Hamilton (2003, 2005) documents that nine out of ten of the U.S. recessions since World War II were preceded by a spike up in oil prices. The "oil-shocks" of the 1970's are the most dramatic example. More interesting, Hamilton (2003) documents that the connection between oil and economic activity is not simple. Oil expenditures account for roughly 4% of US GDP. However, the 7.8% reduction in world oil production in 1973 is associated with a large 3.2% drop in real US GDP. In contrast, the 8.8% drop in production that occurred at the time of the 1991 Persian Gulf War lead only a 0.1% drop on GDP. Many papers have studied the role of oil in real-business-cycle (RBC) models, which turns out to be a difficult task.[2] In this paper we focus on a related question. We investigate the properties of the price of oil (and oil futures) in a general equilibrium model where oil is an important input to the production of the consumption good, and where oil production is characterized by significant adjustment costs and irreversibility.

☆ We thank Luis Alvarez, Leonid Kogan, Duane Seppi, Chester Spatt, Stan Zin, seminar participants at Carnegie Mellon University, University of Madison-Wisconsin, UC Berkeley, University of Southern California, University of Utah, the 2003 LACEA Meeting, the 2005 Q-group meeting, the 2006 Western Finance Association meeting, and the ESEM 2008 Congress.

* Corresponding author.
E-mail address: jcasassus@uc.cl (J. Casassus).

1 Casassus acknowledges financial support from FONDECYT (grant 1141205) and CONICYT through the Anillo Project ACT-1407.

2 Some authors have extended the standard RBC model (e.g. Kydland and Prescott, 1982) to study the response of the economy to exogenous energy shocks. See Kim and Loungani (1992), Finn (1995), Rotemberg and Woodford (1996) and Leduc and Sill (2004) among others.

https://doi.org/10.1016/j.jbankfin.2018.04.001
0378-4266/

Oil prices display a number of interesting characteristics. First, there is tremendous variation in the spot price of oil. Since 1947 the nominal spot price has fluctuated between \$10 and \$140 per barrel. This large volatility is not just present in the spot price, but also in futures prices. Although the volatility is smaller for longer-horizon delivery contracts (the "Samuelson Effect"), even three-year contracts exhibit substantial volatility. More generally, there is substantial variation in the slope of the futures curve. Futures curves tend to be decreasing (i.e., in"backwardation") most of the time (63% in our sample). But at other times futures curves are increasing (i.e., in "contango") or display a single hump. This fact explains why many reduced-form commodity derivative models require at least two factors to fit the futures curve. Typically, these models consider a second factor, the so-called "convenience yield", to capture the variation in the slope of the futures curve. This also implies that the spot price is not a sufficient statistic to characterize the shape of the futures price curve. While oil and futures prices display a number of salient stylized facts, the evidence about whether oil-price risk commands a risk premium is unclear. Gorton and Rouwenhorst (2006) document that there is no risk-premium to holding commodity risk while Erb and Harvey (2006) come to the opposite conclusion. Specifically, the evidence is inconclusive as to whether the futures price of oil is, on average, equal to the future spot price. Understanding the theoretical connection between futures prices and the price of risk will shed some light on this ambiguous evidence.

To understand the properties of the oil price we consider a two-good economy within a general equilibrium model. Such a model is important to connect oil to the fundamentals and to generate an endogenous stochastic discount factor useful to track down the convenience yield and the risk premium in oil spot and futures prices. The consumption good is used for consumption and investment. Production of this consumption good is uncertain and requires as inputs both consumption good and a second good, oil (which could more generally be seen as energy). Oil used in production flows at a fixed rate from existing oil reserves which are also uncertain. New oil can be added to reserves when needed, but this requires an irreversible investment of the consumption good with a fixed-cost component. We first solve for the optimal consumption, investment, and oil reserves policy of a representative agent, and then derive the equilibrium price of oil as well as the term structure of oil futures prices.

Optimal consumption and investment are characterized by the ratio of the stock of oil to the stock of the consumption good (capital), which is the key state variable in the economy. The irreversibility of investment creates an inaction region that makes investment periodic.[3] In this region, investment is postponed until crude oil is scarce enough in the economy. The fixed-cost component in the investment makes the addition of new oil reserves lumpy.[4] Finally, the fixed flow rate of oil in production is crucial to generate a substantial degree of backwardation. These three key assumptions in the oil sector – irreversible investment, a fixed-cost component, and a fixed flow rate of oil — generate significant variation in the state variable that carries over to the equilibrium oil and futures price dynamics.

We define the price of oil as the marginal value of an additional unit of oil reserves. Given this characterization, the central feature of our model is that the lumpy, periodic investment in new oil implies that the spot price of oil is a humped-shaped function of the state variable. Indeed, the equilibrium oil price reaches a maximum price above the marginal cost of adding new oil stocks. This non-monotonic response of the oil price is due exclusively to the fixed-cost component of the investment in new oil stocks. When oil is plentiful (oil/capital ratio is high), investment in new oil reserves is a long way off. As oil is used in production the ratio of oil-to-capital decreases and the price of oil increases. In contrast, when oil is scarce (oil-to-capital ratio is low), we are near the investment boundary. In this region, as oil is used and the oil-to-capital ratio declines, the price of oil falls. Despite the fact the current quantity of oil is smaller, the marginal value of a unit of oil is smaller since the near-term investment will increase the stock of oil. The result is that the dynamic properties of the oil price (and hence futures prices) are notably different in the abundance and scarcity regions. In effect, the price of oil is governed by a process with two regimes.

Our model is similar to partial equilibrium commodity models that focus on the role of commodity storage.[5] These models typically assume an inverse (net) demand function that maps the quantity of the commodity and a shock into a price. The fact that inventories must be non-negative produces an asymmetry in the spot price behavior at zero inventory (stock-out). This generates interesting dynamic properties for the commodity spot price and the related futures prices, in particular, backwardation arises from a fear of stock-out in the near future. Storage models are similar to our general equilibrium model in that they produce a two-regime process for the spot price. Our general equilibrium model adds to the storage approach along a few dimensions. First the zero-inventory or stock-out event that is so central to inventory models is infrequently observed in oil markets. For example, the oil-futures curve is downward sloping 63% of the time. It is hard to see zero (or even near-zero) inventory levels anywhere near this frequency. Of course, this fact by itself may not be troubling as inventories are difficult to measure and the storage technology in these models does not capture the complexities of inventory. More troubling, however, is the difficulty inventory models have explaining longer-horizon dynamics. In order to explain frequent backwardation, models must be calibrated so inventory levels hit zero frequently. As a result, the inventory state variable has little impact over price levels beyond the short horizon of inventory. Routledge et al. (2000), for example, normalize oil-futures prices so that the twelve-month futures price has zero variance. Lastly, due to computational limitations, this literature typically assume risk-neutrality and hence do not have implications for commodity risk premia.

Our paper is also related to other equilibrium models that study the behavior of commodity prices. Carlson et al. (2007) propose an equilibrium model of natural resources and study the effect of adjustment costs in the forward price dynamics. However, in contrast to our paper, they assume risk-neutrality, an exogenous demand function for commodity, and (the main friction in their model) that commodity is exhaustible, whereas in our paper commodity is essentially present in the ground in infinite supply but is costly to extract. Also, Kogan et al. (2009) identify a new pattern of futures volatility term structure that is inconsistent with standard storage models. This new pattern can be explained within their industry equilibrium model that exhibits investment constraints and irreversibility. Similar to our scarcity and abundance regions, they have investment and no-investment regions with different spot price dynamics. Unlike our model, they take the demand side and

[3] This is a well-known result from the investment under uncertainty literature when investment is irreversible, see for example, Abel (1983), Bernanke (1983), Dixit and Pindyck (1994) and Bloom et al. (2007).

[4] The U.S. crude oil and natural gas drilling activity is a result of investment in these energy sectors. Since 1974, the two-year count of rotary rigs in operation has spiked twice by more than 80% (in the 1980–1982 and 2000–2002 periods). We use this as evidence of lumpiness in oil investments.

[5] There are many papers that take this approach. For example, Gustafson (1958), Newbery and Stiglitz (1981), Wright and Williams (1982), Scheinkman and Schechtman (1983), Williams and Wright (1991), Deaton and Laroque (1992), Chambers and Bailey (1996), Routledge et al. (2000), and Bobenrieth et al. (2002).

risk premia as exogenous and focus mainly on the implications for the volatility curve. Casassus et al. (2013) extend a simplified version of the model proposed here to a multi-commodity framework tailored to study the long-term correlation between futures returns. Recently, Ready (2016) and Hitzemann (2016) extend the production-based asset pricing framework with long-run productivity risk of Croce (2014) to study the relation between oil consumption, oil prices and macroeconomic fluctuations. As opposed to our model, Ready (2016) builds a partial equilibrium model with exogenous oil supply, while Hitzemann (2016) assumes that oil enters directly in the utility function of the representative household and not as an input for the production technology, like in our model.[6]

To illustrate the quantitative features and advantages of our general equilibrium model, we undertake three (related) empirical exercises. First we estimate our model using Simulated Method of Moments (SMM) estimation proposed by Lee and Ingram (1991) and Duffie and Singleton (1993). We estimate the parameters of our model to match the sample average term-structure of the oil futures price and the sample volatility of the term-structure along with some aggregate macroeconomic moments. With (arguably) sensible parameters, we match the term-structure of futures prices and volatilities reasonably well. The model generates backwardation (negative slope in the futures curve) on average. The model also captures the level and shape of the term-structure of volatilities pretty well. In theory, the dynamic properties of the oil price are different in the scarcity and abundance regions. This is only relevant empirically if the (endogenous) state variable has a distribution that accesses these regions frequently enough to matter. Under our estimated model parameters, this is indeed the case. The economy is in scarcity of oil 14% of the time and in abundance 86% of the time.

One of the disadvantages of a general equilibrium model is that it is challenging to compute. In practice, commodity pricing is often implemented using an exogenously specified process of the spot price and a convenience yield (a second factor that captures the dynamics of the futures curve slope). In this setting, pricing futures contracts and other derivatives is straightforward (e.g., Gibson and Schwartz, 1990; Schwartz, 1997 and Casassus and Collin-Dufresne, 2005). This approach is very effective since it is tractable enough to price complex commodity-contingent claims, but often suffers from the lack of theoretical justification for the underlying dynamics chosen for the state variables.

Our second empirical exercise is motivated by trying to link both approaches. In the model, the equilibrium process for the oil price differs across the scarcity and abundance regions. However, it turns out that within those regions the spot price behavior is simple to characterize. In particular, within a regime, a linear approximation to the drift and volatility functions is a reasonable approximation of our equilibrium price. We use quasi-maximum likelihood technique of Hamilton (1989) to estimate a two-regime model with crude oil price data. The resulting parameter estimates on the drift and volatility are in line with the calibrated model. Moreover, since the stochastic process is, up to the two-regimes, affine, we expect the reduced-form approximation of our model to be useful for derivative pricing. In addition, we estimate the smoothed inference about whether the state of the economy is in the scarcity or the abundance regions. This is a helpful diagnostic on the model since we can confirm that the data are consistent with a two-regime process that is economically meaningful (both regimes are visited frequently).

One of the advantages of a model in general equilibrium is that we can investigate the risk premium properties of the oil commodity. Interestingly, in our model the risk premium on oil is state-dependent. In the abundance region, a long position in oil (e.g., an oil futures) is risky and commands a risk premium. In the scarcity region, a long position in oil is a hedge and the risk premium is negative. Our third empirical exercise is to investigate the conditional risk premium in oil-futures data. In regressing oil price return on the market return we find that the beta is significantly negative in the estimated scarcity regime and positive (though not statistically significant) in the other regime. The difference across regimes is consistent with our model. More importantly, variation in the commodity risk premium makes unconditional tests of the risk premium more difficult to interpret (e.g., Gorton and Rouwenhorst, 2006 and Erb and Harvey, 2006).

Our paper proceeds with the model developed next in Section 2. The equilibrium properties are characterized in Section 3 (with most of the proofs contained in the appendix). Section 4 contains the empirical implementations of our model where we estimate with SMM, characterize a two-regime affine model, and explore the properties of the commodity risk premium. Finally, Section 5 concludes.

2. The model

We consider an infinite horizon production economy with two goods. The model extends the Cox et al. (1985) production economy to the case where the production technology requires a second input which we interpret as oil. The commodity is not directly consumed but is a required input for the production of the consumption (numeraire) good. We solve the representative agent's optimal consumption/investment policy and use this to characterize equilibrium prices.

2.1. Representative agent characterization

There is a representative agent with time-separable, constant relative risk aversion expected utility preferences over lifetime consumption of a single numeraire good

$$E_0 \int_0^{\infty} e^{-\rho s} \frac{C_s^{1-\gamma}}{1-\gamma} ds$$

The rate of time preference is $\rho > 0$ and the level of relative risk aversion is $\gamma > 0$ (with the usual understanding that $\gamma = 1$ is the log risk aggregator). In particular, note that only one good is consumed. The single consumption good in our economy has stock denoted K_t. There is a second good in our economy, oil, that is necessary for consumption good production. The stock oil reserves is Q_t.[7] The dynamics of the consumption good stock, K_t, and oil stock, Q_t, are given by

$$dK_t = \left(\alpha K_t^{1-\eta}\left(\bar{i} Q_t\right)^{\eta} - C_t\right)dt - \beta(K_t, Q_t, X_t)dI_t + \sigma_K K_t dw_{K,t} \tag{1}$$

$$dQ_t = -(\bar{i} + \delta)Q_t dt + X_t dI_t + \sigma_Q Q_t dw_{Q,t} \tag{2}$$

The growth rate of the stock of consumption good (capital) is determined by capital and the quantity of oil used in production. The two input goods follow a Cobb–Douglas production function with parameter α as the productivity of capital and η determining the importance of oil in the economy.

[6] There are several recent studies that consider equilibrium models to better understand different aspects of commodities, such as, the determinants of futures returns (Acharya et al., 2013; Yang, 2013; David, 2015; Baker and Routledge, 2016) or the so-called "financialization of commodity markets" (Hamilton and Wu, 2014; Sockin and Xiong, 2015; Basak and Pavlova, 2016; Baker, 2016).

[7] One could also think of Q_t as the number of oil wells, each of which produces a flow of oil.

We assume that the amount of oil used in production is proportional to the stock of oil reserves Q_t. A fixed flow rate $\bar{i}$ captures the physical difficulty in adjusting rapidly the extraction flow of known oil reserves.[8] Most importantly, this rigidity in the oil sector is necessary to generate the degrees of backwardation observed in the data.[9] Allowing the rate of flow to be optimally chosen at no cost would imply forward prices rising at the riskless interest rate (this is an implication from Hotelling, 1931). As oil is used in production its stock is depleted. The stock of oil also depreciates at rate δ if not used.

New oil reserves can be created by investing the consumption good. Specifically, X_t units of new oil capacity can be created at t by investing $\beta(K_t,Q_t,X_t)$ units of the consumption good. Investment in new oil stocks is irreversible ($X_t \geq 0$) implying that the addition of new oil will be periodic. We denote $dI_t = 1$ if there is investment at t and 0 otherwise. This investment includes also a fixed cost component. Specifically, we assume that building X_t oil reserves costs

$$\beta(K_t, Q_t, X_t) = \beta_K K_t + \beta_Q Q_t + \beta_X X_t \tag{3}$$

where $\beta_K, \beta_Q, \beta_X > 0$. The fixed cost component is $\beta_K K_t + \beta_Q Q_t$. This cost is scaled by the size of the economy K_t and Q_t to maintain the homotheticity of the model. Practically, if the fixed cost were not scaled its influence over investment behavior would vanish as the economy grows. The importance, of course, of the fixed-cost assumption is that the creation of new oil is an 'impulse control' optimization problem, where the optimal investment decision occurs only at discretely dates (it is lumpy). In order for oil-stock investment to be feasible, we assume $\beta_K < 1$ and $\beta_Q < \beta_X$.

Our model abstracts from technological improvements in the oil sector by assuming that importance of oil to production (η) and investment-cost parameters (β_K, β_Q, and β_X) are constant. This assumption delivers a homothetic (hence, tractable) model. However, they ignore technological improvement in oil use and exploration (e.g., advances in miles-per-gallon and the changes in oil exploration technology). This is surely an important aspect but we leave it to future research.

Finally, uncertainty in our economy is captured by the Brownian motions $w_{Q,t}$ and $w_{K,t}$ which drive the diffusion terms in Eqs. (1) and (2). We allow the shocks to be correlated, denoted ρ_{KQ}. As is standard, we assume that there exists an underlying probability space (Ω, $\mathbf{F}$, P) satisfying the usual conditions, and where Ω is the sample space, $\mathbf{F} = \{\mathcal{F}\}_{t\geq 0}$ is the natural filtration generated by the Brownian motions and P is a function that assigns probabilities to the events.

2.2. Optimal consumption and investment strategies

Given the fixed cost of irreversible investment, it is natural to seek an oil investment policy of the form $\{(X_{T_i}, T_i)\}_{i=0,1,\ldots} \in \mathcal{A}$ where $\{T_i\}_{i=0,\ldots}$ is a sequence of stopping times of the filtration $\mathbf{F}$ such that $I_t = \mathbf{1}_{\{T_i \leq t\}}$ and the X_{T_i} are $\mathcal{F}_{T_i}$-measurable random variables. The set of admissible strategies $\mathcal{A}$ ensure the consumption good stock process is positive ($K_t > 0$ *a.s.*). Lastly, the set of allowable consumption policies $\mathcal{C}$ must be positive integrable $\mathbf{F}$ adapted processes. The optimal consumption–investment policy of the representative agent has the 'current' value function

$$J(K_t, Q_t) = \sup_{\mathcal{C};\mathcal{A}} E_t\left[\int_t^\infty e^{-\rho(s-t)} \frac{C_s^{1-\gamma}}{1-\gamma} ds\right] \tag{4}$$

and subject to the laws-of-motion in Eqs. (1) and (2).

Given that investment in new oil is irreversible ($X_t \geq 0$) and the presence of fixed costs, the optimal investment will be infrequent and 'lumpy' (e.g., Dumas, 1991). This defines two zones of the state space $\{K_t, Q_t\}$: A *no-investment region* where $dI_t = 0$ and an *investment region* where $dI_t = 1$. This is analogous to the shipping cone in Dumas (1992), but with only one boundary because investment is irreversible. Appendix A establishes sufficient conditions on the parameters for a solution to the problem in (4) to exist. We also present there a closed-form solution to the problem when oil investment is perfectly reversible.

2.2.1. Optimal consumption in the no-investment region

When the state variables $\{K_t, Q_t\}$ are in the no-investment region, a portion of the consumption good can be consumed with the remainder used in consumption-good production (along with oil). This region is defined by the condition $J(K_t - \beta(K_t, Q_t, X_t), Q_t + X_t) < J(K_t, Q_t)$ since it is not optimal to use the consumption good to invest in new oil reserves. In this region, the solution of the problem in Eq. (4) is determined by the following the Hamilton–Jacobi–Bellman (HJB) equation:

$$\sup_{\{C\geq 0\}}\left\{-\rho J + \frac{C^{1-\gamma}}{1-\gamma} + \mathcal{D}J\right\} = 0 \tag{5}$$

where $\mathcal{D}$ is the Itô operator

$$\mathcal{D}J(K,Q) \equiv \left(\alpha K^{1-\eta}\left(\bar{i}Q\right)^{\eta} - C\right)J_K - (\bar{i}+\delta)QJ_Q + \frac{1}{2}\sigma_K^2 K^2 J_{KK} + \frac{1}{2}\sigma_Q^2 Q^2 J_{QQ} + \rho_{KQ}\sigma_K\sigma_Q KQJ_{KQ} \tag{6}$$

with J_K and J_Q representing the marginal value of an additional unit of consumption good and oil respectively. Also, J_{ij} are the second-order derivatives with respect to i and j for $i, j \in \{K, Q\}$.

To characterize optimal consumption we derive the first order condition for Eq. (5)

$$C^* = J_K^{-\frac{1}{\gamma}}. \tag{7}$$

At the optimum, the marginal value of consumption is equal to the marginal value of an additional unit of the consumption good.

2.2.2. Optimal investment with fixed costs and irreversibility

It is optimal to invest in new oil reserves when the benefit of additional oil reserves exceeds its cost. This condition is $J(K_t - \beta(K_t, Q_t, X_t), Q_t + X_t) \geq J(K_t, Q_t)$.[10] Investment occurs when crude oil is relatively scarce in the economy. Let $J_1 = J(K_t^*, Q_t^*)$ be the value function immediately before investment and $J_2 = J(K_t^* - \beta(K_t^*, Q_t^*, X_t^*), Q_t^* + X_t^*)$ be the value function immediately after the investment is made.

To determine the investment threshold level of consumption good K_t^*, the amount of oil Q_t^*, and the size of the optimal oil investment X_t^* at the investment boundary we consider the standard value-matching and super-contact (smooth pasting) conditions. The former states that at the threshold levels the agent is indifferent between investing and not investing, therefore,

$$J_1 = J_2 \tag{8}$$

while the latter are higher-order conditions that ensure optimality and imply[11]

$$J_{1K} = (1-\beta_K)J_{2K} \tag{9}$$

[8] Anecdotal stories describe oil wells in Siberia where the flow cannot be interrupted to avoid freezing. Generically, modulating the speed at which oil can be extracted from existing wells and delivered to end-users is a costly and inflexible process.

[9] Litzenberger and Rabinowitz (1995) and Carlson et al. (2007) discuss how frictions in oil production can generate a downward sloping futures curve.

[10] Without loss of generality we assume that the initial capital stocks $\{K_0, Q_0\}$ are in the no-investment region. If this is not the case, there is an initial lumpy investment that takes the state variables into the no-investment zone.

[11] The optimality conditions for an economy without fixed costs (i.e.,$\beta_K = \beta_Q = 0$) result directly from Eq. (8). In that case, two additional super-contact conditions are necessary to characterize the optimal investment: $-J_{1QK} + \beta_X J_{1KK} = 0$ and $-J_{1QQ} + \beta_X J_{1KQ} = 0$. These conditions are due to an 'infinitesimal control' optimiza-

$$(\beta_X - \beta_Q)J_{1K} = (1-\beta_K)J_{1Q} \quad (10)$$

$$\beta_X J_{2K} = J_{2Q} \quad (11)$$

2.2.3. Reduction of number of state variables

Given homogeneous functions, a sufficient statistic characterizing the economy is the ratio of the oil reserves to capital stocks.[12] Specifically, denoting by z_t the log of the oil reserves to capital ratio in period t, $z_t = \log(Q_t/K_t)$ measures the relative abundance of crude oil in the economy. Then the value function in Eq. (4) is reduced to

$$J(K_t, Q_t) = \frac{K_t^{1-\gamma}}{1-\gamma} j(z_t) \quad (12)$$

with $j(z_t)$ being the value function defined under the sufficient statistic.

While the dynamic process for z_t can be obtained from a generalized version of Itô's Lemma

$$dz_t = \mu_{z,t} dt + \sigma_z dw_{z,t} + \Lambda_z dI_t \quad (13)$$

one can characterize two regions such as the no-investment and investment regions solely based on z_t. Let $z_1 = \log(Q_t^*) - \log(K_t^*)$ be the log oil to capital ratio just prior to investment. Similarly, define $z_2 = \log(Q_t^* + X_t^*) - \log(K_t^* - \beta(K_t^*, Q_t^*, X_t^*))$ as the log ratio immediately after the optimal investment in oil occurs. When $z_t > z_1$ it is optimal to postpone investment in new oil. If the state variable z_t reaches z_1, an investment to increase oil stocks by X_t^* is made. z_1 delineates the investment trigger which occurs when the relative scarcity of oil is at its highest level. The result is that the state variable jumps to z_2 which is inside the no-investment region, therefore the size of the jump in z_t is given by $\Lambda_z = z_2 - z_1$. Given the cost function in Eq. (3), the optimal investment in new oil reserves relative to current oil stock,x_t, is just a function of z_1 and z_2

$$x_t^* \equiv \frac{X_t^*}{Q_t^*} = \frac{e^{-z_1} - e^{-z_2} - (\beta_K e^{-z_1} + \beta_Q)}{e^{-z_2} + \beta_X} \quad (14)$$

The jump in oil reserves is

$$\frac{Q_2}{Q_1} \equiv \frac{Q_t^* + X_t^*}{Q_t^*} = 1 + x_t^* \quad (15)$$

and, we can express the jump in the consumption good stock simply as

$$\frac{K_2}{K_1} \equiv \frac{K_t^* - \beta(K_t^*, Q_t^*, X_t^*)}{K_t^*} = 1 - (\beta_K + \beta_X x_t^* e^{z_1} + \beta_Q e^{z_1}) \quad (16)$$

Finally, the optimal consumption from (7) can be rewritten in terms of $j(z_t)$ as

$$c_t^* \equiv \frac{C_t^*}{K_t} = \left(j(z_t) - \frac{j'(z_t)}{1-\gamma} \right)^{-\frac{1}{\gamma}} \quad (17)$$

Plugging the definition of $J(K_t, Q_t)$ into the HJB equation in (5) returns a one-dimensional ordinary differential equation (ODE) for the function $j(z_t)$. This substitution also yields one-dimensional value-matching and super-contact conditions. Appendix B contains these equations as well as the formal proofs of optimality and details on the derivations. The system of equations does not have (to the best of our knowledge) a closed-form solution, but Appendix D sketches the numerical technique used to solve for the equilibrium.

3. Equilibrium prices

The solution to the representative agent's problem of Eq. (4) characterizes equilibrium prices.[13] In this section, we begin by characterizing the stochastic discount factor which exhibits a periodic jump due to the lumpy investment in the oil sector. Then we characterize the price of oil and the net convenience yield as functions of state variables and parameters in the model. Lastly, we derive the partial differential equation for oil futures prices.

3.1. Asset prices and the stochastic discount factor

We assume our model has dynamically complete markets so the stochastic discount factor (SDF) is characterized by the representative agent's marginal utility (see Duffie, 2001)

$$\xi_t = e^{-\rho t} \left(\frac{C_t}{C_0} \right)^{-\gamma} \quad (18)$$

with $\xi_0 = 1$. We replace optimal consumption from Eq. (7) and apply the generalized Itô's Lemma to get the SDF dynamics

$$\frac{d\xi_t}{\xi_t} = -\frac{dB_t}{B_t} - \lambda_{K,t} dw_{K,t} - \lambda_{Q,t} dw_{Q,t} \quad (19)$$

where

$$\frac{dB_t}{B_t} = r_t dt + \Lambda_B dI_t \quad \text{with} \quad B_0 = 1 \quad (20)$$

$$r_t = (1-\eta)\alpha \left(\frac{\bar{i} Q_t}{K_t} \right)^{\eta} - \sigma_K \left(\lambda_{K,t} + \rho_{KQ} \lambda_{Q,t} \right) \quad (21)$$

$$\Lambda_B = -\frac{\beta_K}{1-\beta_K} \quad (22)$$

$$\lambda_{R,t} = -\sigma_R \frac{R_t J_{KR}}{J_K} = \gamma\, \sigma_R \frac{R_t C_R}{C} \quad \text{for} \quad R \in \{K, Q\} \quad (23)$$

The price of the risk-free money market, B_t, in Eq. (20) has a continuous component in the no-investment and a singularity at the investment boundary. Eqs. (21)–(23) represent the endogenously determined interest rate, the jump in the SDF and the market prices of risk. It is straightforward to show that each of these terms are functions solely of the state variable z_t.

The instantaneous risk-free rate,r_t, corresponds to the marginal productivity of capital minus a risk adjustment for shocks in the capital sector. Λ_B in Eq. (22) is a jump in financial market prices (and in the SDF) that occurs when the lumpy investment in the oil industry is made. This jump arises because the representative agent's marginal utility has a singularity at the investment boundary due to the fixed cost β_K.[14] This jump is predictable, in the sense that it triggers when the state variable reaches the investment boundary. To rule out arbitrage, all asset prices in the economy jump by the same amount (see Karatzas and Shreve, 1998).

tion problem where the optimal investment for this particular case is a continuous regulator (Harrison, 1990), i.e., the consumption good and oil stocks are the same before and after investment. See Dumas (1991), Dixit (1991, 1993) for a discussion of value-matching and super-contact (smooth-pasting) conditions.

[12] The consumption good production function is homogeneous of degree one and the utility function is homogeneous of degree $(1-\gamma)$.

[13] The fixed cost of building new oil reserves makes decentralizing the economy not straightforward. Specifically, it is hard to reconcile the "zero profit" condition of perfect competition with the need to recover fixed investment costs. See Guesnerie (1975) and Romer (1986), for examples.

[14] The super-contact condition in (9) shows that J_K – the marginal utility of consumption – has a singularity at the investment boundary for a positive fixed cost for capital (i.e.,$\beta_K > 0$).

Finally, the equilibrium market prices of risk in Eq. (23) measure the sensitivity of the representative agent's marginal utility to the shocks in the capital and oil sectors. In our consumption-based asset pricing model, the prices of risk $\lambda_{K,t}$ and $\lambda_{Q,t}$ increase with the level of uncertainty (σ_K and σ_Q, respectively), the relative risk aversion coefficient γ and the capital and oil elasticities of consumption. To express the market prices of risk in terms of consumption and its derivatives as in the last term of Eq. (23), we use the first order condition in Eq. (7) to replace J_K and J_{KR}.

3.2. Marginal price of oil

We price the value of a barrel of oil using the standard utility indifference argument. We define the spot oil price (S_t) as the marginal value of an incremental unit in oil capacity. Specifically,S_t solves $J(K_t, Q_t) = J(K_t + S_t\epsilon, Q_t - \epsilon)$ as $\epsilon \to 0$. With a Taylor expansion, this implies $S_t = J_Q/J_K$ or, equivalently, in terms of the state variable z_t,

$$S_t = \frac{e^{-z_t} j'(z_t)}{(1-\gamma)j(z_t) - j'(z_t)} \tag{24}$$

Our definition of the oil price,S_t is literally the marginal price of adding one unit of oil to the reserves, in terms of the consumption good. Since one cannot adjust the rate $\bar{i}$ at which oil reserves are used in our model, this seems the appropriate definition. In particular, this implies that the price of oil is different from the marginal productivity of oil, because agents cannot alter the demand rate $\bar{i}$.

The following proposition presents the equilibrium spot oil price process in our economy:

Proposition 1. *The spot price has the following dynamics*

$$\frac{dS_t}{S_t} = \mu_{S,t}dt + \sigma_{SK,t}dw_{K,t} + \sigma_{SQ,t}dw_{Q,t} + \Lambda_S dI_t \tag{25}$$

where

$$\mu_{S,t} = r_t - y_t + \pi_t \tag{26}$$

$$\pi_t = \sigma_{SK,t}(\lambda_{K,t} + \rho_{KQ}\lambda_{Q,t}) + \sigma_{SQ,t}(\lambda_{Q,t} + \rho_{KQ}\lambda_{K,t}) \tag{27}$$

$$\sigma_{SR,t} = \lambda_{R,t} - \theta_{R,t} \quad \text{for} \quad R \in \{K, Q\} \tag{28}$$

$$\theta_{R,t} = -\sigma_R \frac{R_t J_{RQ}}{J_Q} \quad \text{for} \quad R \in \{K, Q\}$$

$$\Lambda_S = \frac{\beta_Q - \beta_K\beta_X}{\beta_X - \beta_Q} \tag{29}$$

Proof. We apply Itô's Lemma to the definition of the spot price $S_t = J_Q/J_K$. To identify the components from the drift $\mu_{S,t}$ in Eq. (26), we use that the oil risk premium π_t is the negative correlation between oil returns and the SDF (see Cochrane, 2005). The net convenience yield y_t will be characterized in Proposition 2. Since the price is a function solely of the state variable z_t, all the variables in the proposition are only functions of z_t. □

Eq. (26) shows that the expected oil return in the no-investment region is the interest rate minus the net convenience yield plus the oil risk premium. The oil risk premium in Eq. (27) is proportional to the oil return volatilities and to the market prices of risk. As Eq. (23) shows, a necessary condition for the existence of the oil risk premium is to have risk-averse agents, i.e.,$\gamma > 0$. The oil return volatilities in Eq. (28) are derived from the volatility of the two components of the oil priceJ_K and J_Q.

As with the SDF, the lumpy investment can result in a jump in the price. To see this, consider the behavior of the oil price at the investment boundary z_1 (recall the subscript "1" indicates just prior to investment and "2" indicates just after investment). Substituting $S_t = J_Q/J_K$ into Eqs. (10) and (11) lets us characterize the oil price at the point of investment. Just prior to investment ($z_{t^-} = z_1$), Eq. (10) implies that $S_1 = (\beta_X - \beta_Q)/(1-\beta_K)$. Also, just after investment is made ($z_t = z_2$), Eq. (11) shows that the value of a unit of oil is equal to its marginal cost, that is,$S_2 = \beta_X$. Therefore, at the investment boundary $dS_t/S_t = \Lambda_S$ (or $S_2 - S_1 = \Lambda_S S_1$). Since oil is not a traded financial asset, the jump in the oil price can differ from the price jump in financial assets. Also, if $\beta_Q = \beta_K\beta_X$ then $S_1 = S_2$, i.e., the oil price has no jump. In this case, the fixed cost of the investment at the investment trigger is proportional to aggregate wealth ($K_t^* + S_2 Q_t^*$). For simplicity of exposition, we focus on this case in the SMM calibration that follows.

3.3. Net convenience yield

The convenience yield is defined as the benefit that accrues to the holder of the physical commodity, but not to the owner of a futures contract. The convenience yield is analogous to a dividend yield on a stock (or an interest payment on a bond). However, unlike for financial assets where the dividend is an explicit monetary payment made to the owner of the underlying, for commodities the convenience yield has to be measured implicitly from the difference between long and short futures prices (or equivalently, from the expected return of the spot price under risk-neutrality). This is analogous to calculating the implicit convenience yield from the "cost-of-carry" and the slope of the futures curve as in Routledge et al. (2000). The net convenience yield is the convenience yield net of depreciation and risk adjustments. The net convenience yield can generate backwardation (downward sloping forward curve) and its relation to spot prices can have an important effect on the degree of mean-reversion in oil prices (see Casassus and Collin-Dufresne, 2005).

The following proposition presents the equilibrium cumulative net convenience yield in our economy:

Proposition 2. *The implicit cumulative net convenience yield Y_t has the following dynamics*

$$dY_t = y_t S_t dt + \Lambda_Y S_t dI_t \tag{30}$$

where

$$y_t = \frac{\bar{i}}{S_t}\left(\alpha\,\eta\left(\frac{K_t}{\bar{i}\,Q_t}\right)^{1-\eta} - S_t\right) - \delta - \sigma_Q(\theta_{Q,t} + \rho_{KQ}\theta_{K,t}) \tag{31}$$

$$\Lambda_Y = \Lambda_B - \Lambda_S. \tag{32}$$

Proof. The net convenience yield is determined implicitly from equilibrium prices using the no-arbitrage condition for tradable assets

$$E_t^*\left[\frac{dS_t}{S_t}\right] = \frac{dB_t}{B_t} - \frac{dY_t}{S_t} \tag{33}$$

where E_t^* is the expectation under the equivalent martingale measure. The relation between this expectation and the expectation under physical measure is:

$$E_t\left[\frac{dS_t}{S_t}\right] = E_t^*\left[\frac{dS_t}{S_t}\right] - \frac{d\xi_t}{\xi_t}\frac{dS_t}{S_t}. \tag{34}$$

Applying Itô's Lemma to $S_t = J_Q/J_K$ and using Eq. (34) we can determine dY_t from Eq. (33). □

Eq. (30) shows two components of the cumulative net convenience yield. The first is the absolutely continuous component y_t. This is a function only of z_t, but as before, we prefer to present this variable under $\{K_t, Q_t\}$ to deliver better economic intuition. Eq. (31) gives a clear interpretation of this rate in terms of the convenience yield itself, its physical depreciation and an adjustment for supply shock risk. The first term in this equation represents the benefit for having oil stocks and depends on the difference between the marginal productivity of oil $(\alpha\,\eta(K_t/(\bar{i}Q_t))^{1-\eta})$ and its price (S_t). Because, the rate of oil extraction $\bar{i}$ is fixed, the instantaneous marginal productivity of oil differs from the price of holding one unit of oil in reserves.[15] This also implies that the net convenience yield is time-varying. When marginal productivity is higher (lower) than the oil price, there is a positive (negative) convenience yield to owning oil. Further, this endogenous convenience yield is increasing in the ratio of capital stock to oil reserves, K_t/Q_t. This implies that the convenience yield is higher near the investment region when the crude oil is relatively scarce. The second term in Eq. (31) is the depreciation of oil reserves δ, which is a cost for holding units of oil. Finally, the third term shows that risk (in the form of both supply and technology shocks) tends to lower the net convenience yield.

The second component of the cumulative net convenience yield in Eq. (30) is a singularity, which is due to the predictable jumps that occur in prices at the time of oil investment. If oil were a traded asset then Λ_Y would represent pure arbitrage profits that can be locked in by trading oil prices against any other financial asset. Instead, the commodity is not a financial asset, and its 'price' is the shadow value to consumers of using it as an input to production.

Richard and Sundaresan (1981) presents the isomorphism between convenience yield and interest rates in a multi-good economy. Comparing Eq. (31) with that for the short rate r in Eq. (21) we see a strong resemblance. Effectively, the net convenience yield y can be interpreted as an interest rate in an economy where we switch numeraire and use the commodity instead of the consumption good. In that economy, r_t would become a 'convenience yield' on the consumption good.

3.4. Oil futures prices

Given the stochastic discount factor and the process for the (oil) spot price, we can characterize the behavior of oil futures prices in our model. Let $F_t = F(z, t, T)$ be the date-t futures price in a contract delivering one unit of oil at maturity T.[16] Here the argument z indicates that the futures price depends on the state of the economy summarized by the relative abundance of the oil. The stochastic process for the futures price is written as

$$\frac{dF_t}{F_t} = \mu_{F,t}dt + \sigma_{FK,t}dw_{K,t} + \sigma_{FQ,t}dw_{Q,t} \tag{35}$$

where $\mu_{F,t}$, $\sigma_{FK,t}$ and $\sigma_{FQ,t}$ are determined in equilibrium following Cox et al. (1985) as satisfying the partial differential equation

$$\frac{1}{2}\sigma_z^2 F_{zz} + (\mu_{z,t} + \sigma_K(\lambda_{K,t} + \rho_{KQ}\lambda_{Q,t}) - \sigma_Q(\lambda_{Q,t} + \rho_{KQ}\lambda_{K,t}))F_z + F_t = 0 \tag{36}$$

with boundary condition

$$F(z, T, T) = S(z). \tag{37}$$

Note that futures prices do not have any jump at the oil-investment point, since $F(z_1, t^-, T) = F(z_2, t, T)$.[17]

4. Model estimation

To illustrate the quantitative features and advantages of our general equilibrium model, we undertake three empirical exercises. First we estimate our model using Simulated Method of Moments (SMM). This lets us explore the quantitative properties of our model with a plausible set of parameters. Second, we estimate a two-regime affine approximation of our general equilibrium model where the two regimes capture the key features of the scarcity/abundance of oil. Third, we use our model to estimate the risk premium process associated with oil. In particular, we exploit the theoretical property that the risk premium on an oil futures position depends on the relative scarcity of oil.

4.1. Moments and SMM estimation

In order to see if our model can capture the interesting features of oil prices, we calibrate our model using SMM (Lee and Ingram, 1991 and Duffie and Singleton, 1993). We focus on matching the unconditional mean level of the term-structure of futures prices and their unconditional volatility from NYMEX prices.[18] Since oil plays an important role in our economy we also try to match the average consumption to GDP ratio, the average oil consumption to GDP ratio, and the average real interest rate.[19] We use quarterly data and two different time periods for the SMM estimation: Panel A (pre-Iraq war) from Q4/1990 to Q1/2003 and Panel B (whole sample) from Q4/1990 to Q2/2008. We do this because since the Iraq war average crude oil prices have risen dramatically and have become more volatile. The unconditional volatility of long-maturity futures prices has also increased compared to the beginning of the sample period. The data sources and construction are listed in Appendix C.

The unconditional moments in our data are sample averages, $G_T = \frac{1}{T}\sum_{t=1}^{T} g_t$. Denote the analogous moments in our model as $G_Z(\psi)$ where ψ are the parameters we are estimating. Given a particular set of parameters, $\hat{\psi}$, we solve the HJB equation in (B5) with the numerical technique described in Appendix D. This determines the endogenous dynamics of our state variable, z_t. To approximate the density function of z, $f(z;\hat{\psi})$ we simulate the model.[20] Using this simulated density, we calculate the moments implied by our model as $G_Z(\hat{\psi}) = E^Z[g(z;\hat{\psi})] \approx \int g(z)f(z)dz$. We choose parameters to solve

$$\psi^* = \underset{\psi\in\Psi}{\operatorname{argmin}}[G_Z(\psi) - G_T]'W_T[G_Z(\psi) - G_T] \tag{38}$$

where Ψ is the set of feasible parameters (i.e., where the model is well-defined). W_T is the weighting or distance matrix. We choose W_T to be the inverse of the diagonal of the unbiased estimate covariance matrix of the sample averages. This weighting matrix ensures that the scale of each moment condition is the same, and gives more weight to less volatile moments.[21]

---

[15] Indeed, if the representative agent could flexibly choose $\bar{i}$, then it will be such that the marginal productivity of oil equals the oil price, making this component zero.

[16] Since the futures contracts are continuously marked-to-market, the value of the futures contract is zero.

[17] This condition derives directly from the fact that futures prices are martingales under the risk-neutral measure. No arbitrage implies that if at maturity $S(z_T) = S_2$, then $F(z_1, T^-, T) = F(z_2, T, T) = S_2$ where T^- is time just before the contract expires. However, if $S(z_{T^-}) = S_1$, then $F(z_1, T^-, T^-) = S_1$. Note that even if there's no jump in the futures price dynamics, there could be a price gap across futures contracts with near maturities, i.e. $F(z_1, T^-, T^-) = S_1$ and $F(z_1, T^-, T) = S_2$.

[18] The discussion about the existence of mean-reversion in oil prices is always controversial. Nevertheless, we make this assumption which justifies the usage of data on futures prices instead of returns. This assumption is also consistent with the prediction of the equilibrium model.

[19] For the calibration we use petroleum consumption as a proxy for oil consumption, even though the petroleum category includes also other products made from coal, natural gas and biomass.

[20] We discretize the state space of $z_t \in [-20, 10]$ in a grid of 30,000 points and then simulate weekly samples of the state variable for 10^5 years.

[21] Cochrane (2005) discusses the pros and cons of using different weighting matrices for the estimation.

To overcome the empirical difficulties during the computationally intensive SMM estimation, we fix five (out of twelve) parameters based on earlier studies. The oil share of income, η, is set to 0.04 which is consistent with recent RBC studies that include energy as a production factor (see Finn, 1995; Finn, 2000 and Wei, 2003). The depreciation rate for oil reserves, δ, is set to 0.1. This is consistent with a storage cost of approximately $4 per barrel per year (in line with Ross, 1997). Regarding the volatility parameters, we fix σ_Q and ρ_{KQ} and leave σ_K free.[22] We fix the volatility of oil stock using the deviation of annual changes of petroleum consumption in our dataset (i.e.,$\sigma_Q = 0.013$), a figure that captures the low variability in this sector. We also assume that the "unexpected" shocks to capital and oil stocks are independent because of the fairly low correlation between the aggregate stocks in both sectors ($\rho_{KQ} = 0$). Lastly, we set the rate of time preference parameter to $\rho = 0.05$.

The cost structure for the production of new oil is central to our model. We estimate both, the variable cost of oil,β_X, and the fixed cost component β_K. We choose the parameter β_Q so that the total fixed cost is proportional to aggregate wealth in the economy, which implies that the oil price is continuous even at the investment boundary. We also estimate the four remaining parameters: the capital productivity, α, the rate of oil extraction,$\bar{i}$, the volatility of the capital sector,σ_K, and the curvature of the representative agent's utility function, γ.

Table 1 reports the SMM estimates for both panels of data. These estimates all seem reasonable and in line with other RBC models. The one notable parameter estimate is the relatively low estimate of risk-aversion close to 0.5.[23] This low estimate of γ is driven by the low real interest rates in the data. The estimation results show that the volatility of capital,σ_K, the marginal cost,β_X, and the extraction rate of oil,$\bar{i}$, increase when considering the whole time period. The volatility of capital increases from 0.39 (Panel A) to 0.42 (Panel B) to match the higher volatility of crude oil prices in recent times. The marginal cost of oil production increases from $18.1 (Panel A) to $26.9 (Panel B). A higher marginal cost yields higher average crude oil prices. This is consistent with the recent entry of low-yield wells that produce with higher costs.[24] Finally, a higher demand rate of oil is necessary to match a higher oil consumption to GDP ratio that arises because oil consumption is measured in dollars.

Table 2 shows the historical and model implied mean term structure of futures prices, mean futures volatilities and mean macroeconomic variables for both time periods (Panels A and B). The average futures curve over is downward sloping in both cases.

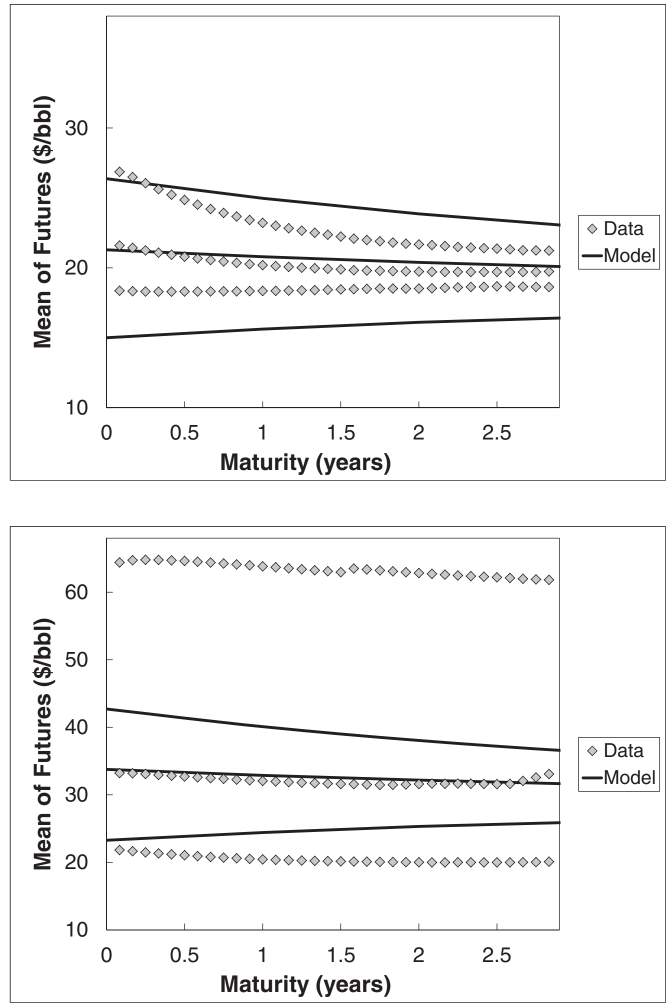


Fig. 1. Historical and model implied moments of futures curves for the periods Q4/1990 to Q1/2003 (top) and Q4/1990 to Q2/2008 (bottom). Each set of curves has three different moments: the unconditional mean, and two conditional means depending whether the spot price is below or above the average spot price.

This "backwardation" is a common feature of oil prices. For example, in the whole sample period, 63% of the time the 6-months maturity contract is below the 1-month maturity contract. The model generates a similar average degree of backwardation than in the data. The unconditional volatility is also decreasing in the horizon for both time periods reflecting the high degree of mean reversion in the oil price. Our model also captures this feature of the data reasonably well, although it is unable to replicate the observed higher short-term volatility.[25] Our model also implies sensible values for consumption to GDP, oil consumption to GDP, and the real interest rate. We match the mean values of these items almost exactly and closely match the standard deviation of these quantities.

Fig. 1 shows the historical and model implied mean futures curve for the periods Q4/1990 to Q1/2003 (top) and Q4/1990 to Q2/2008 (bottom). As reported above, the model correctly fits the unconditional futures curves. The figure also shows the average curves conditional on whether the spot price is below or above the average spot price. These conditional curves are not explicitly included as moments for the SMM estimation. However, the model

[22] In data, it is hard to separately observe shocks to the oil-stock and the capital stock. Also, the key parameter that drives the variability of our variables of interest (e.g., oil price, convenience yield, consumption rate, interest rates, risk associated with the irreversible investment) is $\sigma_z = \sqrt{\sigma_K^2 - 2\rho_{KQ}\sigma_K\sigma_Q + \sigma_Q^2}$. Therefore, we could have fixed any two of σ_K, σ_Q and ρ_{KQ} without changing the volatility structure of these interest (this is not necessarily true for their levels).

[23] The estimation of the risk aversion parameter has always been controversial. Lucas (2003) states that values of γ close to one are typically used in economic applications and suggests an upper bound of 2.5. On the other hand, Mehra and Prescott (1985) show that to account for the entire equity premium,γ has to be close 40. A natural solution to this controversy is to use a more general utility function, i.e., a recursive utility specification (see Epstein and Zin, 1989; Duffie and Epstein, 1992 and Backus et al. (2005)). Indeed, in the CRRA framework we employ here, the risk-aversion parameter is the only curvature in the utility specification and reflects attitudes to both risk and time aggregation. We leave a recursive utility model to future research.

[24] The *World Economic Outlook* from the IMF (2008) reports the annual marginal cost of producing a barrel of oil from 1991 to 2007. This cost is defined as the average of the highest-cost (or bottom quartile) producers, based on a survey of listed oil companies. In recent years, the marginal cost increased steadily from $21.4 in 2002 to $74.2 in 2007. Moreover, the average marginal cost for the 1991–2002 period was $18.8, while for the 1991–2007 period was $28.5. These figures are in line with our SMM estimates.

[25] This implies also that the oil spot return volatility is underestimated. A natural way of fixing this would be to increase both the volatility of the state variable and the degree of mean reversion in the economy, however, this would harm the macroeconomic moments.

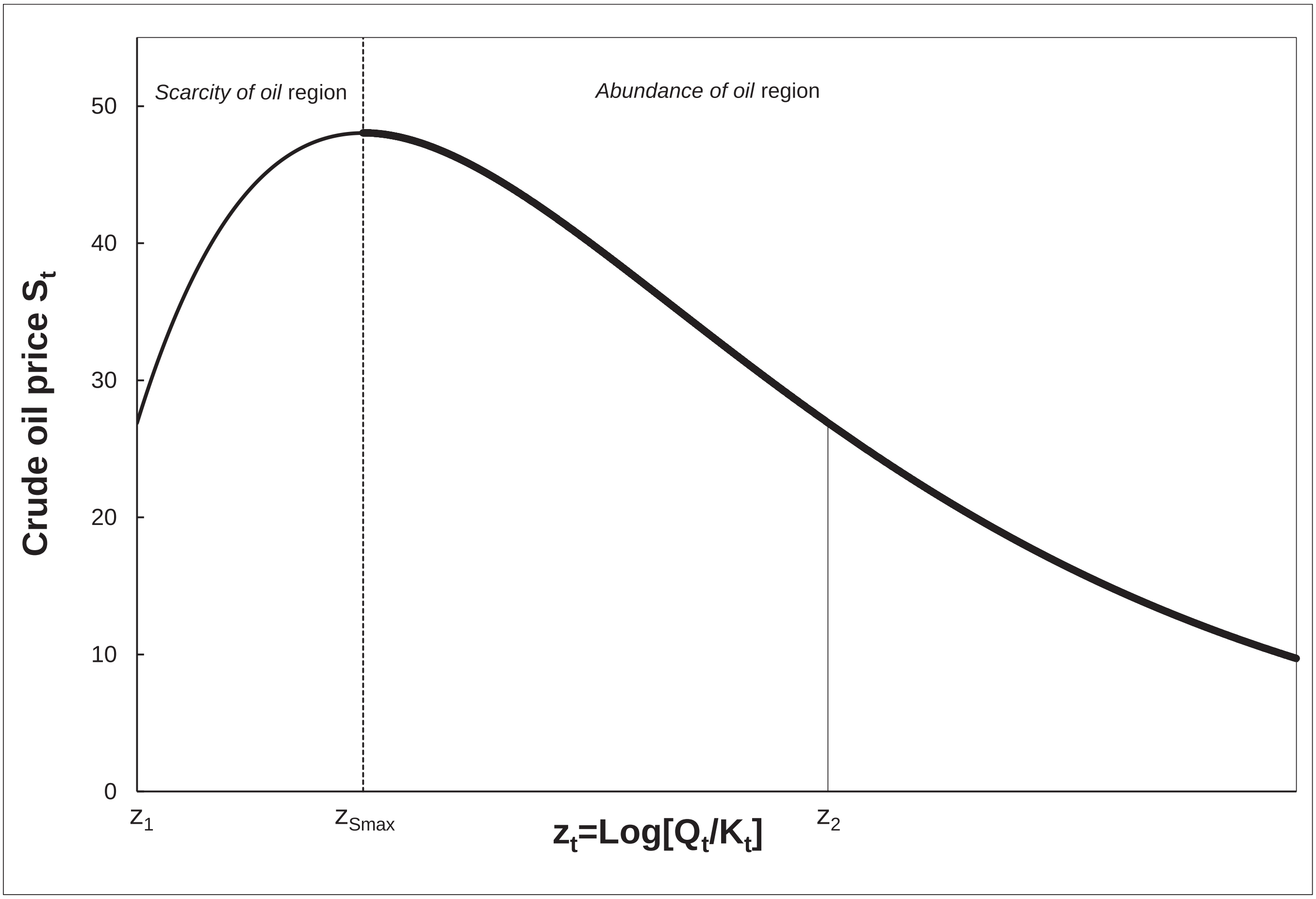


Fig. 2. Oil price S_t as a function of the logarithm of the oil reserves-capital ratio, z_t. The vertical dashed-line is at $z_{_{Smax}}$ and separates two regions. The thin line shows the oil price in the scarcity region ($z_1 < z_t \leq z_{_{Smax}}$) and the thick line is the oil price in the abundance region ($z_t \geq z_{_{Smax}}$). We use the parameters from Panel B in Table 1. In particular, the fixed cost components of the investment are $\beta_{_K} = 0.009$ and $\beta_{_Q} = 0.242$, and the marginal cost of oil is $\beta_{_X} = \$26.9$. The equilibrium critical ratios are $z_1 = -8.91$, $z_{_{Smax}} = -8.269$ and $z_2 = -6.89$.

is able to generate similar conditional curves for the Q4/1990-Q1/2003 period (data vs. model). Indeed, the slope and level of the implied conditional curves are closely related to the ones in the data. For the whole time period, the model fails to replicate the conditional curves. This is especially true for the futures curve conditional on high prices where in the data, the curve is above \$60, while in the model it is around \$40.[26] Nevertheless, the model captures the higher unconditional volatility for the long-maturity futures, which is consistent with the observed data. To see this, compare the long-term spreads between the conditional curves for both sample periods.

Overall, with sensible model parameters, we match the term-structure of futures prices and volatilities reasonably well. The model also produces reasonable quantities for output, oil consumption, and real interest rates. Therefore, we use this parametrization to explore the quantitative implications of the model.

4.1.1. Oil spot price

Fig. 2 plots the equilibrium oil spot price against the state variable,z_t, as in Eq. (24). Recall that z_t, the log of the oil reserves to capital ratio, is a measure of the relative scarcity of oil. Interestingly, the oil price is non-monotonic in the state variable and reaches a maximum level S_{max} that is greater than the marginal extraction cost β_X. The equilibrium maximum price is reached at z_{Smax}, which lies between the investment threshold z_1 and the replenishment level z_2 when there is a fixed cost component ($\beta_K, \beta_Q > 0$).[27] This happens because the oil price is driven by both current and anticipated changes in oil reserves. Since oil investment is periodic, as oil is depleted, the stock of oil falls and this is reflected in a lower value of z_t. When oil is plentiful ($z_t \geq z_{Smax}$), the decreased stock of oil leads to an increase in the spot price of oil. We define this zone as the *abundance region*, because there is a relative abundance of oil in the economy. The fixed cost involved in adding new oil stocks implies that it is not optimal to make a new investment as soon as the spot price (marginal benefit of oil) reaches the marginal cost of adding new oil,β_X. Therefore the spot price rises above β_X as oil is depleted. However, closer to the investment threshold, the oil price reflects the likelihood of a near-term lumpy investment in new oil. In this zone, as oil is depleted and z_t decreases, the probability of hitting the investment increases and the oil price decreases in anticipation of investment in new oil reserves. This defines the *scarcity region* due to the relative scarcity of oil stocks. Lastly, note that in this parametrization with a fixed cost that is proportional to aggre-

[26] This occurs because for the estimated parameters, the model is unable to reach crude oil prices above \$50 (as we will see later, the price has an upper bound at \$48.04.

[27] The maximum price is endogenously determined. A necessary condition for it's existence is to have $S_1 \leq S_2$, because the price function is continuous and it's derivative is negative at z_2.

Table 1
The table reports the SMM estimates for the sample periods Q4/1990 to Q1/2003 (Panel A) and Q4/1990 to Q2/2008 (Panel B). The parameters marked with an asterisk (*) are estimated. The Q-component of the fixed cost, denoted with (†), was set to $\beta_Q = \beta_K \beta_X$.

Production technologies		Panel A: Q4/1990 to Q1/2003	Panel B: Q4/1990 to Q2/2008
Productivity of capital K (*)	α	0.131	0.138
Importance of oil	η	0.04	0.04
Demand rate for oil (*)	$\bar{i}$	0.061	0.091
Volatility of capital (*)	σ_K	0.390	0.420
Volatility of oil stocks	σ_Q	0.013	0.013
Correlation of capital and oil shocks	ρ_{KQ}	0	0
Depreciation of oil	δ	0.10	0.10
Irreversible investment			
Fixed cost (K component) (*)	β_K	0.007	0.009
Fixed cost (Q component) (†)	β_Q	0.132	0.242
Marginal cost of oil (*)	β_X	18.1	26.9
Agents preferences			
Patience	ρ	0.05	0.05
Risk aversion (*)	γ	0.50	0.51

Table 2
The table presents the historical and the model implied moments using the corresponding parameter estimates from Table 1.

Moment conditions	Historical data		Model	
	Sample average	Sample SD	Uncond. Mean	Uncond. SD
Panel A: Q4/1990 to Q1/2003				
Futures prices - 01	21.66	4.93	21.26	6.65
Futures prices - 03	21.28	4.42	21.17	6.47
Futures prices - 06	20.84	3.84	21.04	6.22
Futures prices - 09	20.50	3.41	20.92	5.98
Futures prices - 12	20.25	3.08	20.80	5.76
Futures prices - 18	19.95	2.58	20.59	5.36
Futures prices - 24	19.84	2.28	20.40	5.01
Futures prices - 30	19.75	2.03	20.23	4.71
Futures prices - 36	19.93	1.98	20.07	4.44
Volatility - 01	0.312	0.106	0.199	0.097
Volatility - 03	0.273	0.092	0.196	0.092
Volatility - 06	0.225	0.081	0.190	0.088
Volatility - 09	0.191	0.075	0.183	0.085
Volatility - 12	0.170	0.068	0.176	0.084
Volatility - 18	0.142	0.057	0.163	0.082
Volatility - 24	0.130	0.049	0.150	0.081
Volatility - 30	0.126	0.047	0.139	0.080
Volatility - 36	0.128	0.044	0.129	0.079
Consumption/GDP	0.615	0.008	0.617	0.022
(Petroleum consumption)/GDP	0.016	0.003	0.016	0.005
Real interest rates	0.017	0.022	0.010	0.002
Panel B: Q4/1990 to Q2/2008				
Futures prices - 01	33.81	24.03	33.70	11.28
Futures prices - 03	33.65	24.31	33.53	10.95
Futures prices - 06	33.30	24.45	33.30	10.47
Futures prices - 09	32.94	24.45	33.09	10.03
Futures prices - 12	32.64	24.37	32.89	9.61
Futures prices - 18	32.18	24.12	32.53	8.88
Futures prices - 24	31.87	23.81	32.20	8.24
Futures prices - 30	31.58	23.54	31.90	7.69
Futures prices - 36	31.68	23.35	31.62	7.21
Volatility - 01	0.309	0.093	0.223	0.112
Volatility - 03	0.273	0.079	0.219	0.105
Volatility - 06	0.231	0.072	0.212	0.098
Volatility - 09	0.202	0.069	0.205	0.094
Volatility - 12	0.182	0.066	0.197	0.092
Volatility - 18	0.157	0.063	0.181	0.089
Volatility - 24	0.146	0.060	0.167	0.087
Volatility - 30	0.143	0.060	0.154	0.086
Volatility - 36	0.145	0.058	0.142	0.084
Consumption/GDP	0.618	0.008	0.618	0.023
(Petroleum consumption)/GDP	0.020	0.009	0.022	0.008
Real interest rates	0.011	0.031	-0.001	0.002

Table 3
The table reports the quasi-maximum likelihood (QML) estimates for the regime-switching model in Eqs. (41)–(44) for weekly deflated crude oil prices.

Parameter	Estimate	Std.Error	t-ratio
Panel A: Apr/1983 to Mar/2003			
λ_1	1.132	0.392	2.9
κ_1	0.293	0.146	2.0
σ_1	0.289	0.017	16.7
λ_2	6.347	2.263	2.8
κ_2	0.126	0.591	0.2
σ_2	−0.898	0.101	−8.9
α	−0.173	0.107	−1.6
$\log S_{max}$	4.222	0.050	84.9
Panel B: Apr/1983 to Jun/2008			
λ_1	1.113	0.360	3.1
κ_1	0.019	0.135	0.1
σ_1	0.213	0.010	21.5
λ_2	2.494	0.959	2.6
κ_2	0.082	0.224	0.4
σ_2	−0.496	0.033	−15.2
α	0.049	0.178	0.3
$\log S_{max}$	4.831	0.037	130.9

gate wealth, the oil price is unchanged at the point of investment, i.e.,$S(z_1) = S(z_2)$.

It is important to note that the hump-shaped response of the oil price to the state variable that determines the scarcity/abundance regions is exclusively due to the fixed cost component of the investment in new oil units. Under the absence of these fixed costs (i.e.,$\beta_K = \beta_Q = 0$), investment would be infinitesimal, implying $z_1 = z_2$. In addition, investment would be made as soon as the price reaches the marginal cost and would be large enough to keep the price capped at β_X, causing the maximum oil price to be at the investment boundary. In this case, the price of oil would be only a monotonic function of the state variable, i.e., there would be no hump in the oil price function. The fixed cost component is necessary not only to create significant variation in the crude oil price, but also to generate the two regimes in the economy.

Fig. 2 shows also that the oil price cannot be written as a one-dimensional Markov process, despite that fact that it is a function of z_t, which itself follows a one-dimensional Markov process. However, the figure suggests that the price process can be made Markov if, in addition to the price level S_t we also know a discrete state variable, say ε_t, a proxy for the relative scarcity/abundance regime (i.e., the fact that z_t is less or greater than z_{Smax}). We formalize this observation in the proposition below, which will be useful for our next empirical exercise, where we model the time-series of oil prices as a regime switching model (as we do not observe the state variable z_t).

Proposition 3. *The oil price in Eq. (24) is governed by the following two-regime stochastic process*

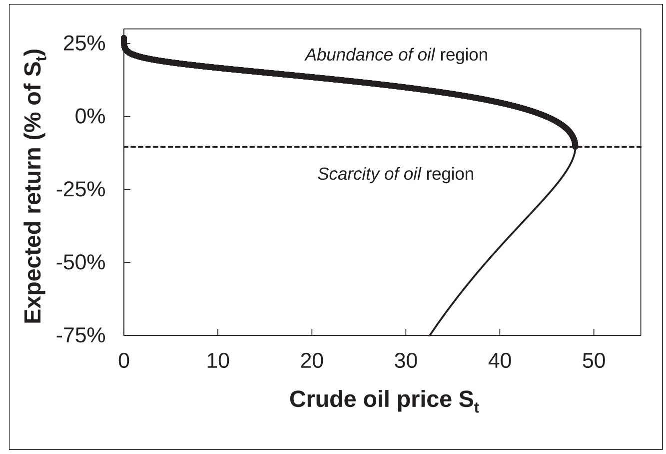
$$\frac{dS_t}{S_t} = \mu_S(S_t, \varepsilon_t)dt + \sigma_{SK}(S_t, \varepsilon_t)dw_{K,t} + \sigma_{SQ}(S_t, \varepsilon_t)dw_{Q,t} + \Lambda_S dI_t \tag{39}$$

where

$$\varepsilon = \begin{cases} 1 & \text{if} \quad z > z_{Smax} \\ 2 & \text{if} \quad z_1 < z \le z_{Smax} \end{cases} \tag{40}$$

and where $\mu_S(S_t, \varepsilon_t) = \mu_{S,t}$, $\sigma_{SK}(S_t, \varepsilon_t) = \sigma_{SK,t}$, $\sigma_{SQ}(S_t, \varepsilon_t) = \sigma_{SQ,t}$ *and* Λ_S *is from Eq. (29).*

Proof. To show that the dynamics of S_t depends only on $\{S_t, \varepsilon_t\}$, we note that there is a one-to-one mapping between $\{S_t, \varepsilon_t\}$ and z_t. Since all the variables in the proposition are only a function of z_t, they can also be expressed in terms of $\{S_t, \varepsilon_t\}$. □



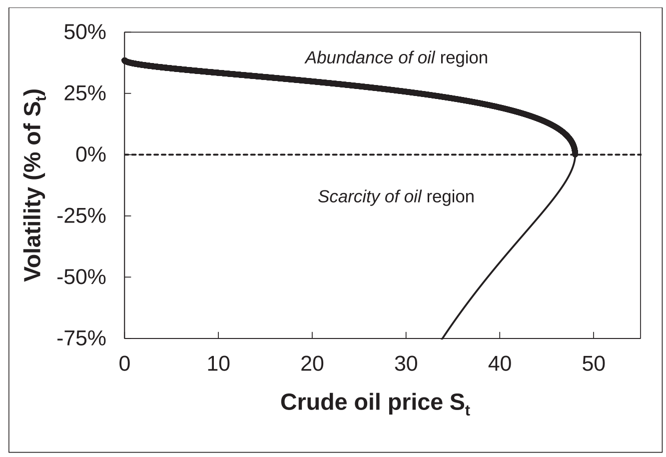


Fig. 3. Expected return and instantaneous volatility of returns in oil price S_t. The horizontal dashed-line separates the two regimes. The thin lines below the dashed-lines show the variables under the scarcity regime and the thick lines under the abundance regime. We use the parameters from Panel B in Table 1. In particular, the fixed cost components of the investment are $\beta_K = 0.009$ and $\beta_Q = 0.242$, and the marginal cost of oil is $\beta_X = \$26.9$. The endogenous upper bound for the price is $S_{max} = \$48.04$.

Fig. 3 plots the conditional instantaneous return and conditional instantaneous volatility of return as a function of S_t. As you would expect from Fig. 2, the drift and volatility differ across the scarcity- and abundance-of-oil regions. Since the oil price is bounded by 0 and S_{max}, it is mean reverting. This can be observed in the top panel of Fig. 3 where the drift is positive for lower prices and mostly negative for higher prices. However, the figure also shows that the rate of mean reversion is much higher in the scarcity region. The bottom panel of Fig. 3 shows that the volatility follows a similar non-linear pattern than the drift. Conditional on the region and in absolute terms, the volatility is decreasing with the oil price. The negative sign of the volatility in the scarcity region implies a negative correlation with the state variable z_t. This is directly important to the risk premium properties of oil futures prices (we will return to this topic below). Lastly, note that the level of the volatility depends on the level of the spot price relative to the maximum spot price. At S_{max} the volatility is zero since the price will decrease almost surely.

4.1.2. Net convenience yield and oil risk premium

Proposition 1 shows that the expected oil spot return can be decomposed into the interest rate, the net convenience yield and the oil risk premium. Here we analyze the two oil-specific components.

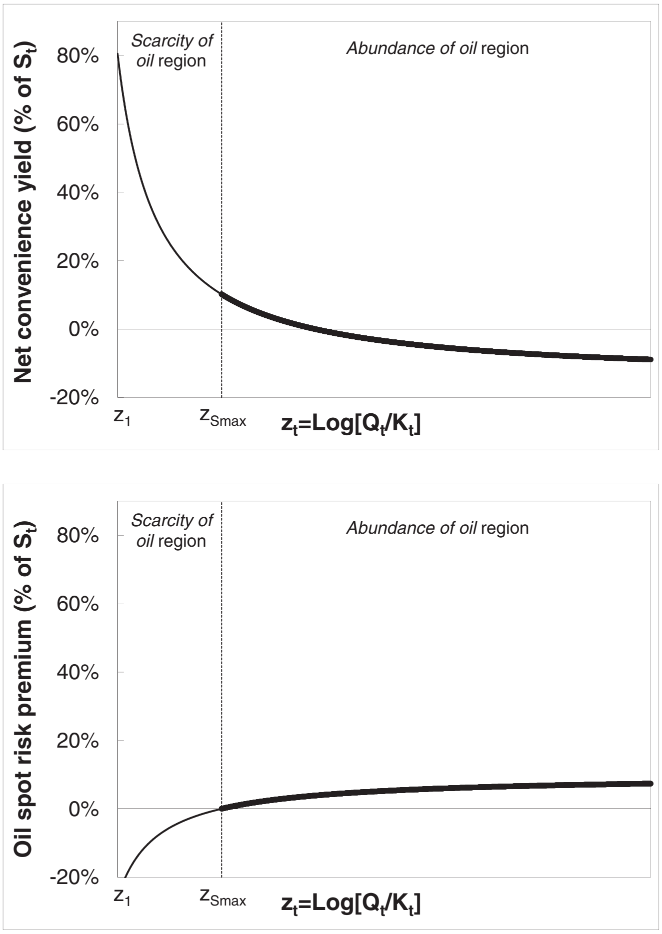


Fig. 4. Net convenience yield (top) and oil spot risk premium (bottom) as functions of the state variable z_t. The thick line is the net convenience yield when the economy is in the abundance region and the thin for the economy in the scarcity region. We use the parameters from Panel B in Table 1 and the endogenous upper bound for the price is $S_{max} = \$48.04$.

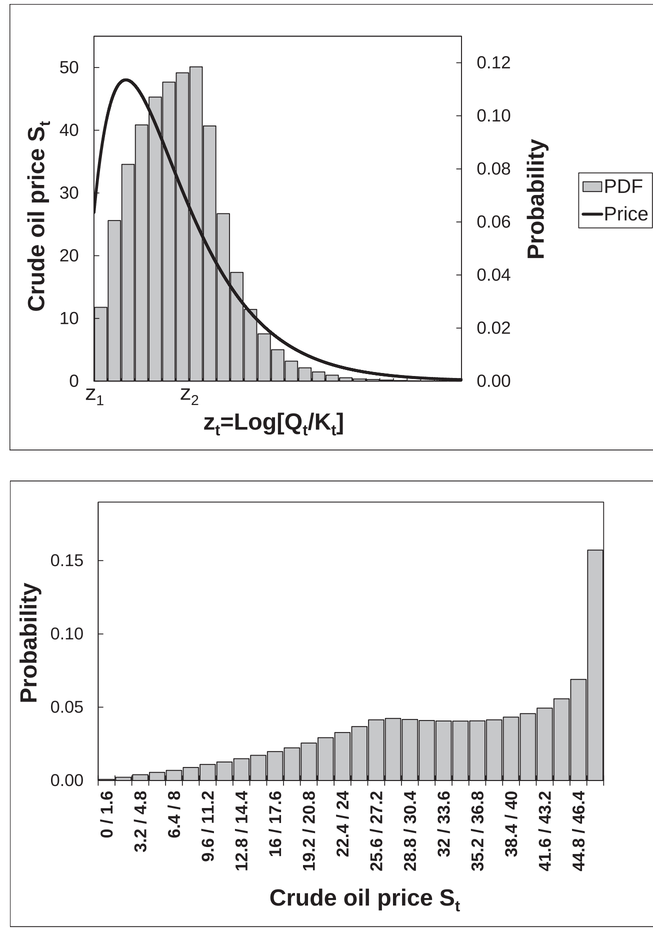


Fig. 5. Simulated probability density function for the state variable z_t and the commodity price S_t using the parameters from Panel B in Table 1.

The top panel of Fig. 4 plots the net convenience yield as a function of the state variable z_t. As anticipated, the convenience yield is higher in the scarcity region, because of the higher marginal productivity of oil. In this region the representative agent would like to demand oil at a higher rate, but she is unable to do so because of the fixed rate $\bar{i}$. In the scarcity region inventories are low and a future oil supply shortfall is more likely. A relation between these fundamentals and the convenience yield is a known result from the literature (e.g., Pindyck, 2001 and Alquist and Kilian, 2010). The convenience yield decreases in the abundance region even to negative values because of the depreciation in the oil sector and a lower productivity of the commodity. As we will see later, the futures curves are steeper when the spot price is in the scarcity region, because the net convenience yield is higher in this region.

Our model also replicates a positive relation between the convenience yield and spot prices, another fact documented in the convenience yield literature and the traditional arbitrage pricing approach.[28] Moreover, using the simulated unconditional distribution for the state variable in top panel of Fig. 5, we can calculate the correlation between these variables as well as the conditional expected prices for each region. Given our parametrization, the correlation between the net convenience yield and the oil price is 0.51. This positive correlation can also be seen by the fact that the expected spot price in the scarcity region (high convenience yield) is \$43.47 while this expectation in the abundance region (low convenience yield) is \$32.01.

The bottom panel of Fig. 4 plots the risk premium on oil as a function of the oil-to-capital state variable. The risk premium is state-dependent, it is positive in the abundance region and negative in the scarcity region. Recall, from Fig. 3 (lower panel) that the sign of the volatility of the spot price differs across the scarcity/abundance regions. This implies that the correlation between shocks to the consumption good and the spot price of oil also changes sign across the regions generating a change in the sign of the risk premium.

4.1.3. Oil futures prices and backwardation

Using the estimated parameters of our model, we can solve for futures prices. Sample futures prices are plotted in Fig. 6. As above, we can characterize the futures prices in terms of the current spot price and the scarcity/abundance region. The thick futures curves indicate the abundance region and the thin lines indicate the scarcity. Broadly, the futures prices in our model capture the futures prices in the data (see Table 2). When the oil price is low, oil is plentiful and the state variable is far from the investment trigger. This means that the supply of oil decreases on average, so the expected price in the future is above the current

[28] For the latter see the reduced-form commodity pricing models of Gibson and Schwartz (1990), Schwartz (1997) and Casassus and Collin-Dufresne (2005), among others.

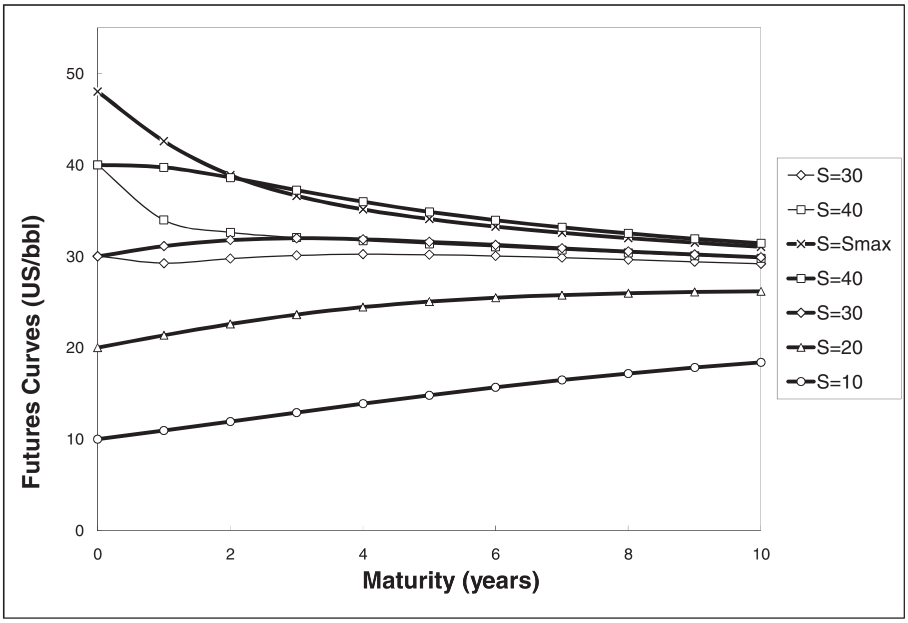


Fig. 6. Futures curves for contracts on oil for different spot prices. The thick curves are for spot prices in the abundance region and the thin lines when the spot price is in the scarcity region. We use the parameters from Panel B in Table 1 and the endogenous upper bound for the price is $S_{max} = \$48.04$.

price.[29] In these situations the futures curves are upward-sloping or in contango.

When the spot price of oil is high, the futures curve is downward sloping or in backwardation since the (risk-adjusted) expected change in spot price is negative. More interestingly for intermediate prices, the futures curve can be hump-shaped. These hump-shaped curves highlight that the spot price is not sufficient to characterize the futures curve. The futures curve are steeper in the scarcity region. This is an implication of the high likelihood of a near-term investment in new oil in this zone. For example, in Fig. 6 consider the curve with a current spot price $S_t = 30$ in the scarcity region (thin line). In the near term, there is an expected drop in the spot price due to the likely new investment. This is followed by an average price increase as the new oil is depleted. In contrast, a current spot price of $S_t = 30$ but in the abundance region (thick line), reflects an anticipated increase in the spot price as oil is depleted. Possible new investment in oil influences the longer horizon prices.

The interesting features of the spot and futures prices all stem from the property that the dynamic properties of the oil price are different in the scarcity and abundance regions. This is only relevant empirically if the (endogenous) state variable has a distribution that accesses these regions frequently enough to matter. The fact that our model matches the unconditional average pattern of futures prices (see Table 2) is encouraging. However, given our parametrization we can directly calculate the relative frequency. The top panel of Fig. 5 plots the unconditional distribution for the state variable,z_t. The mapping from z_t to the spot price is overlaid to highlight that both, the scarcity and the abundance regions, are visited frequently. The economy is in the scarcity region 14% of the time and in the abundance region 86% of the time. The bottom panel of Fig. 5 show the related distribution for the spot price. The spot price is frequently near its maximum. This implies that, consistent with the data, futures prices in our model are often downward sloping in our calibration.

4.1.4. Oil futures risk premium and volatility

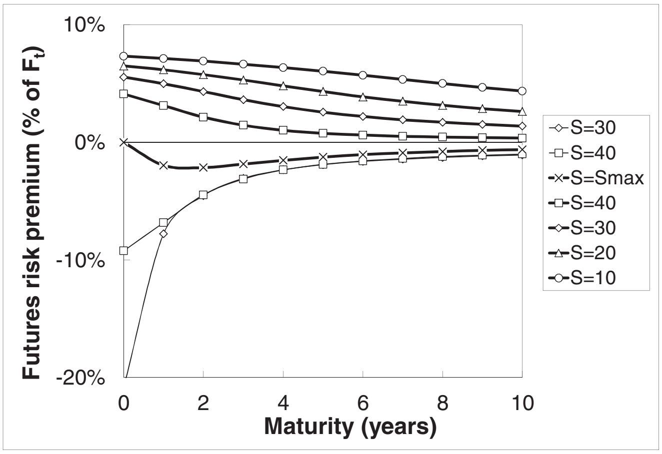
The top panel of Fig. 7 shows the futures risk premium for different spot prices, states and maturities of the contracts. Since oil futures prices are martingales under equivalent martingale measure, the futures risk premium are the expected futures returns. Again, we observe that these premia are positive when the economy is in the abundance region and negative in the scarcity region. In the scarcity region, a long position in a near-term oil futures hedges negative shocks to the consumption good. A negative shock to the stock of the consumption good in the scarcity region implies an increased oil price (and an increased price in near-term futures prices) since the relative stocks of commodity decreases. In the abundance region, the opposite is true and a long-position in futures earns a positive risk premium. When the spot price is S_{max} the shortest maturity contract has no risk premium, because the instantaneous volatility is zero at this point. For longer maturities, the futures risk premia are negative because the economy is expected to be in the scarcity region.

In general, the futures risk premia tend to decrease (in absolute terms) with the contract's maturity, because the volatility of the futures decreases with maturity (Fig. 7 bottom panel). If we plot the size of the futures risk premium relative to the corresponding spot risk premium we will get the same figure as the one with the relative volatilities. This occurs because the Sharpe ratio is the same across maturities of the contracts.[30]

The volatility of the futures contract are shown in the bottom panel of Fig. 7. To compare the futures volatility for different oil spot prices we show the *relative volatility* which we define as $\sigma_F(S_t, \varepsilon_t; T-t)/\sigma_S(S_t, \varepsilon_t)$. This ratio corresponds to the inverse of the optimal *hedge ratio*, which is the number of futures contracts in a portfolio that minimizes the risk exposure of one unit of oil. Since the futures price with zero maturity is the spot price, this ratio is 1 when $t = T$.

[29] For intuition, we discuss futures shape in terms of expected future prices, but this is not the full story in this model where risk premia play a central role (and futures prices are risk-adjusted expected future spot prices). In fact, it is straightforward to interpret the properties of the futures curve slope in terms of the net convenience yield.

[30] Indeed, since $\mu_{F,t,T1}/\sigma_{F,t,T1} = \mu_{F,t,T2}/\sigma_{F,t,T2}$, the relative risk premia are the same as the relative volatilities, i.e.,$\mu_{F,t,T1}/\mu_{F,t,T2} = \sigma_{F,t,T1}/\sigma_{F,t,T2}$.



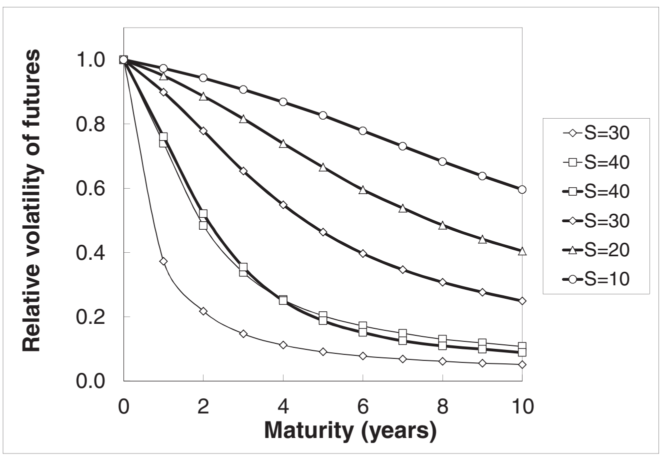


Fig. 7. Futures risk premium (top) and relative volatility of futures contracts on oil to spot price volatility (bottom) for different spot prices and maturities. The thick curves are for spot prices in the abundance region and the thin lines when the spot price is in the scarcity region. We use the parameters from Panel B in Table 1 and the endogenous upper bound for the price is $S_{max} = \$48.04$.

As above, we can characterize the volatilities in terms of the current spot price and the scarcity/abundance region. The thick futures curves indicate the abundance region and the thin lines indicate the scarcity region. In general, the volatilities are much lower for higher maturities. This is a consequence of mean reversion in spot prices. The figure highlights that the hedge-ratio depends on both the level of the spot price and the scarcity/abundance region. In contrast, many commodity price models impose log-linear relation between spot and futures prices that implies a hedge ratio that is independent of price (e.g., Schwartz, 1997). Second, the curves are non-monotonic in the maturity horizon. For high prices, the expected investment in oil (rise in supply) is reflected in the futures contract and also in the volatility. For short maturities and very high prices the *relative volatility* has an abrupt behavior because the volatility of the spot price is very low (recall that $\sigma_S(S_{max}, \varepsilon_t) = 0$). Eventually, the relative volatility could be negative at longer horizons. This reflects the expected transition of the state-variable across the two regimes where the spot price volatilities change sign.

4.2. Regime-switching estimation

While our structural model captures many interesting features of the data, it is challenging to compute. In comparison, commodity pricing models in practice are often based on an exogenous specification of a spot price and other factors typically related to the slope of a futures curve (e.g., the convenience yield). These two-factor models are flexible enough to capture many salient features of the data while remaining tractable. Our second empirical exercise links our general equilibrium model to the more tractable reduced-form contingent claims approach.

Following the predictions of our structural model, we propose the following two-factor model

$$dS_t = \mu_s(S_t, \varepsilon_t)S_t dt + \sigma_S(S_t, \varepsilon_t)S_t dw_{s,t} \tag{41}$$

where

$$\mu_s(S, \varepsilon) = \alpha + \kappa_\varepsilon(\log[S_{max}] - \log[S]) \tag{42}$$

$$\sigma_S(S, \varepsilon) = \sigma_\varepsilon\sqrt{\log[S_{max}] - \log[S]} \tag{43}$$

and ε_t is a two-state Markov chain with transition (Poisson) probabilities

$$P_t = \begin{bmatrix} 1-\lambda_1 dt & \lambda_1 dt \\ \lambda_2 dt & 1-\lambda_2 dt \end{bmatrix} \tag{44}$$

As seen in Proposition 3 and Fig. 3, the stochastic process for the equilibrium spot price differs across the scarcity and abundance regions. However, within those regions a linear approximation to the drift and volatility functions works quite well. Hence we can estimate a two-factor model where the second factor is an indicator for the scarcity/abundance regions, although we do not impose a priori which are these two states.[31] An important characteristic of our reduced-form model is that since the stochastic process is, up to the two-regimes, affine, one can apply standard derivative pricing techniques.

To estimate the model we use the quasi-maximum likelihood technique of Hamilton (1989) for two-regime specifications. A by-product of the estimation technique are the smoothed inferences for each regime.[32] The data here consists of weekly crude oil prices for two different time periods: Panel A (pre-Iraq war) from Apr/1983 to Mar/2003 and Panel B (whole sample) from Apr/1983 to Jun/2008. We deflate oil prices by the US Consumer Price Index. The average price is 16.29 dollars per barrel in 1983 prices (or 31.53 dollars per barrel in 2005 prices). When we consider the whole sample period, we find that the annualized standard deviation of weekly returns is 37%, the skewness in crude oil returns is −0.05 and their excess kurtosis is 0.19.

The parameter estimates and standard errors of our model for both time periods are given in Table 3. Most parameters in Panel A are significant and vary across regimes implying that there are clearly two different regimes in the data. The parameter $\alpha = -0.173$ is negative and somewhat significant implying that the process for the price has an upper bound at S_{max} which is estimated at \$68.17 per barrel ($\log S_{max} = 4.222$). Although we do not impose this directly in the estimation, the price dynamics in regime 1 is similar to the price dynamics in the abundance region in the general equilibrium model. A similar thing happens with the dynamics in regime 2 and in the scarcity region in the equilibrium model. Consistently, regime 1 (abundance) has a higher degree of mean reversion ($\kappa_1 > \kappa_2$) and a lower volatility ($|\sigma_1| < |\sigma_2|$) than regime 2 (scarcity). The economy stays, on average 10.6 months in the abundance state ($12/\lambda_1 = 10.6$) and 1.89 months in the scarcity state ($12/\lambda_2 = 1.89$). The resulting parameter estimates on the drift and volatility are in line with the model. Also, regime 1 (abundance state) is the most frequent regime with the economy staying approximately 84.9% of the time in it (i.e.,$\lambda_2/(\lambda_1 + \lambda_2) =$

[31] A simplification of our reduced-form model is that it considers constant transition probabilities for the Markov chain.

[32] We follow Kim (1994) algorithm, which is a backward iterative process that starts from the smoothed probability of the last observation.

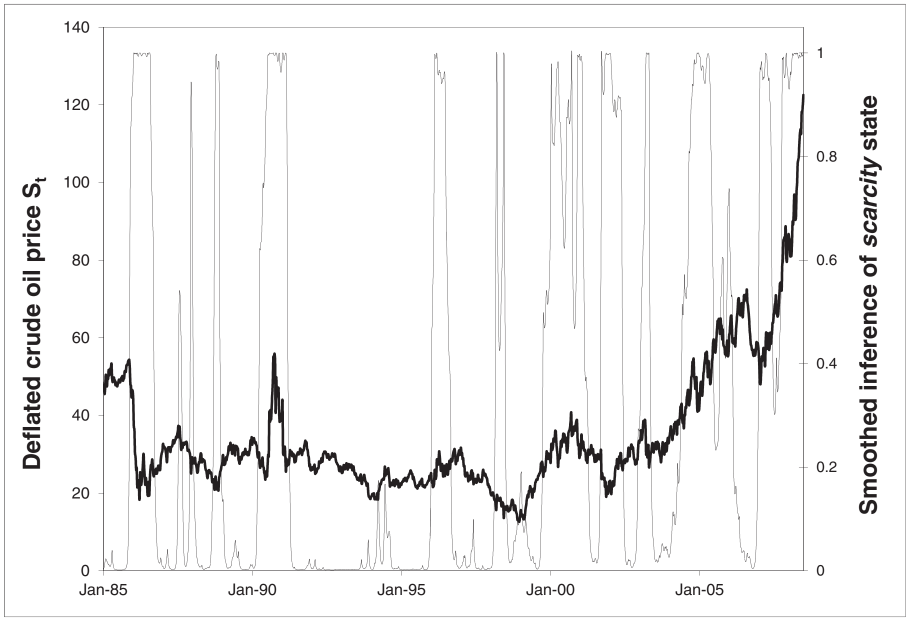


Fig. 8. Historical crude oil prices between Apr/1983 and Jun/2008 deflated by the US Consumer Price Index (thick line) and inferred probability of being in the scarcity state (thin line).

0.849). This mirrors closely the calibration of our general equilibrium model in the prior section.

When we observe the estimates from Panel B, we find that again the economy stays more time in regime 1 (abundance state) (i.e.,$\lambda_1 < \lambda_2$), however, regime 2 (scarcity) becomes more frequent than with the restricted sample period. The economy now stays on average 4.81 months ($12/\lambda_2 = 4.81$), a results that is consistent with the higher prices observed between 2003 and 2008. Accordingly, the estimation delivers a higher upper bound for the oil price,$S_{max} = \$125.34$ ($\log S_{max} = 4.831$). Finally, regime 1 (abundance) again has a lower volatility than regime 2 (scarcity) ($|\sigma_1| < |\sigma_2|$), although now both mean reversion parameters are insignificant.

Fig. 8 shows the crude oil price and the inferred probability of being in regime 2 (scarcity state) for the whole sample period. Consistently with the previous result, the smoothed inference about whether the state of the economy is in regime 1 or 2 are also consistent with our general equilibrium model. We use these inferred probabilities to investigate the risk premium properties of the model next.

4.3. Time-varying oil risk premium

Section 4.1.2 showed that in our model the risk premium on oil is state-dependent. Can we detect this time variation in the data? To investigate this we consider a simple linear regression between crude oil returns and aggregate risk

$$r^e_{j,t+1} = a_t + b_t\, r^e_{M,t+1} + \epsilon_{t+1} \tag{45}$$

where we condition a_t and b_t on the scarcity/abundance regime of the economy. In particular, we assume a linear relation between these variables and $\hat{p}_t$, our inferred (smoothed) probability of being in the scarcity state estimated using the regime switching model in the previous section (see Fig. 8)

$$a_t = a_0 + a_1\, \hat{p}_t \tag{46}$$

$$b_t = b_0 + b_1\, \hat{p}_t \tag{47}$$

A significant estimate for b_1 will imply that the oil risk premium is sensitive to the scarcity of oil. The return on a long position in oil

Table 4
The table presents parameter estimates for the regressions: $r^e_{j,t+1} = a_t + b_t\, r^e_{M,t+1} + \epsilon_t$ where $a_t = a_0 + a_1\, p_t$, $b_t = b_0 + b_1\, p_t$. $\hat{p}_t$ is the conditioning variable and is the smoothed inferred probability of being in the scarcity regime. $r^e_{j,t+1}$ is the (log) monthly real excess return of crude oil spot prices (columns (1) and (2)) or the (log) monthly real excess return of the collateralized futures strategy (columns (3) and (4)). $r^e_{M,t+1}$ is the real monthly excess return of the value-weighted CRSP index. The scaling variables $\hat{p}_t$ is the smoothed inferred probability of being in the scarcity regime (from the estimation in Table 3). The number of observations for all regressions is 240. The t-statistics are in parentheses.

		Oil spot price returns		Oil futures returns	
		(1)	(2)	(3)	(4)
Constant	a_0	−0.07	0.17	1.05	1.12
		(0.11)	(0.23)	(1.67)	(1.54)
$\hat{p}_t$	a_1		−0.88		−0.17
			(0.45)		(0.03)
$r^e_{M,t+1}$	b_0	−0.20	0.05	−0.17	0.06
		(1.47)	(0.30)	(1.28)	(0.41)
$\hat{p}_t\, r^e_{M,t+1}$	b_1		**−1.23**		**−1.17**
			(2.95)		(2.85)
R^2		0.01	0.05	0.01	0.04

is $r^e_{j,t+1}$ and is computed two ways. We use the return on a spot investment in oil (analogous to a buy and hold strategy). As an alternative, we use a fully collateralized long futures position in oil.[33] We proxy aggregate risk with U.S. equities (CRSP value weighted index).

The OLS results are consistent with the model in that the risk premium is sensitive to the scarcity and abundance regimes. Columns 1 and 3 in Table 4 document the well known fact that risk premium in oil markets are, if they exist, not easy to detect. Unconditionally, an investment in oil does not bear systematic risk. The intercept from the unconditional regression is also not sig-

[33] Full collateralization means that when an investor buys a futures contract with price $F(t, T)$, she simultaneously invests this amount in T-bills. Therefore, the total return earned by the investor over a given time period, will be the change in price of the futures contract plus the interest on the investment. This is the more common approach for investors to take positions in energy markets. See Gorton and Rouwenhorst (2006) and Erb and Harvey (2006).

nificant. Including the inferred probability of being in the scarcity regime, reported in columns 2 and 4 of Table 4, has a significant impact. Including the regime improves the fit of the model (the R^2 increase from 1% to 5% and from 1% to 4%, respectively). Consistent with our model, the risk premium appears lower and negative in scarcity regime. The coefficient estimate of b_1 is negative and significant. In contrast, the risk premium in the abundance region is not different from zero.

We tried several alternative specifications for the returns and proxy for the scarcity/abundance regime. The results are broadly consistent. The risk premium is negative and significant in the (proxy to) the scarcity region. These results may help understand why detecting risk premia in commodity trading strategies has thus far produced ambiguous results (e.g., Gorton and Rouwenhorst, 2006 and Erb and Harvey, 2006). However, we leave a joint estimation of the risk premium properties and all the parameters of our model for future research.[34]

5. Conclusion

We developed a two-good general equilibrium model where one of the goods, oil, is not directly consumed but is essential for production of the consumption good. Two key assumptions in the model are that oil investment is irreversible and occurs at a cost that includes a fixed component. These assumptions imply that oil investment is periodic and lumpy. Another important assumption in the oil sector is that the flow rate of oil to production is fixed. As a result, the implied oil price dynamics differ when the economy is near the point of new oil investment, where the crude oil is relatively scarce, than far from the investment threshold, when there is plenty of oil. This implies a rich behavior for the term structure of futures oil price. We estimate key parameters of our model using a Simulated Method of Moments estimation to match features of the economy (oil-to-GDP ratio, interest rates) and properties of oil prices. In particular, we match the term structure of oil futures average prices and price volatility. Since much of the predictive power of the model comes from the differing spot-price behavior across the scarcity/abundance regions, we consider a simpler "reduced-form" model. We estimate a two-regime switching model where the spot price within the regime is affine. We find that not only does this empirical specification fit the data reasonably well, but it is also a sensible approximation of the general equilibrium model. This characterization, then, captures the key properties of the general equilibrium model and maintains much of the tractability of standard commodity derivative pricing. Lastly, we use our general equilibrium setting to explore the risk premium characteristics of a long investment in oil futures. Immediately, the two-regime property for oil prices translates to a risk premium that is state dependent. In fact, in our model the risk premium is positive in one state and negative in the other. Empirically, controlling for the scarcity/abundance regime helps to identify the risk premium in an oil position.

Appendix

A. Sufficient conditions for existence of a solution

We note that our problem is slightly different than in traditional models without fixed costs such as Dumas (1992) or Kogan (2001). Indeed, unlike in these models the no-transaction cost problem does not provide for a natural upper bound. Indeed, in our case, if we set $\beta_K = \beta_Q = \beta_X = 0$ the value function becomes infinite, since it is then optimal to build an infinite number of oil wells (at no cost). Thus unlike in these papers, it is natural to expect that sufficient conditions on the parameters for existence of the solution should depend on the marginal cost of building an oil well (as well as other parameters). Indeed, intuitively, if the marginal costs of an additional oil well is too low relative to the marginal productivity of oil in the K-technology one would expect the number of oil wells built (and thus the value function) to be unbounded. To establish reasonable conditions on the parameters we consider the case where there are only variable costs ($\beta_K = \beta_Q = 0$ and $\beta_X > 0$), but where the investment decision is perfectly reversible. Let us denote $J_u(t, K, Q)$ the value function of the perfectly reversible investment/consumption problem. Clearly, the solution to that problem will be an upper bound to the value function of (4).

When the investment decision is perfectly reversible then it becomes optimal to adjust the oil stock continuously so as to keep $\frac{J_{uQ}}{J_{uK}} = \beta_X$. This suggests that one can reduce the dimensionality of the problem, and consider as the unique state variable $W_t = K_t + \beta_X Q_t$ the 'total wealth' of the representative agent (at every point in time the agent can freely transform Q oil units into $\beta_X Q$ units of consumption good and vice-versa). Indeed, the dynamics of W are:

$$dW_t = \left(\alpha K_t^{1-\eta}\left(\bar{i}Q_t\right)^{\eta} - C_t - \beta_X\left(\bar{i}+\delta\right)Q_t\right)dt + \sigma_K K_t dw_{K,t} + \beta_X \sigma_Q Q_t dw_{Q,t} \tag{A1}$$

Since along each path, the agent can freely choose to adjust the ratio of oil to capital stock $Z_t = \frac{Q_t}{K_t}$, the Cobb–Douglas structure suggests that it will be optimal to maintain a constant ratio,$Z_t = Z^*$. We may rewrite the dynamics of W_t as

$$\frac{dW_t}{W_t} = \left(\mu_W^u(Z^*) - c_t^u\right)dt + \sigma_W^u(Z^*)dw_{W,t} \tag{A2}$$

where $w_{W,t}$ is a standard Brownian motion and we define

$$C_t = c_t^u W_t, \tag{A3}$$

$$\mu_W^u(Z) = \frac{\alpha(\bar{i}Z)^{\eta} - (\bar{i}+\delta)\beta_X Z}{1+\beta_X Z} \tag{A4}$$

and

$$\sigma_W^u(Z) = \frac{\sqrt{\sigma_K^2 + 2\rho_{KQ}\sigma_K\sigma_Q\beta_X Z + (\beta_X Z)^2\sigma_Q^2}}{1+\beta_X Z}. \tag{A5}$$

The proposition below verifies that if the function

$$f(Z) = \frac{\rho}{1-\gamma} - \mu_W^u(Z) + \gamma\frac{\sigma_W^u(Z)^2}{2} \tag{A6}$$

admits a global minimum at Z^* such that

$$a^u := \frac{1}{\gamma}\left\{\rho - (1-\gamma)\left(\mu_W^u(Z^*) - \gamma\frac{\sigma_W^u(Z^*)^2}{2}\right)\right\} > 0 \tag{A7}$$

then the optimal strategy is indeed to consume a constant fraction of total wealth $c_t^u = a^u$ and to invest continuously so as to keep $Q_t/K_t = Z^*$.

Proposition A1. *Assume that there are no fixed costs ($\beta_K = \beta_Q = 0$), and that investment is costly ($\beta_X > 0$), but fully reversible. If the function f(Z) defined in (A6) admits a global minimum Z^* such that condition (A7) holds then the optimal value function is given by*

$$J_u(t, K, Q) = e^{-\rho t}\frac{(a^u)^{-\gamma}(K+\beta_X Q)^{1-\gamma}}{1-\gamma} \tag{A8}$$

The optimal consumption policy is

$$C_t^* = a^u(K_t^* + \beta_X Q_t^*) \tag{A9}$$

[34] In particular, the use of the smoothed probabilities from our linear regime shifting model are estimated from the whole data set and are subject to a look-ahead bias (see Lettau and Ludvigson, 2001.)

and the investment policy is characterized by:

$$\frac{Q_t^*}{K_t^*} = Z^*. \tag{A10}$$

Proof. Applying Itô's Lemma to the candidate value function we have:

$$\frac{dJ_u(t, K_t, Q_t) + U(t, C_t)dt}{J_u(t, K_t, Q_t)} = (1-\gamma)\{h(c_t) - f(Z_t)\}dt + (1-\gamma)\sigma_W^u(Z_t)dw_{W,t} \tag{A11}$$

where we have set $C_t = c_t(K_t + \beta_X Q_t)$. $\sigma_W^u(Z)$ and $f(Z)$ are defined in Eqs. (A5) and (A6), respectively, and we have defined:

$$h(c) = (a^u)^\gamma \frac{(c)^{1-\gamma}}{1-\gamma} - c.$$

Note that the function $h(c)$ is concave and admits a global maximum $c_t^* = a^u$ with $h(a^u) = \frac{a^u \gamma}{1-\gamma}$. Suppose the function $f(Z)$ is strictly convex and admits a global minimum at Z^*. Then, if we pick the constant a^u such that $h(a^u) = f(Z^*)$, we have for any c, Z:

$$h(c) - f(Z) \le h(c^*) - f(Z^*) = 0$$

Thus integrating Eq. (A11) we obtain:

$$J_u(T, K_T, Q_T) + \int_0^T U(t, C_t)dt \le J_u(0, K_0, Q_0) + \int_0^T (1-\gamma)J_u(t, K_t, Q_t)\sigma_W^u(Z_t)dw_{W,t} \tag{A12}$$

Taking expectation and using the fact that the stochastic integral is a positive local martingale we obtain:

$$E\left[J_u(T, K_T, Q_T) + \int_0^T U(t, C_t)dt\right] \le J_u(0, K_0, Q_0) \tag{A13}$$

Further we note that for when we choose the controls $c_t = a^u$ and $Z_t = Z^*$ then we obtain equality in Eq. (A12) and further have:

$$\frac{dJ_u}{J_u} = -a^u dt + (1-\gamma)\sigma_W^u(Z^*)dw_{W,t} \tag{A14}$$

which implies that the local martingale is a martingale and thus (A13) obtains with equality. Further we have

$$\lim_{T\to\infty} E[J_u(T, K_T, Q_T)] = \lim_{T\to\infty} J_u(0, K_0, Q_0)e^{-a^u T} = 0$$

under the assumption (A7). Letting $T \to \infty$ in (A13) shows that our candidate value function indeed is the optimal value function and confirms that the chosen controls are optimal. □

We note that in the case where $\eta = 0$, then Oil has no impact on the optimal decisions of the agent and the value function J_u is the typical solution one obtains in a standard Merton (1973) or Cox et al. (1985) economy. In that case, the condition on the coefficient a^u becomes:

$$a_0 = \frac{1}{\gamma}\left\{\rho - (1-\gamma)(\alpha - \gamma\frac{\sigma_K^2}{2})\right\} > 0. \tag{A15}$$

A lower bound to the value function is easily derived by choosing to never invest in new oil reserves (i.e., setting $dI_t = 0\ \forall t$) and by choosing an arbitrary feasible consumption policy $C_t^l = \alpha K_t^{1-\eta}(\bar{i}Q_t)^\eta$. Indeed, in that case we have:

$$\frac{dK_t}{K_t} = \sigma_K dw_{K,t} \tag{A16}$$

It follows that if the following condition holds:

$$a^l := \rho + (1-\gamma)\left\{(1-\eta)\gamma\frac{\sigma_K^2}{2} + \eta\left(\bar{i} + \delta + \gamma\frac{\sigma_Q^2}{2}\right)\right\} + (1-\eta)\eta(1-\gamma)^2\left\{\frac{\sigma_K^2}{2} - \rho_{KQ}\sigma_K\sigma_Q + \frac{\sigma_Q^2}{2}\right\} > 0 \tag{A17}$$

then, we have

$$J_l(0, K_0, Q_0) := E\left[\int_0^\infty e^{-\rho t}\frac{(C_t^l)^{1-\gamma}}{1-\gamma}dt\right] = \frac{1}{a^l}\frac{(C_0^l)^{1-\gamma}}{1-\gamma} \tag{A18}$$

We collect the previous results and a few standard properties of the value function in the following proposition.

Proposition A2. *If a^l, $a^u > 0$, the value function of problem (4) has the following properties.*

1. *$J_l(t, K, Q) \le J(t, K, Q) \le J_u(t, K, Q)$.*
2. *$J(t, K, Q)$ is increasing in K, Q.*
3. *$J(t, K, Q)$ is concave homogeneous of degree $(1-\gamma)$ in Q and K.*

For the estimation we select parameters such that conditions (A7) and (A17) are satisfied, i.e., that a^l, $a^u > 0$.

B. Homogeneity and optimality

Section 2.2.3 presents how to use the homogeneity to reduce the number of states variables. In this section we offer more details needed to characterize the solution and present a proposition that defines the optimal consumption and investment strategy.

We start by recalling that $j(z)$ satisfies

$$J(K, Q) = \frac{K^{1-\gamma}}{1-\gamma}j(z) \tag{B1}$$

where z is the log of the oil reserves to capital ratio $z = \log\left(\frac{Q}{K}\right)$. The dynamic process for z_t is in Eq. (13) where

$$\mu_{z,t} = \left(-(\bar{i}+\delta) - \frac{1}{2}\sigma_Q^2\right) - \left(\alpha(\bar{i}e^z)^\eta - c_t^* - \frac{1}{2}\sigma_K^2\right), \tag{B2}$$

$$\sigma_z = \sqrt{\sigma_K^2 - 2\rho_{KQ}\sigma_K\sigma_Q + \sigma_Q^2}, \tag{B3}$$

$$\Lambda_z = z_2 - z_1, \tag{B4}$$

and the consumption rate, $c_t^* = C_t^*/K_t$, is a function of z_t.

Plugging consumption in Eq. (17) and the definition of $J(K, Q)$ into the HJB equation in (5) we obtain one-dimensional ODE for the function j.

$$\theta_0 j(z) + \theta_1 j'(z) + \theta_2 j''(z) + \gamma\left(j(z) - \frac{j'(z)}{1-\gamma}\right)^{1-\frac{1}{\gamma}} + \alpha(\bar{i}e^z)^\eta\left((1-\gamma)j(z) - j'(z)\right) = 0 \tag{B5}$$

where

$$\theta_0 = -\rho - \gamma(1-\gamma)\frac{\sigma_K^2}{2},$$

$$\theta_1 = -(\bar{i}+\delta) + \gamma\sigma_K(\sigma_K - \rho_{KQ}\sigma_Q) - \frac{\sigma_z^2}{2}, \quad \theta_2 = \frac{\sigma_z^2}{2} \tag{B6}$$

To determine the investment policy, $\{z_1, z_2\}$, the value-matching condition of Eq. (8) becomes:

$$(1 + e^{z_2}\beta_X)^{1-\gamma} j(z_1) - \left(1 - \beta_K + e^{z_1}(\beta_X - \beta_Q)\right)^{1-\gamma} j(z_2) = 0 \tag{B7}$$

Lastly, using the homogeneity there are only two super-contact conditions to determine that are[35]

[35] In a similar way, if $\beta_K = \beta_Q = 0$ the two super-contact conditions become the same condition $(1 + (1-\gamma)e^{z_1}\beta_X)j'(z_1) - (1 + e^{z_1}\beta_X)j''(z_1) = 0$.

$$(1-\gamma)e^{z_1}(\beta_X-\beta_Q)j(z_1)-\left(1-\beta_K+e^{z_1}(\beta_X-\beta_Q)\right)j'(z_1)=0 \tag{B8}$$

$$(1-\gamma)e^{z_2}\beta_X j(z_2)-(1+e^{z_2}\beta_X)j'(z_2)=0 \tag{B9}$$

The following proposition summarizes the above discussion and offers a verification argument. Let us define the functions:

$$a(z) := j(z)-\frac{j'(z)}{1-\gamma} \tag{B10}$$

$$F(x,y) := \left(\frac{1-\beta_K+e^x(\beta_X-\beta_Q)}{1+\beta_X e^y}\right)^{1-\gamma}\frac{j(y)}{1-\gamma}-\frac{j(x)}{1-\gamma} \tag{B11}$$

Proposition B1. *Suppose that we can find two constants z_1, z_2 $(0\le z_1\le z_2)$ and a function $j(\cdot)$ defined on $[z_1,\infty)$, which solve the ODE given in Eq. (B5) with boundary conditions (B7), (B8), and (B9), such that the following holds:*

$$0 < a(z)^{-1/\gamma} < M_1 \tag{B12}$$

$$0 < \frac{a(z)}{j(z)} < M_2 \tag{B13}$$

$$F(x,y) \le 0, \qquad \forall y \ge x \ge z_1 \tag{B14}$$

$$0 = F(z_1,z_2) \ge F(z_1,y), \qquad \forall y \ge z_1 \tag{B15}$$

where M_1, M_2 are constants.

Then the value function is given by

$$J(t,K,Q)=e^{-\rho t}\frac{K^{1-\gamma}}{1-\gamma}j(z) \tag{B16}$$

where $z=\log\frac{Q}{K}$. Further the optimal consumption policy is to set

$$c(z_t)=a(z_t)^{-1/\gamma}.$$

The optimal investment policy consists of a sequence of stopping times and investment amounts,$\{(T_i,X_{T_i})\}_{i=0,2\dots}$ given by $T_0=0$ and:

- *If $z_0\le z_1$ then invest (to move z_0 to z_2):*

$$X_0^*=Q_0\frac{e^{-z_0}(1-\beta_K)-e^{-z_2}-\beta_Q}{e^{-z_2}+\beta_X} \tag{B17}$$

Then start with new initial values for the stock of consumption good $K_0-\beta(K_0,Q_0,X_0^*)$ and new oil stock $Q_0+X_0^*$.

- *If $z_0>z_1$ then set $X_0^*=0$ and define the sequence of* **F***-stopping times:*

$$T_i=\inf\{t>T_{i-1}: z_{t^-}=z_1\}\ \ i=1,2,\dots \tag{B18}$$

and corresponding $\mathcal{F}_{T_i}$-measurable investments in oil:

$$X_{T_i}^*=Q_{T_i}\frac{e^{-z_1}(1-\beta_K)-e^{-z_2}-\beta_Q}{e^{-z_2}+\beta_X}. \tag{B19}$$

Proof. We define our candidate value function as $J(K,Q,t)=e^{-\rho t}\frac{K^{1-\gamma}}{(1-\gamma)}j(z)$, where $z=\log(Q/K)$ as before and where we define $j(z)$ as in the proposition for $z\ge z_1$ and where we set

$$j(z)=\left(\frac{1-\beta_K+e^z(\beta_X-\beta_Q)}{1+\beta_X e^{z_2}}\right)^{1-\gamma}j(z_2), \quad \forall z<z_1.$$

Applying the generalized Itô's Lemma to our candidate value function for some arbitrary controls we find:

$$\begin{aligned}
&dJ(t,K_t,Q_t)+U(t,C_t)dt\\
&= e^{-\rho t}K_{t^-}^{1-\gamma}\Bigg\{\Bigg[\frac{\hat\theta_0(z_t)j(z_t)+\hat\theta_1(z_t)j'(z_t)+\theta_2 j''(z_t)}{1-\gamma}\\
&\quad+\frac{(c_t)^{1-\gamma}}{1-\gamma}-a(z_t)\,c_t\Bigg]dt+a(z_t)\sigma_K dw_{K,t}\\
&\quad+\{j(z_t)-a(z_t)\}\sigma_Q dw_{Q,t}+F(z_{t^-},z_t)\Bigg\}
\end{aligned} \tag{B20}$$

where for simplicity we have defined $\hat\theta_0(z)=\theta_0+(1-\gamma)\alpha(\bar{i}e^z)^\eta$ and $\hat\theta_1(z)=\theta_1-\alpha(\bar{i}e^z)^\eta$ and $C_t=c_tK_t$.

Now the definition of the function $j(z)$ implies that

$$\frac{\hat\theta_0(z)j(z)+\hat\theta_1(z)j'(z)+\theta_2 j''(z)}{1-\gamma}+\sup_c\left[\frac{(c)^{1-\gamma}}{1-\gamma}-a(z)\,c\right]\begin{cases}=0 & \forall z\ge z_1\\ <0 & \forall z<z_1\end{cases}$$

Further,$F(x,\ y)\le 0\,\forall x\le y$ with equality only if $x\le z_1$ and $y=z_2$. Thus we have that for arbitrary controls

$$\begin{aligned}
&J(T,K_T,Q_T)+\int_0^T U(t,C_t)dt\\
&\le J(0,K_0,Q_0)+\int_0^T e^{-\rho t}K_{t^-}^{1-\gamma}a(z_t)\sigma_K dw_{K,t}+\\
&\int_0^T e^{-\rho t}K_{t^-}^{1-\gamma}\{j(z_t)-a(z_t)\}\sigma_Q dw_{Q,t}.
\end{aligned} \tag{B21}$$

Taking expectation (using the fact that the stochastic integral is a positive local martingale hence a supermartingale) we obtain that for arbitrary controls

$$E\left[J(T,K_T,Q_T)+\int_0^T U(t,C_t)dt\right]\le J(0,K_0,Q_0) \tag{B22}$$

For the controls proposed in the proposition equation (B21) holds with equality. Further, we have for these particular controls:

$$\frac{dJ(t,K_t,Q_t)}{J(t,K_t,Q_t)}=-a(z_t)^{-1/\gamma}\frac{a(z_t)}{j(z_t)}dt+\sigma_J\left(\frac{a(z)}{j(z)}\right)dw_{J,t} \tag{B23}$$

where $w_{J,t}$ is a standard Brownian motion and

$$\sigma_J(x)=(1-\gamma)\sqrt{x^2\sigma_K^2+2x(1-x)\rho_{KQ}\sigma_K\sigma_Q+(1-x)^2\sigma_Q^2}. \tag{B24}$$

This implies that (using the assumptions that $\frac{a(z)}{j(z)}\in(0,M_1)$ and $a(z_t)^{-1/\gamma}\in(0,M_2)$) the stochastic integral in (B21) is a martingale and that

$$\lim_{T\to\infty}E[J(T,K_T,Q_T)]=\lim_{T\to\infty}J(0,K_0,Q_0)\tilde{E}\left[e^{-\int_0^T a(z_t)^{-1/\gamma}\frac{a(z_t)}{j(z_t)}dt}\right]=0.$$

where we have defined a new measure $\tilde{P}\sim P$ by the Radon–Nikodym derivative

$$\frac{d\tilde{P}}{dP}=e^{-\int_0^T\frac{1}{2}\sigma_J\left(\frac{a(z_t)}{j(z_t)}\right)^2dt+\int_0^T\sigma_J\left(\frac{a(z_t)}{j(z_t)}\right)dw_{J,t}} \tag{B25}$$

□

C. Data

For the SMM estimation we use quarterly time series from Q4/1990 to Q2/2008. We build the series of crude oil futures prices and interest rates, private consumption, GDP and petroleum consumption from OECD countries. Crude oil futures prices are obtained from the New York Mercantile Exchange (NYMEX). We use contracts with maturities of 1, 3, 6, 9, 12, 18, 24, 30 and 36 months. If a specific contract is missing, we select the one with the nearest maturity. For the quarterly figures we use the average prices within that period. To get the (annualized) volatility of futures returns, we sample quarterly observations of a GARCH(1,1) estimated separately for each (log) futures series using weekly prices. Consumption and output data is from http://www.oecd.org. The aggregate data is available from Q1/1995 for all OECD members (30 countries). For the initial years we build a proxy for the series with the G7 countries data available from the same site. We assume that the GDP ratio of the G7 countries and all OECD members was constant from Q1/1990 to Q1/1995. Petroleum consumption data for OECD countries is from the U.S. Energy Information Administration site (http://www.eia.doe.gov). Finally, the interest rate data is from Federal Reserve FRED site (http://www.research.stlouisfed.org/fred2). To build the real interest rate time series we also use the CPI series, which are obtained from the same site.

D. Numerical techniques

In this appendix we delineate the numerical algorithm used to solve the HJB equation in (5) with boundary conditions represented by Eqs. (8)–(11).

The first step is to use the homogeneity of the solution to reduce the state space (see Section 2.2.3). After this is done, the solution of the problem is represented by a nonlinear second order ODE in the state variable $z_t = \log(Q_t/K_t)$. The boundary conditions are also expressed in terms of z_t. Now, we need to determine the value function $j(z)$ in Eq. (12). The nonlinear HJB equation for $j(z)$ depends on (i) the optimal control c_t^*, and on (ii) the optimal investment strategy $\{z_1, z_2\}$ determined by the boundary conditions. Unfortunately, the optimal control itself depends on the value function $j(z)$. This implies that $j(z), c_t^*, z_1$ and z_2 need to be simultaneously determined.

We use an iterative method to solve for $j(z)$. The main idea is to build a conditionally linear ODE for $j(z)$ so it is possible to apply a finite-difference scheme. The selection of the initial guess is extremely important for the convergence of the iteration. We assume that $j^0(z) = 1$ which corresponds to the solution when the oil is not relevant for the production technology ($\eta = 0$). In this case we also know that it is never optimal to invest $z_1^0 \to \infty$.

For every iteration m (for $m = 0\ldots\infty$) we do the following steps:

- Determine the optimal consumption c^{*m} as a function of $j^m(z)$ using Eq. (17).
- We recognize that the ODE for $j^{m+1}(z)$ determines the value function when it is optimal *not* to invest in new stocks of commodity. We name this function as $j^{m+1}_{noinv}(z)$. We calculate the coefficients of the ODE for $j^{m+1}_{noinv}(z)$. It is important to notice that this ODE is linear conditional on c^{*m}.
- Determine the optimal oil to capital ratio z_2^{m+1} using the super contact condition in Eq. (B9). Conditional that it is optimal to invest in new commodity stocks, the returning point is always z_2^{m+1} independent of what was the value of z_t before investment was made. Using this argument we define the extended value matching condition as

$$j^{m+1}_{inv}(z) = j^m(z_2^{m+1})\left(\frac{1-\beta_\kappa + e^z(\beta_x - \beta_Q)}{1+e^{(z_2^{m+1})}\beta_x}\right)^{1-\gamma}. \tag{D1}$$

  This equation represents the value function when the representative agent is forced to invest.
- Use a finite-difference scheme to solve for the value function $j^{m+1}_{noinv}(z)$. The finite difference discretization defines a tridiagonal matrix that needs to be inverted to determine the value of $j^{m+1}_{noinv}(z)$. Instead of doing this, we eliminate the upper diagonal of this matrix. At this point the value of $j^{m+1}_{noinv}(z)$ depends only on the value of $j^{m+1}_{noinv}(z-\Delta z)$. We choose a z_{min} negative enough to ensure that at that level it is optimal to invest, and then we solve the value function for higher z_t. At every point we choose the maximum of the value from investing ($j^{m+1}_{inv}(z)$) and the value of no investing which comes from the finite-difference scheme. This maximum determines the value of $j^{m+1}(z)$. The optimal trigger z_1^{m+1} is endogenously determined when the representative agent is indifferent between investing and postponing the investment. The algorithm described above is a more efficient way than solving independently for $j^{m+1}_{inv}(z)$ and $j^{m+1}_{noinv}(z)$ and then choosing $j^{m+1}(z) = \max(j^{m+1}_{inv}(z), j^{m+1}_{noinv}(z))$.
- Check for the convergence condition. If it not satisfied we start a new iteration with the updated value of $j^{m+1}(z)$.

Once $j(z)$ has converged it is straight forward to calculate spot commodity prices from Eq. (24). For the futures prices we use an implicit finite-difference technique. This is simpler than the solution for $j(z)$ since the coefficients of the PDE and boundary conditions and boundaries $\{z_1, z_2\}$ are known at the beginning of the scheme.

References

Abel, A.B., 1983. Optimal investment under uncertainty. Am. Econ. Rev. 73 (1), 228–233.

Acharya, V.V., Lochstoer, L.A., Ramadorai, T., 2013. Limits to arbitrage and hedging: Evidence from commodity markets. J. Financ. Econ. 109 (2), 441–465. doi:10.1016/j.jfineco.2013.03.003.

Alquist, R., Kilian, L., 2010. What do we learn from the price of crude oil futures? J. Appl. Econometr. 25 (4), 539–573.

Backus, D.K., Routledge, B.R., Zin, S.E., 2005. Exotic preferences for macroeconomists. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2004, 19. MIT Press, Cambridge, MA, pp. 319–391.

Baker, S. D., 2016. The Financialization of Storable Commodities. Working paper, University of Virginia.

Baker, S. D., Routledge, B. R., 2016. The Price of Oil Risk. Working paper, Carnegie Mellon University.

Basak, S., Pavlova, A., 2016. A model of financialization of commodities. J. Financ. 71 (4), 1511–1556. doi:10.1111/jofi.12408.

Bernanke, B.S., 1983. Irreversibility, uncertainty, and cyclical investment. Q. J. Econ. 98 (1), 85–106.

Bloom, N., Bond, S., van Reenen, J., 2007. Uncertainty and investment dynamics. Rev. Econ. Stud. 74 (2), 391–415.

Bobenrieth, E.S.A., Bobenrieth, J.R.A., Wright, B.D., 2002. A commodity price process with a unique continuous invariant distribution having infinity mean. Econometrica 70, 1213–1219.

Carlson, M., Khokher, Z., Titman, S., 2007. Equilibrium exhaustible resource price dynamics. J. Financ. 62 (4), 1663–1703. doi:10.1111/j.1540-6261.2007.01254.x.

Casassus, J., Collin-Dufresne, P., 2005. Stochastic convenience yield implied from commodity futures and interest rates. J. Financ. 60 (5), 2283–2332.

Casassus, J., Liu, P., Tang, K., 2013. Economic linkages, relative scarcity, and commodity futures returns. Rev. Financ. Stud. 26 (5), 1324–1362. doi:10.1093/rfs/hhs127.

Chambers, M., Bailey, R., 1996. A theory of commodity price fluctuations. J. Polit. Econ. 104, 924–957.

Cochrane, J.H., 2005. Asset Pricing. Princeton University Press.

Cox, J.C., Ingersoll Jr., J.E., Ross, S.A., 1985. An intertemporal general equilibrium model of asset prices. Econometrica 53 (2), 363–384.

Croce, M.M., 2014. Long-run productivity risk: a new hope for production-based asset pricing? J. Monet. Econ. 66, 13–31. doi:10.1016/j.jmoneco.2014.04.001.

David, A., 2015. Exploration Activity, Long Run Decisions, and the Risk Premium in Energy Futures. Working paper, University of Calgary.

Deaton, A., Laroque, G., 1992. On the behaviour of commodity prices. Rev. Econ. Stud. 59 (1), 1–23.

Dixit, A., 1991. A simplified treatment of the theory of optimal regulation of brownian motion. J. Econ. Dyn. Control 15 (4), 657–673.

Dixit, A., 1993. The Art of Smooth Pasting. Fundamentals of Pure and Applied Economics, 55. Harwood Academic Publishers.

Dixit, A.K., Pindyck, R.S., 1994. Investment under Uncertainty. Princeton University Press, Princeton.

Duffie, D., 2001. Dynamic Asset Pricing Theory. Princeton, NJ: Princeton University Press.

Duffie, D., Epstein, L.G., 1992. Stochastic differential utility. Econometrica 60 (2), 353–394.

Duffie, D., Singleton, K., 1993. Simulated moments estimation of markov models of asset prices. Econometrica 61 (4), 929–952.

Dumas, B., 1991. Super contact and related optimality conditions. J. Econ. Dyn. Control 15 (4), 675–685.

Dumas, B., 1992. Dynamic equilibrium and the real exchange rate in a spatially separated world. Rev. Financ. Stud. 5 (2), 153–180.

Epstein, L.G., Zin, S.E., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica 57 (4), 937–969.
Erb, C.B., Harvey, C.R., 2006. The strategic and tactical value of commodity futures. Financ. Anal. J. 62 (2), 69–97.
Finn, M.G., 1995. Variance properties of Solow's productivity residual and their cyclical implications. J. Econ. Dyn. Control 19 (5-7), 1249–1281.
Finn, M.G., 2000. Perfect competition and the effects of energy price increases on economic activity. J. Money, Credit Bank. 32 (3), 400–416.
Gibson, R., Schwartz, E.S., 1990. Stochastic convenience yield and the pricing of oil contingent claims. J. Financ. 45 (3), 959–976.
Gorton, G., Rouwenhorst, K.G., 2006. Facts and fantasies about commodity futures. Financ. Anal. J. 62 (2), 47–68.
Guesnerie, R., 1975. Pareto optimality in non-convex economies. Econometrica 43 (1), 1–29.
Gustafson, R. L., 1958. Carryover levels for grains: a method for determining amounts that are optimal under specified conditions.Technical Bulletin No. 1178, U.S. Department of Agriculture, Washington, DC.
Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57 (2), 357–384.
Hamilton, J.D., 2003. What is an oil shock? J. Econometr. 113 (2), 363–398.
Hamilton, J. D., 2005. Oil and the Macroeconomy. Working Paper, University of California, San Diego.
Hamilton, J.D., Wu, J.C., 2014. Risk premia in crude oil futures prices. J. Int. Money Financ. 42, 9–37. doi:10.1016/j.jimonfin.2013.08.003. Understanding International Commodity Price Fluctuations
Harrison, J.M., 1990. Brownian Motion and Stochastic Flow Systems. Krieger Publishing Company.
Hitzemann, S., 2016. Macroeconomic fluctuations, oil supply shocks, and equilibrium oil futures prices. Working paper, Ohio State University.
Hotelling, H., 1931. The economics of exhaustible resources. J. Polit. Econ. 39 (2), 137–175.
IMF, 2008. World Economic Outlook: Financial Stress, Downturns, and Recoveries. International Monetary Fund, Washington DC.
Karatzas, I., Shreve, S.E., 1998. Methods of Mathematical Finance. Springer New York.
Kim, C.-J., 1994. Dynamic linear models with markov-switching. J. Econometr. 60 (1-2), 1–22.
Kim, I.-M., Loungani, P., 1992. The role of energy in real business cycle models. J. Monet. Econ. 29 (2), 173–189.
Kogan, L., 2001. An equilibrium model of irreversible investment. J. Financ. Econ. 62 (2), 201–245.
Kogan, L., Livdan, D., Yaron, A., 2009. Oil futures prices in a production economy with investment constraints. J. Financ. 64 (3), 1345–1375.
Kydland, F.E., Prescott, E.C., 1982. Time to build and aggregate fluctuations. Econometrica 50 (6), 1345–1370.
Leduc, S., Sill, K., 2004. A quantitative analysis of oil-price shocks, systematic monetary policy, and economic downturns. J. Monet. Econ. 51 (4), 781–808.
Lee, B.-S., Ingram, B.F., 1991. Simulation estimation of time-series models. J. Econometr. 47 (2-3), 197–205.
Lettau, M., Ludvigson, S.C., 2001. Consumption, aggregate wealth, and expected stock returns. J. Financ. 56 (3), 815–849.
Litzenberger, R.H., Rabinowitz, N., 1995. Backwardation in oil futures markets: theory and empirical evidence. J. Financ. 50 (5), 1517–1545.
Lucas, R.E., 2003. Macroeconomic priorities. Am. Econ. Rev. 93 (1), 1–14. doi:10.1257/000282803321455133.
Mehra, R., Prescott, E.C., 1985. The equity premium: a puzzle. J. Monet. Econ. 15 (2), 145–161. doi:10.1016/0304-3932(85)90061-3.
Merton, R.C., 1973. An intertemporal capital asset pricing model. Econometrica 41 (5), 867–887.
Newbery, D.M., Stiglitz, J.E., 1981. The Theory of Commodity Price Stabilization: A Study of the Economics of Risk. Oxford University Press, Oxford.
Pindyck, R.S., 2001. The dynamics of commodity spot and futures markets: a primer. Energy J. 22 (3), 1–29.
Ready, R., 2016. Oil Consumption, Economic Growth and Oil Futures. Working paper, University of Rochester.
Richard, S.F., Sundaresan, M., 1981. A continuous time equilibrium model of forward prices and futures prices in a multigood economy. J. Financ. Econ. 9 (4), 347–371.
Romer, P.M., 1986. Increasing returns and long-run growth. J. Polit. Econ. 94 (5), 1002–1037.
Ross, S.A., 1997. Hedging long run commitments: exercises in incomplete market pricing. Econ. Notes 26, 99–132.
Rotemberg, J.J., Woodford, M., 1996. Imperfect competition and the effects of energy price increases on economic activity. J. Money, Credit Bank. 28 (4), 549–577.
Routledge, B.R., Seppi, D.J., Spatt, C.S., 2000. Equilibrium forward curves for commodities. J. Financ. 55 (3), 1297–1338.
Scheinkman, J.A., Schechtman, J., 1983. A simple competitive model with production and storage. Rev. Econ. Stud. 50, 427–441.
Schwartz, E.S., 1997. The stochastic behavior of commodity prices: implications for valuation and hedging. J. Financ. 52 (3), 923–973.
Sockin, M., Xiong, W., 2015. Informational frictions and commodity markets. J. Financ. 70 (5), 2063–2098. doi:10.1111/jofi.12261.
Wei, C., 2003. Energy, the stock market, and the putty-clay investment model. Am. Econ. Rev. 93 (1), 311–323.
Williams, J.C., Wright, B.D., 1991. Storage and Commodity Markets. Cambridge University Press, England.
Wright, B.D., Williams, J.C., 1982. The economic role of commodity storage. Econ. J. 92(367), 596–614.
Yang, F., 2013. Investment shocks and the commodity basis spread. J. Financ. Econ. 110 (1), 164–184. doi:10.1016/j.jfineco.2013.04.012.