What Does the Convenience Yield Curve Tell Us about the Crude Oil Market?*

Ron Alquist
Kings Peak Asset Management
ralquist@kingspeakam.com

Gregory H. Bauer
Bank of Canada
gbauer@bankofcanada.ca

Antonio Diez de los Rios
Bank of Canada
diez@bankofcanada.ca

November, 2015

Abstract

Using the prices of crude oil futures contracts, we construct the term structure of crude oil convenience yields out to one-year maturity. The crude oil convenience yield can be interpreted as the interest rate, denominated in barrels of oil, for borrowing one barrel of oil, and it measures the value of storing crude oil over the borrowing period. We show that the convenience yield curve is well explained by a level and a slope factor. Consistent with the theory of storage, convenience yields have predictive power over future crude oil inventories, production, global real economic activity and the price of oil.

JEL classification: C53, G12, G13, Q43
Keywords: Convenience yield; crude oil futures contracts; crude oil inventories; Working curve.

*We thank David Finer, Sandra Ramirez and Argyn Toktamyssov for their excellent research assistance. We also extend a special thanks to Jean-Sebastien Fontaine, Christiane Baumeister, and Lutz Kilian for several useful discussions. Finally, we would like to thank Bahattin Büyükşahin, Sebastian Fossati, Sermin Gungor, Fulvio Pegoraro, Gabriel Power and seminar participants at the University of New South Wales, the 3rd Joint Bank of Spain–Bank of Canada Workshop on International Financial Markets (Madrid, 2014), the First Annual Conference of the International Association for Applied Econometrics (London, 2014), the 2014 European Meeting of the Econometric Society (Toulouse), and the Northern Finance Association 2014 Conference (Ottawa). The views expressed in the paper represent those of the authors and do not necessarily reflect those of the Bank of Canada, its Governing Council, or any other organizations with which the authors are affiliated. **Address for correspondence:** Antonio Diez de los Rios, Bank of Canada, Financial Markets Department, 234 Laurier Ave. West, Ottawa (Ontario), K1A 0G9, Canada.

1 Introduction

Elevated and sometimes volatile crude oil prices have become enduring features of the international economy and preoccupy policy-makers, financial analysts, and the broader public. The run-up and collapse in the price of crude oil between 2003 and 2008 and its persistently high level since 2009 have reinvigorated interest in the question of the fundamental forces that drive crude oil prices (see, e.g., Hamilton 2009; Kilian 2009; Juvenal and Petrella 2014; Tang and Xiong 2012; Hamilton and Wu 2014).

One of the fundamental drivers of the price of crude oil is inventories (Alquist and Kilian 2010; Kilian and Lee 2014; Kilian and Murphy 2014). Since crude oil is a storable commodity, stocks play a central role in the intertemporal relationship linking current demand and supply to expectations of future demand and supply. Storing oil is intrinsically valuable because of the operational flexibility that stocks provide to refiners by reducing the costs of changing production and helping them to avoid stockouts. Consequently, the optimal levels of production and inventories are jointly determined given the spot price of oil and the value of storage (Pindyck 2001).

While the value of storage is not directly observable, it is closely related to the crude oil convenience yield. The convenience yield can be thought of as the interest rate paid in barrels of oil for borrowing one barrel of oil, and it can be constructed from the prices of crude oil futures contracts. The borrower of a barrel of oil is, in essence, supplying storage in the form of crude oil inventories to the lender. As a result, the lender must be compensated for forgoing the benefits associated with holding the barrel of oil. In equilibrium, this condition links the convenience yield to the value of storage, and periods of relative scarcity of the commodity are related to high convenience yields. Several papers have examined this relationship using futures markets for industrial commodities, including crude oil (see Fama and French 1987; Fama and French 1988; Ng and Pirrong 1994; Pindyck 1994; Pindyck 2001; Geman and Ohana 2009). These papers focus on the prices of short-term (e.g., one-month) futures contracts to examine the contemporaneous relationship between the short-term convenience yield and the current level of inventories, or the so-called Working curve (Working 1933).[1]

At the same time that oil prices have reached such elevated levels, there has been a substantial increase in the liquidity of the market for oil futures contracts. On the sell side, financial institutions have become more actively involved in commodity derivatives markets, including futures contracts of longer maturity (Büyükşahin et al. 2008; Spector 2013). On the buy side, increased investor interest has resulted in large quantities of financial capital flowing into these markets during the past decade (Büyükşahin and Harris 2011; Alquist and Gervais 2013). Over a sample period between April 1989 and June 2013, we exploit the increase in the liquidity of longer-maturity oil futures contracts to construct the term structure of convenience yields out to the one-year horizon, something that, to the best of our knowledge, is new to the literature. This approach enables us to examine the information contained in the term structure of convenience yields, which, in equilibrium, should be determined by financial markets participants' views of the future scarcity of oil.

Our analysis sheds new light on the term structure of convenience yields and its relationship with the theory of storage. First, we show that the cross-section of crude oil convenience yields across maturities can be explained by a small number of principal components. Similar to the term structure of interest rates (i.e., Litterman and Scheinkman 1991), the first component resembles a level factor that is common across maturities. The second component is related to the slope of the curve. Second, we find that both the level and slope components have predictive power for future changes of crude oil stocks up to the one-year horizon. This finding differentiates our work from other analyses that have focused on the contemporaneous relationship between convenience yields and the level of inventories only (e.g., Gorton, Hayashi and Rouwenhurst 2012). To the best of

[1] Working (1933) originally documented this relationship using data from the wheat futures market in Chicago. For this reason, there have been many tests for the existence of a stable Working curve in agricultural futures markets. In general, the evidence obtained from such tests indicates that there is relationship between the level of inventories and the convenience yield consistent with the one Working originally identified (e.g., Carter and Revoredo-Geha 2007; and Joseph et al. 2014).

our knowledge, this paper is the first to show that, consistent with the theory of storage, longer-maturity convenience yields are forward-looking variables related to the scarcity of crude oil. Third, the term structure of convenience yields contains information about future crude oil production, global real economic activity and the real price of crude oil. To assess the statistical significance of the results, we use a bootstrap procedure that accounts for the fact that the principal components are generated regressors, as well as for the well-known small-sample biases that plague long-run predictive regressors.

Overall, the evidence suggests that the term structure of crude oil convenience yields contains information about the value of holding crude oil stocks over different horizons in the same way that the dollar yield curve contains information about future economic activity, and inflation.[2]

Our paper is related to, but distinct from, several papers that examine the ability of different types of models to forecast the nominal and real prices of crude (e.g., Alquist and Kilian 2010; Alquist et al. 2013; Baumeister and Kilian 2012; Baumeister et al. 2014). For example, Baumeister et al. (2014) show that inventories forecast the real price of oil, which is consistent with the notion that inventories contain a forward-looking element related to conditions in the oil market. Our findings show that the reason inventories are able to forecast the real price of oil is related to the convenience yield, which summarizes the information contained in changes in inventories.

Finally, our results are also related to those found in the literature that assesses the predictive content of asset markets (see, e.g., Stock and Watson 2003, and the references within). Hong and Yogo (2012) show that open interest in the crude oil futures market contains information about future economic activity and inflation expectations that is not immediately reflected in asset prices. Gospodinov and Ng (2013) find that the first two principal components of a panel of short-term (e.g., one-month) commodity convenience yields contain important predictive power for inflation. Rather than examining the prin-

[2] See, for example, Mishkin (1990), Estrella and Hardouvelis (1991), and Ang et al. (2006) for evidence on the forward-looking nature of the term structure of dollar bond yields.

cipal components extracted from a cross-section of commodity short-term convenience yields across different commodities, we extract the principal components from the cross-section of convenience yields at different maturities for a given commodity and examine whether they contain information about variables specific to the crude oil market. Our approach is thus complementary to theirs.

The remainder of the paper is organized as follows. Section 2 discusses the theory of storage, which provides the theoretical basis for the empirical relationship between crude oil inventories and the term structure of convenience yields. Section 3 provides summary statistics of the data. The empirical evidence about the relationship between convenience yields and crude oil stocks is discussed in section 4. Section 5 reports the results obtained from the predictive regressions for crude oil production, global real economic activity and the real price of crude oil. Section 6 concludes. Technical details regarding the bootstrap methods employed in this paper are provided in the appendix.

2 The Information Contained in the Term Structure of Convenience Yields

We begin by discussing the theoretical basis for the empirical relationship between crude oil inventories and the term structure of convenience yields. The term structure of convenience yields is analogous to the term structure of interest rates: it represents the cost that investors pay in barrels of oil for borrowing a single barrel of oil at different horizons.

2.1 The theory of storage

According to competitive storage models of commodity price determination, convenience yields arise endogenously as the result of the interaction between the demand for the commodity with the supply and storage decisions of the producer (see, e.g., Working 1949; Brennan 1958; Ng and Ruge-Murcia 2000; Routledge, Seppi and Spatt 2000). In such models, inventories play a fundamental role in the formation of a commodity price because holding stocks is intrinsically valuable given the operational flexibility they provide. For example, owing to technological constraints, an oil refinery has the incentive to hold stocks to optimize its output of petroleum products (National Petroleum Council 2004). In addition, the capital investments required to establish a crude oil refinery are much longer lived than the horizon over which a refinery makes plans about storage and production. Adjusting crude oil inventories rather than the capital stock is therefore a key way for a refinery to change its variable costs. The value that the refinery assigns to its ability to expand its product mix can be represented as a convenience yield (e.g., Considine 1997).

To understand the relationship between convenience yields and inventories implied by the theory of storage, we observe that by borrowing a barrel of crude oil, the borrower is supplying storage in the form of crude oil inventories to the lender. Consequently, the lender must be compensated for forgoing the benefit associated with holding the barrel of oil. In equilibrium, this requirement links the convenience yield to the price of storage – that is, the marginal value of the flow of services that accrue from holding an additional unit of inventory net of the cost of physically storing crude oil (see Pindyck 2001).[3]

The theory of storage also suggests that (i) the marginal benefit for holding inventories increases at a decreasing rate with the scarcity of a commodity, and (ii) the marginal cost of physically storing oil can be treated as constant over the relevant range of inventories (see Brennan 1958; Telser 1958; Fama and French 1988). That is, the one-period convenience yield (i.e., the marginal benefit of holding inventories minus the storage cost), $\delta_t^{(1)}$, is assumed to be a function of the level of inventories, I_t, such that

$$\delta_t^{(1)} = C(I_t), \tag{1}$$

where $C' < 0$ and $C'' > 0$.

We therefore expect to observe a negative and monotonic relationship between the convenience yield and the current level of crude oil stocks. This empirical relationship was first documented in the market for wheat by Working (1933), and is commonly referred to as the Working curve. We verify below that this relationship exists in the market for

[3]In situations where the value of holding stocks is small, it is possible to observe negative convenience yields, given that the marginal cost can exceed the marginal benefit of physically storing oil.

U.S. crude oil inventories.

Still, the assumption that the cost of physically storing oil is constant is likely to be violated in practice. For example, storage costs may be increasing in inventories during periods when storage facilities are near capacity. Unfortunately, consistent storage cost data that would permit us to identify separately the marginal benefit of storage from its marginal cost in crude oil convenience yields are not readily available.[4] For this reason, when we refer to convenience yields in the remainder of the paper, we mean the convenience yield net of storage costs.

2.2 Constructing the convenience yield curve

While convenience yields are not directly observable, they can be synthetically replicated by taking simultaneous positions in money, crude oil spot and futures markets. Let S_t be the spot price of oil at time t and $F_t^{(n)}$ be the price at time t of a futures contract that matures at time $t+n$. Also, let $y_t^{(n)}$ be the nominal interest rate at which investors can borrow between period t and $t+n$. Time is measured in months.

An investor can synthetically borrow one barrel of oil by

1. borrowing S_t dollars at time t,
2. using the amount borrowed to buy one barrel of oil at time t,
3. selling $S_t \exp\left[ny_t^{(n)}\right]/F_t^{(n)}$ futures contracts that mature at time $t+n$.

By taking these three positions, an investor receives a barrel of oil at time t and has to pay $S_t \exp\left[ny_t^{(n)}\right]/F_t^{(n)}$ barrels of oil at time $t+1$.[5] This position is a synthetic loan of a barrel of oil. In the absence of arbitrage opportunities, the following cost-of-carry

[4] Ederington et al. (2012) contacted several pipeline and storage operators at Cushing, Oklahoma, the delivery point for the West Texas Intermediate futures contract, to obtain an estimate of the cost of storage there. They report it to be about \$0.40 per barrel per month.

[5] The net flow of cash is zero at $t+1$ given that the payout of the futures contract $S_t \exp\left[ny_t^{(n)}\right]$ at time $t+1$ is used to repay the loan for the S_t dollars borrowed at time t.

equation implies that the price of an oil futures contract that expires in n months satisfies:

$$\begin{aligned} \exp\left[n\delta_t^{(n)}\right] &= S_t \exp\left[n y_t^{(n)}\right]/F_t^{(n)}, \\ f_t^{(n)} - s_t &= n y_t^{(n)} - n\delta_t^{(n)}, \end{aligned} \tag{2}$$

where $f_t^{(n)} = \log F_t^{(n)}$, $f_t^{(n)} - s_t$ is the basis, and $\delta_t^{(n)}$ is the n-month (log) convenience yield (net of storage costs) associated with having access to physical oil for the duration of the contract.

We postulate that equation (2) holds continuously because of the presence of investors who simultaneously trade in the oil futures and the dollar money markets and ensure that the two markets are fully integrated with each other. This assumption is necessary because (2) only holds if investors can take simultaneous long and short positions in the oil futures market and the money markets to eliminate arbitrage possibilities. If the condition were violated, some firms would be able to earn riskless profits. Given the liquidity of the West Texas Intermediate (WTI) futures and money markets, the absence of arbitrage is a plausible assumption during the period we analyze in this paper.

It is important to recognize that the no-arbitrage relationship (2) holds for oil forwards but not oil futures contracts. However, the empirical literature shows that the differences are small between the prices of forwards and futures for a variety of commodities (Chow, McAleer and Sequeira 2000). We thus treat equation (2) as a maintained hypothesis throughout the paper.[6]

2.3 The forward-looking nature of convenience yields

In the traditional presentation of the theory of storage, the focus is on the one-period convenience yield and on the contemporaneous relationship between convenience yields and the level of inventories (i.e., the Working curve). Several papers have examined this relationship in the markets for industrial commodities and, in some cases, crude oil (see

[6] In addition, using the results of the model in Alquist, Bauer and Diez de los Rios (2014), we show that, under the assumption of monthly marking to market, the root-mean-squared price difference between the prices of oil forwards and oil futures is less than one cent.

Fama and French 1987; Fama and French 1988; Ng and Pirrong 1994; Pindyck 1994; Pindyck 2001; Geman and Ohana 2009; Gorton, Hayashi and Rouwenhurst 2012). The theory predicts that periods of relative scarcity of the commodity are related to high convenience yields.

In this paper, we go a step further and analyze the information contained in longer-term convenience yields. By analyzing futures contracts with different expiration dates, we assess the information contained in the term structure of convenience yield about the implicit benefit of physical storage over different horizons. For example, an upward-sloping convenience yield curve indicates a situation in which refineries assign a higher value to future inventories than they do to today's inventories. Such periods indicate that oil inventory is expected to be more scarce in the future. The slope of the convenience yield curve should, therefore, predict changes in inventories.

In fact, by appealing to an expectations hypothesis argument, we have that the n-month oil convenience yield must equal the average of the current and expected future one-month convenience yields plus a risk-premium term, $\psi_t^{(n)}$:

$$\delta_t^{(n)} = \frac{1}{n} E_t \sum_{i=0}^{n-1} \delta_{t+i}^{(1)} + \psi_t^{(n)}, \tag{3}$$

Substituting the postulated relationship between short-term convenience yields and inventories in equation (1) into (3), we obtain:

$$\delta_t^{(n)} = \frac{1}{n} E_t \sum_{i=0}^{n-1} C(I_{t+i}) + \psi_t^{(n)}, \tag{4}$$

which reveals the forward-looking nature of long-term convenience yields.

This relationship is exactly analogous to the forward-looking nature of bond yields implied by the expectations hypothesis of the term structure (see, i.e., Bekaert and Hodrick 2001). For example, if the central bank follows a Taylor rule, then the term structure of dollar bond yields reflects market participants' expectations of future output and inflation (see, e.g., Ang, Dong and Piazzesi 2007). In a similar fashion, long-term convenience yields should contain information about future crude oil inventories.

Moreover, because convenience yields are determined by the interaction of storage decisions with the supply and demand of crude oil, we also expect convenience yields to contain information about future conditions in the physical market for crude oil and future oil prices.

3 Data and Summary Statistics

3.1 Crude oil futures and convenience yields

Crude oil futures. The cost-of-carry equation (2) relies on the premise that the spot and futures markets are linked together in a way consistent with the absence of arbitrage opportunities. Because this assumption requires the existence of liquid oil futures and money markets, we limit the sample to the period between April 1989 and June 2013 and focus on the monthly prices of WTI futures contracts traded on the NYMEX and CME exchanges. During the sample period, liquid futures markets existed for maturities up to 12 months. The WTI contracts are the most liquid in the world and are fully physically deliverable, making them a natural choice for examining the dynamics of the convenience yield. To compute the spot price, we select the futures contract that is closest to delivery (see, e.g., Trolle and Schwartz 2009; Szymanowska et al. 2014). Finally, we use the end-of-month observations of these contracts.

Table 1a shows the summary statistics for the spot and futures prices. Over the sample period, the oil futures curve has been flat with an average difference of only $0.25 between the spot and one-year futures prices. Longer dated futures are approximately as volatile as shorter dated ones.

The time series of the monthly price data are plotted in Figure 1a. The figure shows the spot, 1-, 3-, and 12-month futures contracts over the sample period. There is wide variation in the nominal spot price of oil, ranging from less than $20 per barrel to more than $140 per barrel. From the figure, the tight relationship between the prices of the crude oil futures contracts and the spot price is evident.

Shape of oil futures curve. Figure 2a shows the oil futures curves drawn for the end-of-quarter observations. The spot price has an important effect on the oil futures curve and acts as a level factor in the oil market. Because the prices of crude oil futures contracts are linked to the price of spot oil by the cost-of-carry equation, movements in the spot price of oil result in parallel shifts in the level of the oil futures curve. This observation is confirmed by a principal component analysis of the crude oil futures curve. The first component of the cross-section of futures prices accounts for 99.7 per cent of the cross-sectional variation in the oil futures curve, and its correlation with the spot price of oil is 99.6 per cent.

Consistent with other studies of the crude oil market (e.g., Litzenberger and Rabinowitz 1995), we find that crude oil futures prices can exhibit strong backwardation, in which futures prices are below the current spot price (i.e., the crude oil futures curve is downward sloping). In the sample that we consider, strong backwardation occurs approximately 55 per cent of the time. For example, the one-year futures price is below the spot price of oil in 56.4 per cent of the months in the sample.

Some studies find that crude oil futures prices exhibit behavior consistent with the so-called Samuelson (1965) effect – namely, that the variability of oil futures prices decreases with the maturity of the contract under consideration (e.g., Bessembinder et al. 1995; Casassus and Collin-Dufresne 2005). The theoretical explanation for this effect is the smoothing of expectations of the spot price of crude oil. The spot price of crude oil overshoots in the short run, given that supply takes time to respond to a demand shock. The spot price is thus more volatile than the expected oil prices in subsequent periods and oil futures prices, which, under risk neutrality, equal these expectations.

Unlike these studies, there is no evidence that the sensitivity of oil futures prices to changes in the spot price of oil decreases with the maturity of the contract in our data set (see Table 1a and Figure 2a). This evidence is also consistent with that presented in Alquist and Kilian (2010), who show that the inaccuracy of oil futures-based forecasts is related to the variability of such forecasts rather than their bias. While there are several theoretical explanations for the violation of the Samuelson effect, the theory of storage predicts that such violations can occur when inventories are high (see Fama and French 1988; Routledge, Seppi and Spatt 2000).[7] If crude oil is abundant, the spot price does not need to overshoot, because the initial effect of the demand shock can be absorbed by decreasing stocks. The earlier studies that documented the existence of the Samuelson effect predate the persistent increases in the price of oil and the level of inventories observed in recent years. The persistent price increases may explain why we do not observe a Samuelson effect for crude oil futures. We return to this point in the next section when we introduce the data on inventories.

Interest rates. We use LIBOR data for maturities of 1, 2, 3, 6 and 12 months rather than U.S. Treasury bill data, because the former represent a better measure of the borrowing costs incurred by oil companies.[8] Table 1a reports the summary statistics. The LIBOR curve was, on average, upward sloping during the sample period. On the other hand, short-term rates exhibit greater volatility than long-term rates.

Convenience yields. Given that the futures, spot and dollar interest rates are observable, we use the no-arbitrage relationship (2) to construct the set of convenience yields of maturities of 1, 2, 3, 6 and 12 months. The summary statistics are also shown in Table 1a. The average convenience yield is similar in magnitude to the average LIBOR yield. For example, the average one-month LIBOR yield is 3.81 per cent, while the average of the one-month convenience yield is 4.11 per cent. Similarly, the term structure of convenience yields is upward sloping with a difference between one-month and one-year yields of 280 basis points.

Convenience yields are, however, more volatile than LIBOR yields, and they are less persistent. In addition, the 12-month convenience yield is less volatile, but more persis-

[7] Routledge, Seppi and Spatt (2000) provide a list of the explanations for why there can be violations of the Samuelson effect.

[8] Government bonds can embody large liquidity premia due to favorable taxation treatment, repo specials, scarcity premia and benchmark status (see Fontaine and Garcia 2012).

tent, than short-term convenience yields.

These points are illustrated in Figure 1b, which depicts the 1-, 3- and 12-month convenience yields, measured in per cent per annum. The magnitude of the short-term convenience yield for oil is large and reaches, in a couple of cases, values as high as 75 per cent. If we interpret the convenience yield as the interest rate on an oil bond, these values may seem too large to be economically plausible, but they are similar in size to those obtained by Casassus and Collin-Dufresne (2005) for the one-week convenience yields.[9] The magnitude of the one-year convenience yield seems to be bounded between -25 and 40 per cent. Furthermore, this pattern is consistent with the supply of storage being inelastic in the short run, which, in turn, causes the price of storage to overshoot. It is only possible to increase the supply of storage and, hence, oil inventories in the short run if more oil is produced, but the best available estimates of the short-run elasticity of the crude oil supply indicate that the curve is highly inelastic (see, e.g., Hamilton 2009; Kilian and Murphy 2014). We return to this point below when we describe the shape of the convenience yield curve.

Periods when the convenience yield is positive imply that the discounted futures price lies below the current spot price. In such cases, the oil futures curve is said to exhibit weak backwardation, and t occurs about 70 per cent of the time in our sample. For instance, the one-year discounted futures price is below the spot price of oil (i.e., the one-year convenience yield is positive) in 70.8 per cent of the months in the sample. On the other hand, a negative convenience yield is possible when oil inventories are plentiful, and it is costly to hold and carry forward oil stocks. Oil refineries that hold inventories must be compensated for doing so by an upward-sloping futures curve. Consistent with the predictions of the theory of storage (examined below), periods during which the convenience yield is negative are precisely those during which oil stocks are plentiful and, consequently, a stockout is unlikely. The marginal benefit of having crude oil inventories on hand is low relative to the marginal cost of physically storing oil.

[9]See footnote 40 in Casassus and Collin-Dufresne (2005).

Overall, the differences in the time-series behavior between the short- and long-term convenience yields suggest the presence of an important slope factor in the term structure of convenience yields. The fact that the front and back ends of the convenience yield curve exhibit different patterns of behavior suggests that it is necessary to account for the relative movements between the two. A slope factor is the natural way to do so.

Shape of the crude oil convenience yield curve. The convenience yield curve (Figure 2b) exhibits a funnel shape, indicating that the sensitivity of long-term convenience yields to movements in the short-term convenience yield decays with the maturity of the oil bond. When short-term convenience yields are high, the curve tends to be downward sloping. The curve tends to be upward sloping when convenience yields are low, unlike the dollar bond curve, which is mainly upward sloping during the sample period. For example, the one-year LIBOR is above the one-month rate 82.8 per cent of the time, while the convenience yield curve is downward sloping almost 35 per cent of the time.

Interestingly, the funnel shape implies that the volatility of convenience yields is a decreasing function of the maturity of the contract. In other words, the term structure of crude oil convenience yields exhibits a Samuelson effect unlike crude oil futures prices. This is a direct consequence of the assumption that the oil convenience yield is mean reverting, which arises naturally from the effect of the supply of the commodity on inventories. Given a shortfall of current oil supply relative to current oil demand, the current marginal benefit of having inventories on hand is expected to be high (i.e., the short-run convenience yield), since the supply of storage takes time to respond to such a shock. The price of storage overshoots in the short run. However, as both crude oil production and inventories are accumulated over the medium term, the value associated with holding inventories in the future (i.e., the long-run convenience yield) decreases.

Similar to the literature on dollar bond yields (see Litterman and Scheinkman 1991), three principal components explain over 99.9 per cent of the variation of the term structure of convenience yields. One can interpret the three principal components as the level, slope and curvature factors (see Figure 3). The first component accounts for 96.31 per cent of the variation, and it increases all yields, thereby changing the level of the convenience yield curve. The factor loadings of the level factor decrease with maturity. This finding reflects the sensitivity of the long-term convenience yields to movements in the short-term convenience yield that decays with maturity. It also explains the funnel shape of the convenience yield curve (see Panel b, Figure 2).

The second principal component loads negatively on short-maturity yields and positively on long-maturity ones, thereby changing the slope of the convenience yield curve. In contrast to the level component, the slope component accounts for only 3.41 per cent of the cross-sectional variation of convenience yields.

The third component has the interpretation of a curvature factor. While it loads negatively on short- and long-term convenience yields (i.e., 1-, 6- and 12-month) and positively on medium-term yields (i.e., 2- and 3-month), it explains less than one per cent of the variation of yields.[10]

Correlations. Table 2 reports the correlations between the principal components of the LIBOR curve, the oil convenience yield curve and the monthly change in the price of oil. Of course, the level, slope and curvature components of the term structure of oil convenience yields are zero by construction, as they are for the LIBOR curve. The low correlations between the components of the convenience yield curve with those of the LIBOR curve and the change in the nominal price of oil indicates that each of these variables represent very different sources of information. For instance, the largest correlation is only 0.37 (between the levels of the LIBOR and the convenience yield curves). Thus, examining the term structure of convenience yields separately reveals information about the crude oil market above and beyond that already contained in the change in the nominal spot price of oil and interest rates.

[10] For comparison, the first three principal components of the LIBOR curve explain 99.61 per cent, 0.35 per cent and 0.03 per cent, respectively, of the variation of yields.

3.2 Crude oil market variables

To examine whether convenience yields contain information about conditions in the physical market for crude oil, we also use data on the following variables. Table 1b shows the summary statistics for these variables.

Production and inventories. We use data on monthly crude oil production and stocks (last day of the month) from PADD 2, the administrative region in the United States oil distribution network where Cushing, Oklahoma (the delivery point for the WTI futures contract) is located. These data are obtained from the U.S. Energy Information Administration.

On average, crude oil stocks are four times the monthly production in the PADD 2 region, and almost three times more volatile (see Table 1b). Both variables are highly persistent. This persistence can also be seen in Figure 4, which depicts the evolution of the two variables' time series. In fact, neither the (log) levels of crude oil production nor inventories seem to be covariance stationary. This feature of the data is likely related to the persistent increase in both the level of inventories and crude oil production observed in recent years. However, both crude oil production and stocks seem to have a common trend, which suggests that the two variables are cointegrated. Using Johansen's (1988) trace test, we reject the null hypothesis of no cointegration between the two variables at the one per cent level.

Given that convenience yields are covariance stationary but inventories are not, we also focus on a normalized version of inventories that has no trend when analyzing the empirical implications of the theory of storage. While it is possible to use the Hodrick-Prescott filter, we use the cointegration relationship between production and crude oil stocks to detrend crude oil inventories, given that it exploits an implication of the theory of storage: higher production of oil should be associated with higher inventories (see Dvir and Rogoff 2014).[11] We estimate the cointegration relationship between stocks and

[11] In addition, it is not possible to reject the null hypothesis that the cycle component of inventories obtained using the Hodrick-Prescott filter contains a unit root.

production using Johansen's (1988) full information maximum-likelihood technique and use the estimated error-correction term as our normalized version of inventories. We obtain the following relationship:

$$abundance_t = \log(stocks_t) - \underset{(0.0741)}{0.3585} \log(production_t), \tag{5}$$

where the standard error of the coefficient is reported in parentheses, $stocks_t$ are the crude oil stocks (thousands of barrels) and $production_t$ is the crude oil production (thousands of barrels per month) from PADD 2.[12] We plot a demeaned time series of $abundance_t$ in Figure 5. Periods during which the error-correction term ($abundance_t$) is below its unconditional mean are interpreted as periods during which crude oil is relatively scarce. The recent period has been characterized by a persistent increase in the level of inventories that can explain why we do not observe a Samuelson effect for crude oil prices. While still relatively persistent when compared with convenience yields, the measure of crude oil abundance is less persistent than PADD 2 stocks (see Table 1b).[13]

Finally, it is possible to use alternative measures of inventories. For example, Hamilton (2009) focuses on total U.S. crude oil inventories, while Kilian and Murphy (2014) proxy global crude oil stocks by scaling U.S. crude oil inventories by the ratio of OECD over U.S. petroleum stocks. However, given that the contract specification requires the owner of the futures contract to take physical delivery, futures prices should reflect the perceived relative scarcity of the amount of crude oil that is available for immediate and future delivery at Cushing. An investor based at a non-Cushing location, for instance, faces additional transportation costs when trying to arbitrage a deviation from the cost-of-carry relationship in equation (2). This friction, in turn, creates additional basis risk that

[12] The cointegration vector is only identified up to a scale factor (see, i.e., Hamilton 1994). Thus, some normalization of the coefficients is required to uniquely identify the long-run relationship between (log) stocks and (log) production. In particular, we normalize the coefficient on $\log(stocks_t)$ in equation (5) to be equal to one, so that $abundance_t$ can be understood as a suitably detrended measure of inventories.

[13] Rather than deseasonalizing the crude oil stocks data, we deal with the (potential) seasonal variation of inventories by estimating long-run predictive regressions (see section 4.2). In particular, we focus on the determinants of the h-month (log) change in crude oil stocks for $h = 3$, 12, which implicitly takes care of quarterly and monthly seasonality, respectively. In addition, unlike natural gas and agricultural commodities, seasonality does not appear to be a first-order driver of crude oil inventories.

is not present when storage and delivery occurs at Cushing (Ederington et al., 2012). We therefore focus on crude oil stocks from PADD 2 when analyzing WTI futures given that data on crude oil stocks at Cushing are not available pre-2004. As shown below, our results are broadly robust to the use of alternative measures of inventories.

Global demand for commodities. Several papers show that the price of oil should be treated as endogenous with respect to global macroeconomic conditions (e.g., Kilian 2008; Kilian 2009; Kilian and Murphy 2012; Lippi and Nobili 2012; Kilian and Murphy 2014; Baumeister and Peersman 2013). We therefore investigate whether the convenience yield curve has information about the index of global real economic activity, rea_t, constructed by Kilian (2009) as a proxy for the global demand of industrial commodities. Kilian's (2009) index of global real economic activity is constructed from data on dry cargo single-voyage ocean freight rates to capture shifts in the demand for industrial commodities. More importantly, rea_t has been shown to have predictive power for the real price of oil (see, e.g., Baumeister and Kilian 2012; Alquist, Kilian and Vigfusson 2013). This index is centered at zero (see Table 1b) and has been normalized to lie between plus one and minus one.

Real price of oil. Given that the spot price of oil is a nominal variable, we deflate its value by the U.S. CPI (seasonally adjusted) obtained from the Bureau of Labor Statistics. The real price of oil is both less volatile and less persistent than the nominal price of oil (see Table 1a and 1b). By tying the value of the spot price of oil to the CPI over the long run, we are assuming that the real price of oil ($rpo_t = s_t - p_t$) is stationary. The stationarity of the real price of oil is consistent with equilibrium models that predict that the U.S.-dollar oil price should follow the aggregate U.S. price level if the nominal price of oil is flexible (e.g., Gillman and Nakov 2009). This prediction stands in contrast to the assumption made by some studies that posit mean reversion in the nominal price of oil. For instance, Schwartz (1997) posits that mean reversion arises naturally in models of commodity price determination given the effect of relative prices of the supply of the commodity, although it may take time for supply to respond to the price movement. This argument is more plausible if it is applied to the real price of oil. Because the price of crude oil is denominated in U.S. dollars, changes in the U.S. price level imply a one-for-one change in the nominal price of crude oil. To the extent that the U.S. price level is non-stationary, so too will be the nominal price of crude oil. For this reason, economic models that include the price of crude oil inevitably need to be specified in terms of the real price of oil (Alquist, Kilian and Vigfusson 2013).

4 Convenience Yields and Crude Oil Stocks

4.1 The Working curve

The main implication of competitive storage models of commodity price determination is the negative and monotonic relationship between the convenience yield and the current level of inventory of a storable commodity, the Working (1933) curve. Periods of relative scarcity of crude oil are related to high convenience yields.

This prediction suggests a natural test for the existence of the Working curve in the oil market using the inventory data. We regress the detrended measure of crude oil stocks $abundance_t$ on the first principal component (i.e., the level component) of the cross-section of convenience yields.[14] The results are reported in the first column of Table 3. As predicted by the theory of storage, the coefficient of the first principal component of the convenience yields on crude oil abundance is statistically significant (at the 1 per cent level) and negative. Moreover, the R^2 statistic is close to 50 per cent. The evidence thus indicates that the first principal component (i.e., the level) of the convenience yield contains information about current crude oil scarcity.

We also examine whether the second principal component (i.e., the slope component) is related to current crude oil stocks. The results are reported in the second column of Table 3. The coefficient on the second principal component is also negative and significant

[14] The results are robust to the use of individual convenience yields and the use of the cycle component of inventories obtained using the Hodrick-Prescott filter.

(at the 10 per cent level), but the explanatory power for current crude oil abundance is low. The R^2 statistic is only 3 per cent. Moreover, this result is robust to including both components as regressors (see the third column of Table 3).

It is unsurprising that the slope component contains little information regarding current convenience yields given that it captures the relative difference between the implicit benefit of physical storage over a long period (i.e., twelve months) versus a short period (i.e., one month). It is, however, possible that they have predictive power over future inventories. We investigate this possibility in the next section.

4.2 The term structure of convenience yields as predictors of crude oil stocks

To test whether the term structure of convenience yields contains information about the future path of inventories in PADD 2, we focus on the following predictive equation:

$$\begin{aligned} y_{t+h} &= \alpha + \boldsymbol{\theta}'\mathbf{f}_t + \boldsymbol{\gamma}'\mathbf{z}_t + \varepsilon_{t+h}, \\ &= \alpha + \boldsymbol{\beta}'\mathbf{x}_t + \varepsilon_{t+h}, \end{aligned} \tag{6}$$

where y_{t+h} is the variable we are trying to predict (e.g., crude oil stocks), $\mathbf{f}_t$ denotes the first two principal components of the term structure of convenience yields, $\mathbf{z}_t$ is a vector of other observable predictors (e.g., $abundance_t$, rea_t, and so on), $\mathbf{x}_t = (\mathbf{f}_t', \mathbf{z}_t')'$, and $\boldsymbol{\beta} = (\boldsymbol{\theta}', \boldsymbol{\gamma}')'$. We are interested in testing the null hypothesis of no predictability ($H_0 : \boldsymbol{\beta} = \mathbf{0}$).[15]

Given that long-run predictive regressions suffer from small-sample biases, we use bootstrap methods to conduct statistical inference on the parameters in equation (6). In particular:

1. The bootstrap algorithm is based on a recursive wild bootstrap design as in Gonçalves

[15] We focus on the in-sample evidence of predictability because we are primarily concerned with examining the evidence regarding the forward-looking nature of convenience yields. In other words, we are most interested in determining whether a predictive relationship exists in the population as suggested by economic theory. In that respect, in-sample tests of predictability are more appropriate than out-of-sample tests given that they are statistically more powerful when the appropriate critical values are used (Inoue and Kilian 2005; Cochrane 2007).

and Kilian (2004; 2007), which deals with the presence of conditional heteroskedasticity in the error term.

2. As in Gospodinov and Ng (2013), we recompute the principal components of the convenience yield curve at each bootstrap iteration to take account of the fact that the principal components, $\mathbf{f}_t$, are generated regressors.[16]

3. As in Kilian (1998), we also bias correct the parameters of the data-generating process prior to bootstrapping the distribution of the test statistics for the null hypothesis of no predictability to deal with the (potential) persistence of the regressors.

4. Finally, given the overlapping nature of this predictive regression, the errors ε_{t+h} have a MA($h-1$) structure when $h > 1$. To address this issue, we use the West (1997) heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator when computing the sequence of Wald statistics for the null hypothesis of no predictability for each bootstrap replication.[17]

More specific details on the bootstrap methods are provided in the appendix.

The parameter estimates (and bootstrap p-values) from equation (6) with $h = 1, 3$, and 12 months for $y_{t+h} = \Delta^h \log(stocks_{t+h})$ (i.e., the h-month log change in crude oil stocks) are reported in Table 4.[18] Figures in bold are statistically significant at the 10 per cent level, the cut-off level used in Gospodinov and Ng (2013). We let $\mathbf{z}_t$, the set of additional predictors, be the detrended measure of crude oil stocks ($abundance_t$), given its role as the error-correction term between crude oil stocks and production, and the real

[16] The standard errors do not need to be adjusted for the fact that $abundance_t$ is used as a generated regressor given that the estimates of the coefficients of the cointegration regression are superconsistent; that is, the estimates converge to their true values at a rate proportional to the sample T rather than the usual $\sqrt{T}$ (see, e.g., Stock 1987).

[17] West (1997) proposes a HAC estimator of the covariance matrix of the parameter estimates that is applicable when the regression disturbance follows a moving average (MA) process of known order. The estimator is $\sqrt{T}$-consistent and is asymptotically more efficient than non-parametric estimators used in the literature such as Newey-West (1987).

[18] These results are robust to the use of $abundance_{t+h}$ as the variable of interest (y_{t+h}) in equation (6), instead of the change in crude oil inventories.

price of oil (rpo_t). The coefficients of the first principal component of convenience yields (i.e., the level) are negative and significant for forecasting horizons $h = 1$ and 3 months. The estimated coefficient is not statistically different from zero for $h = 12$ months. High convenience yields (i.e., a high level in the level of the curve) are not only related to periods of relative scarcity of crude oil today (see results in Table 3), but also related to the scarcity of crude oil in the near future (up to three months). This result is consistent with the fact that, due to the delivery arrangements in the WTI crude oil futures market, there is usually a delay between trades in the futures market and actual delivery of the crude oil (see Ederington et al., 2012).

The estimated coefficient asssociated with the second principal component (i.e., the slope) is negative and statistically significant for $h = 1$ and 3 months for both specifications of equation (6). It also remains negative and significant for $h = 12$ months when the real price of oil is included as a regressor. This result is also consistent with the theory of storage: an upward-sloping convenience yield curve indicates a situation in which future inventories have a higher value than today's inventories, which indicates that oil is expected to be more scarce in the future. This reasoning explains why the sign of the estimated coefficient is negative.

As expected, the error-correction term $abundance_t$ is statistically significant. The negative sign indicates that when crude oil is abundant today, crude oil stocks are expected to decrease in the future. By contrast, we find that inventories tend to increase when real oil prices are high. This effect is statistically significant. All else equal, refiners are willing to store more crude oil when its price is high than when it is low (see Pindyck 2001).

The goal of the bootstrap procedure is to account for the well-known small-sample biases that plague long-run predictive regressions (see, e.g., Mark 1995; Kilian 1999). The distribution of the bootstrap test statistics thus tends to be more conservative than the one implied by conventional asymptotic theory. It is therefore important to stress that the estimated coefficients for the two principal components of the convenience yield curve remain statistically significant even after accounting for the statistical biases mentioned above.

4.3 Sensitivity to other measures of inventories

As noted above, the physical delivery feature of the WTI crude oil futures contract implies that futures prices should reflect the perceived relative scarcity of the amount of crude oil that is available for immediate and future delivery at Cushing. Consequently, the relevant measure of inventories when analyzing WTI futures should be the amount of crude oil stocks in PADD 2. Still, it is possible that convenience yields contain information about alternative measures of inventories such as the total amount of U.S. crude oil inventories used in Hamilton (2009), or the proxy for global crude oil stocks, which scales U.S. crude oil inventories by the ratio of OECD over U.S. petroleum stocks, used in Kilian and Murphy (2014), since they tend to share a common time-series evolution.[19]

Our conclusions are broadly robust to using the alternative inventory data.[20] For example, the coefficients of both the first and second principal component of convenience yields on the (log) change of U.S. crude oil stocks are negative and significant for forecasting horizons $h = 1$ and 3 months. This evidence makes sense insofar as the WTI futures contract represents a claim on physical oil that is deliverable in PADD 2 and the WTI contract is the main U.S. oil benchmark used for pricing in the North American market (see Fattouh 2011).

On the other hand, while the coefficients of the level component of convenience yield on the (log) change of global inventories, as proxied by Kilian and Murphy (2014), are also negative and significant for forecasting horizons $h = 1$ and 3 months, the slope coefficients are not significant. Thus, the term structure of WTI convenience yields contains information about global crude oil inventories, but it is relatively less informative about global stocks than about stocks in North America. We conjecture that this difference

[19] The three measures of inventories (i.e., PADD 2, U.S. and global) are highly correlated. The correlation ranges from 0.84 for PADD 2 and global inventories to 0.92 for U.S. and global inventories.

[20] The regression results are available from the authors upon request.

in the results might be related to the fact that North American (i.e., WTI) and global markets (i.e., Brent crude) are not fully integrated due, in part, to infrastructure logistics (see Fattouh 2007; Büyükşahin et al. 2013). Similarly, consistent with the segmentation story, these results seem to indicate that the level factor can be interpreted as a measure of the global scarcity of crude oil, while the slope factor seems to capture future scarcity of crude oil in the North American market only.

5 What Other Variables Can the Term Structure of Convenience Yields Predict?

Because inventories are jointly determined in equilibrium with supply and demand decisions, it is important to examine whether the term structure of convenience yields contains information about these variables and the observed price of crude oil. We investigate this hypothesis in this section.[21]

5.1 Convenience yields as predictors of crude oil production

We first focus on whether the term structure of convenience yields contains information about the future path of crude oil production in PADD 2. Table 5 reports parameter estimates (and bootstrap p-values) for the coefficients equation (6) with $h = 1, 3$, and 12 months and $y_{t+h} = \Delta^h \log(production_{t+h})$ (i.e., the h-month log change in crude oil production). As in the case of crude oil stocks, we examine the detrended measure of crude oil stocks ($abundance_t$) and the real price of oil (rpo_t) as the set of additional predictors ($\mathbf{z}_t$).

According to theory, the sign of the effect of crude oil scarcity on production is indeterminate. On the one hand, crude oil can be abundant because of an unexpected decrease in current demand relative to current supply. In this case, the marginal benefit of storing oil is small given that there is plenty of oil available. That is, production decreases when

[21] We can replicate the results obtained by Gospodinov and Ng (2013) that convenience yields are related to future U.S. headline inflation, especially the food and energy component. The results are available from the authors upon request.

crude oil is abundant and increases when it is scarce. This effect is consistent with the positive sign of the estimated coefficient associated with the first principal component of convenience yields. In addition, the coefficient is statistically different from zero for all three forecasting horizons.

But, as Kilian and Murphy (2014) observe, if agents expect a shortfall of future oil supply relative to future oil demand, they will increase their demand for crude oil stocks today in anticipation of the shortfall in the net oil supply. On the supply side, the optimal response to this situation is to increase the production of crude oil, although the increase in production is likely to take time given the inelasticity of the oil supply curve. This set of forces creates a situation in which an increase in inventories today is followed by an increase in crude oil production to meet the additional demand for crude oil stocks. The second effect seems to be captured by the positive and significant sign of the error-correction term, $abundance_t$. Of course, it is important to stress that this relationship is a predictive relationship and hence not necessarily causal. It is challenging to disentangle the separate contributions of each effect without a fully structural model.

These results are robust to including the real price of oil as an additional predictor. In particular, the variable rpo_t is positive and statistically significant for $h = 3$ and 12 but not for $h = 1$. This evidence is consistent with the view that production is inelastic in the short term and that, all else equal, oil producers are willing to produce more of a higher-priced commodity than of a lower-priced one.

5.2 Convenience yields as predictors of the global real economic activity

We next examine the information contained in convenience yields about the global real economic activity. As recent research has shown, global real economic activity is reflected in the demand for industrial commodities and, historically, has been an important driver of crude oil prices (see, e.g., Baumeister and Kilian 2012; Alquist, Kilian and Vigfusson 2013). Therefore, global real economic activity should be a good proxy for the demand of crude oil in the PADD 2 region – that is, the relevant but unobserved demand variable for the pricing of WTI futures contracts.

Table 6 reports parameter estimates (and bootstrap p-values) for the coefficients equation (6) with $h = 1, 3$, and 12 months and $y_{t+h} = rea_{t+h}$. In particular, we control for the current level of global real economic activity (rea_t) and the real price of oil (rpo_t) in the set of additional predictors ($\mathbf{z}_t$).

The coefficients of the first principal component of convenience yields (i.e., the level) are not significant for any of the three horizons under consideration. However, the second principal component is positive and significant for $h = 1$ and 3 months. This finding indicates that firms assign a higher value to future inventories than they do to today's inventories (positive slope) due to the expectation of higher global demand for industrial commodities in general and crude oil in particular.

5.3 Convenience yields as predictors of the price of crude oil

Finally, we turn our attention to the price of crude oil. Table 7 reports parameter estimates (and bootstrap p-values) for the coefficients in equation (6) with $h = 1, 3$, and 12 months and $y_{t+h} = rpo_{t+h}$ (i.e., the real price of oil), while Table 8 reports parameter estimates for the case of $y_{t+h} = \Delta^h s_{t+h}$ (i.e., the h-month log change in the nominal spot price of oil). In both cases, we control for the current level of global real economic activity (rea_t) and the real price of oil (rpo_t) in the set of additional predictors ($\mathbf{z}_t$).

Both tables deliver the same message: the first principal component of convenience yields is negatively related to future crude oil prices in real and nominal terms. When convenience yields are high (i.e., crude oil is scarce), crude oil prices are predicted to subsequently fall. Moreover, the informational content of the first component of convenience yields is robust to including rea_t as an additional predictor.

We attribute this effect to the sluggishness of the supply response and the mean reversion of convenience yields and crude oil scarcity. As discussed above, an unexpected increase in demand causes the spot price of oil to overshoot in the short run, given that supply takes time to respond fully to such a change. In the meantime, inventories are drawn down to compensate for the slow adjustment in production, which, in turn, causes the marginal value of holding inventories and, thus, convenience yields to increase. However, as oil suppliers respond by increasing production over the medium term, the spot price of oil falls, which explains the negative and significant sign of the first principal component of convenience yields in Tables 7 and 8.

6 Final Remarks

In this paper, we construct and analyze the term structure of crude oil convenience yields to assess the implications of the theory of storage. Overall, the evidence supports the theory.

This conclusion is based on three main pieces of evidence. First, the cross-section of convenience yields can be explained using the familiar level and slope principal components. As predicted by the theory of storage, the level component is negatively related to U.S. crude oil inventories. This finding is consistent with the existence of a Working curve in the crude oil market. Second, the two components have in-sample predictive power for future crude oil stocks. Third, the term structure of crude oil convenience yields contains information for future crude oil production, an index of global demand for industrial commodities and the price of oil. These results make sense insofar as inventory holdings are jointly determined in equilibrium with production and consumption decisions.

Taken together, this evidence underscores the importance of assessing the implications of the theory of storage for the crude oil market, and shows how one can use the term structure of convenience yields to interpret developments in the fundamentals that drive these markets. Above all, the evidence demonstrates that there is a forward-looking element embedded in convenience yields that contains information about subsequent developments in the crude oil market.

An area that deserves further investigation is the modelling of the risk-premium component in the term structure of convenience yields, as suggested by equation (3). A complete model of the risk premium embedded in oil futures prices would permit us to isolate the expectations component embedded in the convenience yield curve and therefore better understand the implications of the theory of storage in a world with risk-averse agents. Some progress along these lines can be found in, for example, Alquist, Bauer and Diez de los Rios (2014), who propose a joint model of the term structure of U.S. interest rates, convenience yields and the spot price of crude oil.

References

[1] Alquist, R., G.H. Bauer and A. Diez de los Rios. 2014. "Macroeconomic Drivers of Crude Oil Risk Premia," Bank of Canada manuscript.

[2] Alquist, R. and O. Gervais. 2013. "The Role of Financial Speculation in Driving the Price of Crude Oil," *Energy Journal* 34: 35-54.

[3] Alquist R. and L. Kilian. 2010."What Do We Learn from the Price of Crude Oil Futures?" *Journal of Applied Econometrics* 25: 539-573.

[4] Alquist, R., L. Kilian and R.J. Vigfusson. 2013. "Forecasting the Price of Oil." in: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, 2, Amsterdam: North-Holland: 427-507.

[5] Ang, A., S. Dong and M. Piazzesi. 2007. "No Arbitrage Taylor Rules," Columbia University manuscript.

[6] Ang, A., M. Piazzesi and M. Wei. 2006. "What does the yield curve tell us about GDP growth?" *Journal of Econometrics* 131: 359-403.

[7] Bauer, M.D., G.D. Rudebusch and C. Wu. 2012. "Correcting estimation bias in dynamic term structure models," *Journal of Business and Economic Statistics* 30: 454-467.

[8] Baumeister, C., P. Guerin and L. Kilian. 2014. "Do High-Frequency Financial Data Help Forecast Oil Prices? The MIDAS Touch at Work." Bank of Canada Working Paper 2014-11.

[9] Baumeister, C. and L. Kilian. 2012. "Real-Time Forecasts of the Real Price of Oil," *Journal of Business and Economic Statistics* 30: 326-336.

[10] Baumeister, C. and G. Peersman. 2013. "The Role of Time-Varying Price Elasticities in Accounting for Volatility Changes in the Crude Oil Market," *Journal of Applied Econometrics* 28: 1087-1109.

[11] Bekaert, G. and R.J. Hodrick. 2001. "Expectations Hypotheses Tests," *Journal of Finance* 56: 1357-1394.

[12] Bekaert, G., R.J. Hodrick and D. Marshall. 1997. "On biases in tests of the expectation hypothesis of the term structure of interest rates," *Journal of Financial Economics* 44: 309-348.

[13] Bessembinder, H., J.F. Coughenour, P.J. Seguin and M.M. Smoller. 1995. "Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure." *Journal of Finance* 50: 361-375.

[14] Brennan, M.J. 1958. "The Supply of Storage." *American Economic Review* 48: 54-72.

[15] Büyükşahin, B., M.S. Haigh, J.H. Harris, J.A. Overdahl and M. A. Robe. 2008. "Fundamentals, Trader Activity and Derivative Pricing." CFTC working paper.

[16] Büyükşahin, B. and J.H. Harris. 2011. "Do Speculators Drive Crude Oil Futures Prices?" *Energy Journal* 32: 167-202.

[17] Büyükşahin, B., T.K. Lee, J.T. Moser, and MA. Robe. 2013. Physical Markets, Paper Markets and the WTI-Brent Spread. *The Energy Journal* 34: 129-153.

[18] Carter, C.A. and C.L. Revoredo-Giha. 2007. "The Working Curve and Commodity Storage under Backwardation," *American Journal of Agricultural Economics* 89: 864-872.

[19] Casassus, J. and P. Collin-Dufresne. 2005. "Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates," *Journal of Finance* 60: 2283-2331.

[20] Chow, Y.F., M. McAleer and J. Sequeira. 2000. "Pricing of Forward and Futures Contracts," *Journal of Economic Surveys* 14*:* 215–253.

[21] Cochrane, J.H. 2007. "The Dog that Did Not Bark: A Defense of Return Predictability," *Review of Financial Studies* 21: 1533–1575.

[22] Considine, T.J. 1997. "Inventories under Joint Production: An Empirical Analysis of Petroleum Refining." *Review of Economics and Statistics* 79: 493-502.

[23] Dvir, E. and K. Rogoff. 2014. "Demand Effects and Speculation in Oil Markets: Theory and Evidence," *Journal of International Money and Finance* 42: 113-128.

[24] Ederington, L.H., C.S. Fernando, K. Holland, and T.K. Lee. 2012. "Financial Trading, Spot Oil Prices, and Inventory: Evidence from the U.S. Crude Oil Market." University of Oklahoma manuscript.

[25] Estrella, A. and G.A. Hardouvelis. 1991. "The term structure as predictor of real economic activity," *Journal of Finance* 46: 555–576.

[26] Fama, E. and K. French. 1987. "Commodity Futures Prices: Some Evidence on Forecast Power, Premia, and the Theory of Storage." *Journal of Business* 60: 55–73.

[27] Fama, E. and K. French. 1988. "Business Cycles and the Behavior of Metals Prices." *Journal of Finance* 43: 1075–93.

[28] Fattouh, B. 2007. "WTI Benchmark Temporarily Breaks Down: Is it Really a Big Deal?" *Oxford Energy Comment*, Oxford Institute for Energy Studies.

[29] Fattouh, B. 2011. "An Anatomy of the Crude Oil Pricing System." Oxford Institute for Energy Studies Working Paper WPM40.

[30] Fontaine, J.S. and R. Garcia. 2012. "Bond Liquidity Premia," *Review of Financial Studies* 25: 1207-1254.

[31] Galbraith, J.W. and V Zinde-Walsh. 1994. "A simple noniterative estimator for moving average models." *Biometrika* 81: 143-155.

[32] Geman, H. and S. Ohana. 2009. "Forward Curves, Scarcity and Price Volatility in Oil and Natural Gas Markets." *Energy Economics* 31: 576–585.

[33] Gillman, M. and A. Nakov. 2009. "Monetary effects on nominal oil prices." *The North American Journal of Economics and Finance* 20: 239–254.

[34] Gonçalves, S. and L. Kilian. 2004. "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form," *Journal of Econometrics* 123: 89–120.

[35] Gonçalves, S. and L. Kilian. 2007. "Asymptotic and Bootstrap Inference for AR(inf) Processes with Conditional Heteroskedasticity," *Econometric Reviews* 26: 609–641.

[36] Gorton, G., F. Hayashi and G. Rouwenhurst. 2012. "The Fundamentals of Commodity Futures Returns," *Review of Finance* 17: 35–105.

[37] Gospodinov, N. and S. Ng. 2013. "Commodity Prices, Convenience Yields, and Inflation," *Review of Economics and Statistics* 95: 206-219.

[38] Hamilton, J.D. 1994. *Time-Series Analysis.* Princeton University Press, Princeton, New Jersey.

[39] Hamilton, J.D. 2009. "Causes and Consequences of the Oil Shock of 2007–08," *Brookings Papers on Economic Activity*, Spring: 215-261.

[40] Hamilton, J.D. and J.C. Wu. 2014. "Risk Premia in Crude Oil Futures Prices," *Journal of International Money and Finance* 42: 9–37.

[41] Hong, H. and M. Yogo. 2012. "What does futures market interest tell us about the macroeconomy and asset prices?," *Journal of Financial Economics* 105: 473-490.

[42] Inoue, A. and L. Kilian. 2005. "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?," *Econometric Reviews* 23: 371-402.

[43] Johansen, S. 1988. "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control* 12: 231–254.

[44] Joseph, K., S.H. Irwin, P. Garcia. 2014. "Commodity Storage under Backwardation: Does the Working Curve Still Work?" Mimeo. University of Illinois Urbana-Champaign.

[45] Juvenal, L. and I. Petrella. 2014. "Speculation in the Oil Market," *Journal of Applied Econometrics*, http://dx.doi.org/10.1002/jae.2388.

[46] Kilian, L. 1998. "Small-Sample Confidence Intervals for Impulse Response Functions," *Review of Economics and Statistics* 80: 218–230.

[47] Kilian, L. 1999. "Exchange Rates and Monetary Fundamentals: What Do We Learn from Long-Horizon Regressions?," *Journal of Applied Econometrics* 14: 491–510.

[48] Kilian, L. 2008. "The Economic Effects of Energy Price Shocks." *Journal of Economic Literature* 46: 871–909.

[49] Kilian, L. 2009. "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review* 99: 1053–1069.

[50] Kilian, L. and T. K. Lee. 2014. "Quantifying the Speculative Component in the Real Price of Oil: The Role of Global Oil Inventories," *Journal of Applied Econometrics* 42: 71–87.

[51] Kilian, L. and D.P. Murphy. 2012. "Why Agnostic Sign Restrictions are Not Enough: Understanding the Dynamics of Oil Market VAR Models," *Journal of the European Economic Association* 10: 1166–1188.

[52] Kilian, L. and D.P. Murphy. 2014. "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil," *Journal of Applied Econometrics* 29: 454–478.

[53] Lippi, F. and A. Nobili. 2012. "Oil and the Macroeconomy: A Quantitative Structural Analysis," *Journal of the European Economic Association* 10: 1059–1083.

[54] Litterman, R. and J.A. Scheinkman. 1991. "Common Factors Affecting Bond Returns," *Journal of Fixed Income* June: 54-61.

[55] Litzenberger, R. and N. Rabinowitz. 1995. "Backwardation in Oil Futures Markets: Theory and Empirical Evidence." *Journal of Finance* 50: 1517–45.

[56] Mark, N. 1995. "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review* 85: 201–218.

[57] Mishkin, F.S. 1990. "What Does the Term Structure Tell Us about Future Inflation?," *Journal of Monetary Economics* 25: 77–95.

[58] National Petroleum Council. 2004. *Observations on Petroleum Product Supply.* Washington, DC: National Petroleum Council.

[59] Newey, W.K. and K.D. West. 1987. "A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica* 55: 703-708.

[60] Ng, V. and S. Pirrong. 1994. "Fundamentals and Volatility: Storage, Spreads, and the Dynamics of Metals Prices." *Journal of Business* 67: 203–230.

[61] Ng, S. and F. Ruge-Murcia. 2000. "Explaining the Persistence of Commodity Prices," *Computational Economics* 16: 149-171.

[62] Pindyck, R.S. 1994. "Inventories and the Short-Run Dynamics of Commodity Prices." *RAND Journal of Economics* 25: 141-159.

[63] Pindyck, R.S. 2001. "The Dynamics of Commodity Spot and Futures Markets: A Primer." *Energy Journal* 22: 1–30.

[64] Routledge, B., D. Seppi and C. Spatt. 2000. "Equilibrium Forward Curves for Commodities." *Journal of Finance* 55: 1297–1338.

[65] Samuelson, P. A. 1965. "Proof that properly anticipated prices fluctuate randomly," *Industrial Management Review* 6: 41–49.

[66] Schwartz, E.S. 1997 "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging", *Journal of Finance* 52: 923-973.

[67] Spector, K. 2013. "The Post-2008 Commodity Trade: Market Liquidity and the Three "C's"," *Commodities in Focus*, Canadian Imperial Bank of Commerce, May 17.

[68] Stock, J.H. 1987. "Asymptotic properties of least squares estimators of cointegrating vectors," *Econometrica* 55: 113-144.

[69] Stock, J.H. and M.W. Watson. 2003. "Forecasting output and inflation: the role of asset prices," *Journal of Economic Literature* 41: 788-829.

[70] Szymanowska, M., F. de Roon, T. Nijman and R. van den Goorbergh. 2014. "An Anatomy of Commodity Futures Risk Premia." *Journal of Finance* 69: 453-482.

[71] Tang, K. and W. Xiong. 2012. "Index Investment and Financialization of Commodities," *Financial Analysts Journal* 68: 54-74.

[72] Telser, L.G. 1958. "Futures trading and the storage of cotton and wheat," *Journal of Political Economy* 66: 133–144.

[73] Trolle, A.B. and E. S. Schwartz. 2009. "Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives," *Review of Financial Studies* 22: 4423-4461.

[74] West, K.D. 1997. "Another heteroskedasticity- and autocorrelation-consistent covariance matrix estimator," *Journal of Econometrics* 76: 171-191.

[75] Working, H. 1933. "Price Relations Between July and September Wheat Futures at Chicago Since 1885." *Wheat Studies of the Food Research Institute* IX (6): 1297–1338.

[76] Working, H. 1949. "The Theory of the Price of Storage." *American Economic Review* 39: 1254–1262.

Appendix

A Bootstrap Inference

In this appendix, we provide additional details on the bootstrap methods employed in the main text of the paper, which are adapted from Kilian (1999). We want to test the null hypothesis of no predictability ($H_0 : \boldsymbol{\beta} = \mathbf{0}$) in predictive regresions such as

$$\begin{aligned} y_{t+h} &= \alpha + \boldsymbol{\theta}'\mathbf{f}_t + \boldsymbol{\gamma}'\mathbf{z}_t + \varepsilon_{t+h}, \qquad (7) \\ &= \alpha + \boldsymbol{\beta}'\mathbf{x}_t + \varepsilon_{t+h}, \end{aligned}$$

where y_{t+h} is the variable of interest (i.e., the real price of oil), $\mathbf{f}_t$ denotes the first $M(=2)$ principal components of the term structure of convenience yields, $\mathbf{z}_t$ denotes an additional set of observable predictors, and $\boldsymbol{\beta} = (\boldsymbol{\theta}', \boldsymbol{\gamma}')'$ and $\mathbf{x}_t = (\mathbf{f}_t', \mathbf{z}_t')'$.

Given that long-run predictive regressions suffer from well-known small-sample biases, we use bootstrap methods to conduct statistical inference about the parameters in equation (7). The bootstrap algorithm is based on a recursive wild bootstrap design as in Gonçalves and Kilian (2004; 2007), which allows us to deal with the presence of (potential) conditional heteroskedasticity in the error term.

A.1 Bootstrap data-generating process

A valid bootstrap algorithm can be obtained under the auxiliary assumption that the bootstrap data-generating process is well described by the following state-space model. Under the null hypothesis of no predictability, we have

$$\boldsymbol{\delta}_t - \boldsymbol{\mu}_\delta = \mathbf{B}(\mathbf{f}_t - \boldsymbol{\mu}_f) + \mathbf{u}_{\boldsymbol{\delta}t}, \qquad (8)$$

$$y_t - \mu_y = u_{yt}, \qquad (9)$$

$$\begin{pmatrix} \mathbf{f}_t - \boldsymbol{\mu}_f \\ \mathbf{z}_t - \boldsymbol{\mu}_z \end{pmatrix} = \sum_{j=1}^{p} \begin{bmatrix} \boldsymbol{\Phi}_{ff}^{(j)} & \boldsymbol{\Phi}_{fz}^{(j)} \\ \boldsymbol{\Phi}_{zf}^{(j)} & \boldsymbol{\Phi}_{zz}^{(j)} \end{bmatrix} \begin{pmatrix} \mathbf{f}_{t-j} - \boldsymbol{\mu}_f \\ \mathbf{z}_{t-j} - \boldsymbol{\mu}_z \end{pmatrix} + \begin{pmatrix} \mathbf{u}_{ft} \\ \mathbf{u}_{zt} \end{pmatrix}, \qquad (10)$$

where the order of the vector autoregression (VAR) (i.e., the transition equation of the state-space model) in equation (10), p, is chosen using the Akaike information criterion. The mean parameters $\boldsymbol{\mu}_\delta$, μ_y, $\boldsymbol{\mu}_f$ and $\boldsymbol{\mu}_z$ are forced to be the unconditional sample means of $\boldsymbol{\delta}_t$, y_t, $\mathbf{f}_t$ and $\mathbf{z}_t$, respectively.

An estimate of the matrix of factor loadings, $\mathbf{B}$, can be obtained using principal component analysis. In particular, let $\widehat{\mathbf{B}}$ denote such an estimate.

Similarly, let

$$\mathbf{\Phi}^{(j)} = \begin{bmatrix} \mathbf{\Phi}_{ff}^{(j)} & \mathbf{\Phi}_{fz}^{(j)} \\ \mathbf{\Phi}_{zf}^{(j)} & \mathbf{\Phi}_{zz}^{(j)} \end{bmatrix},$$

and $\boldsymbol{\phi} = vec\left[\mathbf{\Phi}^{(1)}, \mathbf{\Phi}^{(2)}, ..., \mathbf{\Phi}^{(p)}\right]$, where vec denotes the column stacking operator, and let $\widehat{\boldsymbol{\phi}}$ denote the ordinary least squares (OLS) estimates of these parameters in equation (10).

A.2 Preliminary bias-correction

Given the potential persistence of the predictive regressors $\mathbf{x}_t = (\mathbf{f}_t', \mathbf{z}_t')'$, the OLS estimates of the VAR dynamics of $\mathbf{x}_t$ in equation (7) are likely to be subject to small-sample biases (see, i.e., Bekaert, Hodrick and Marshall 1997; Bauer, Rudebusch and Wu 2012). For this reason, we follow Kilian (1998) in using bootstrap methods to bias correct $\widehat{\boldsymbol{\phi}}$, prior to bootstrapping the test statistics for the null hypothesis $H_0 : \boldsymbol{\beta} = \mathbf{0}$ in equation (7). The proposed method is as follows. Start by using $\widehat{\boldsymbol{\phi}}$, i.e., the OLS estimates of $\boldsymbol{\phi}$ obtained above, to generate $J = 5,000$ artificial samples of the regressors $\left\{\mathbf{x}_t^{*(j)}\right\}_{t=1}^{T}$ using a recursive wild bootstrap design as in Gonçalves and Kilian (2004; 2007), in conjunction with equation (10).[22] Denote by $\widehat{\boldsymbol{\phi}}^{*(j)}$ the OLS estimate of $\boldsymbol{\phi}$ obtained using the artifical sample j. Then, compute the bias-corrected estimates of the VAR(p) model in equation (10) as $\widehat{\boldsymbol{\phi}}^{bc} = 2\widehat{\boldsymbol{\phi}} - \frac{1}{J}\sum_{j=1}^{J} \widehat{\boldsymbol{\phi}}^{*(j)}$.[23]

A.3 Bootstrap distribution of the Wald statistic under the null

Using $\widehat{\mathbf{B}}$ (the estimates of the factor loadings obtained using principal component analysis), $\widehat{\boldsymbol{\phi}}^{bc}$ (the bias-corrected estimates of $\boldsymbol{\phi}$ obtained in the previous section), and a recursive wild bootstrap design in conjunction with equations (9) and (10), we generate a new set of $J = 5,000$ artificial samples $\left\{\boldsymbol{\delta}_t^{**(j)}\right\}_{t=1}^{T}$, $\left\{y_t^{**(j)}\right\}_{t=1}^{T}$, $\left\{\mathbf{x}_t^{*(j)}\right\}_{t=1}^{T}$.

Further, we follow Gospodinov and Ng (2013) to re-estimate the first M principal components of $\left\{\boldsymbol{\delta}_t^{**(j)}\right\}_{t=1}^{T}$ (i.e., the artificial term structure of convenience yields) for each bootstrap sample j. Let $\left\{\widehat{\mathbf{f}}_t^{**(j)}\right\}_{t=1}^{T}$ denote the re-estimated components for the artificial sample j.[24] By doing so, the resulting bootstrap distribution of the test statistic takes into account the estimated uncertainty associated with the principal components being generated regressors.

[22] We initialize the artificial sample j at their value on the first date from the original sample.

[23] We use Kilian's (1998) adjustment when the resulting bias-corrected dynamics of the VAR(p) process in equation (10) becomes explosive.

[24] As in Gospodinov and Ng (2013), we set the sign of $\widehat{\mathbf{f}}_t^{**(j)}$ to be consistent with the dynamics of $\mathbf{f}_t$ estimated from the original sample.

Finally, these artificial (re-estimated) components are plugged into the predictive regression for the artificial data:

$$y_{t+h}^{**(j)} = \alpha^{**(j)} + \boldsymbol{\beta}^{**(j)\prime}\widehat{\mathbf{x}}_t^{**(j)} + \varepsilon_{t+h}^{**(j)},$$

where $\widehat{\mathbf{x}}_t^{**(j)\prime} = (\widehat{\mathbf{f}}_t^{**(j)\prime}, \mathbf{z}_t^{**(j)\prime})'$ and the Wald test statistic for the hypothesis that H_0 : $\boldsymbol{\beta} = \mathbf{0}$ is saved for each j. Denote by $W_{H_0}^{**(j)}$ the corresponding Wald statistic for the artificial sample j.[25] This gives a random sample $\left\{W_{H_0}^{**(j)}\right\}_{j=1}^{J}$ of observations of the conditional distribution of the Wald statistic for the null hypothesis of no predictability. We can therefore compute the percentage of these artificial observations that exceed the actual test statistic W_{H_0} to compute a bootstrap p-value such as

$$\widehat{p}_J = \frac{1}{J}\sum_{j=1}^{J} \mathbf{1}\left(W_{H_0}^{**(j)} > W_{H_0}\right),$$

where $\mathbf{1}(\cdot)$ is an indicator function. As usual, if the value of this bootstrap p-value falls below the usual 10 per cent, 5 per cent or 1 per cent value, then we will reject the null hypothesis of no predictability at that level.

[25] We use West's (1997) heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator when computing the Wald test statistic which, in turn, requires the estimation of the coefficients of a moving average (MA) process of order $h-1$. We use a modified version of the noniterative approach proposed by Galbraith and Zinde-Walsh (1994) to obtain estimates of the MA process.

Table 1
Summary Statistics

Panel A: Futures price of crude oil and yields

	Mean	Standard Deviation	Skewness	Excess Kurtosis	Autocorrelation 1	Autocorrelation 12
Spot price of crude oil (in US$)	43.14	30.02	0.98	-0.31	0.987	0.841
Futures price of crude oil (in US$)						
1-month	43.27	30.29	0.97	-0.36	0.988	0.852
2-month	43.31	30.50	0.96	-0.40	0.989	0.860
3-month	43.31	30.68	0.95	-0.43	0.989	0.866
6-month	43.20	31.03	0.94	-0.49	0.990	0.879
12-month	42.89	31.25	0.92	-0.59	0.992	0.895
Bond yields (in % per year)						
1-month	3.81	2.49	0.05	-0.91	0.993	0.805
2-month	3.86	2.47	0.04	-0.93	0.995	0.805
3-month	3.90	2.46	0.03	-0.94	0.995	0.803
6-month	4.02	2.40	0.03	-0.95	0.995	0.799
12-month	4.23	2.34	0.06	-0.96	0.993	0.794
Convenience yields (in % per year)						
1-month	4.11	24.85	-0.16	3.22	0.759	0.109
2-month	5.13	21.44	0.11	1.73	0.819	0.143
3-month	5.75	19.31	0.19	1.02	0.845	0.166
6-month	6.61	15.50	0.25	0.43	0.874	0.207
12-month	6.91	11.62	0.23	0.31	0.900	0.236

Note: Data are sampled monthly from April 1989 to June 2013.

Panel B: Crude oil variables

	Mean	Standard Deviation	Skewness	Excess Kurtosis	Autocorrelation 1	Autocorrelation 12
PADD2 production	18.40	5.72	1.88	3.86	0.982	0.917
PADD2 stocks	73.39	13.88	1.39	1.61	0.976	0.789
Crude oil abundance	7.68	0.11	0.41	-0.63	0.925	0.443
Kilian's (2009) index	0.00	0.24	0.44	-0.41	0.958	0.510
(Real) price of oil in 1982-1984 US$	22.49	12.09	0.95	-0.14	0.981	0.787

Note: Our measure of crude oil abundance is defined as $Abundance_t = \log(stocks_t) - 0.3585 \times \log(production_t)$ and it is the error-correction term between crude oil *stocks* (thousand barrels) and crude oil *production* (thousand of barrels per month) from PADD 2, and where the cointegration coefficient has been estimated using Johansen's (1988) full information maximum-likelihood approach. Kilian's (2009) index of global economic activity is constructed from data on dry cargo single-voyage ocean freight rates to capture shifts in the demand for industrial commodities in global business markets. Data are sampled monthly from April 1989 to June 2013.

Table 2
Correlations between components

	b_{1t}	b_{2t}	c_{1t}	c_{2t}	Δs_t
b_{1t}	1				
b_{2t}	0	1			
c_{1t}	0.372	0.152	1		
c_{2t}	0.172	0.011	0	1	
Δs_t	-0.020	-0.113	0.224	0.108	1

Note: Data are sampled monthly from April 1989 to June 2013.

Table 3
Working Curve Estimates

$$Abundance_t = \beta_0 + \beta_1 \times pc1_t + \beta_2 \times pc2_t + \varepsilon_t$$

	(1)	(2)	(3)
constant	7.70	7.70	7.70
(t-stat)	(648.33)	(418.39)	(572.23)
[p-value]	[<0.001]	[<0.001]	[<0.001]
$pc1_t$	-2.21	-	-2.22
(t-stat)	(-8.76)		(-9.50)
[p-value]	[<0.001]		[<0.001]
$pc2_t$	-	-2.51	-2.51
(t-stat)		(-1.81)	(-2.64)
[p-value]		[0.071]	[0.008]
R^2	0.473	0.022	0.494

Note: Data are sampled monthly from April 1989 to June 2013. $Abundance_t = \log(stocks_t) - 0.3585 \times \log(production_t)$ is the (estimated) error-correction term between crude oil *stocks* (thousand barrels) and crude oil *production* (thousand of barrels per month) from PADD 2. The variables $pc1_t$ and $pc2_t$ are, respectively, the first two principal components of the term structure of crude oil convenience yields. t-stats computed using Newey-West (1987) HAC standard errors with six ($\simeq T^{1/3}$) lags are shown in parentheses. Asymptotic p-values are shown in square brackets.

Table 4
Convenience yields as predictors of crude oil stocks in PADD 2

$$\Delta^h \log(stocks_{t+h}) = \beta_0 + \beta_1 \times pc1_t + \beta_2 \times pc2_t + \beta_3 \times abundance_t + \beta_4 \times rpo_t + \varepsilon_{t+h}$$

	$h=1$		$h=3$		$h=12$	
$pc1_t$	**-0.367**	**-0.384**	**-0.643**	**-0.708**	-0.207	-0.347
[p-value]	[<0.001]	[<0.001]	[0.013]	[0.004]	[0.724]	[0.482]
$pc2_t$	**-0.835**	**-0.895**	**-2.561**	**-2.797**	-2.980	**-3.792**
[p-value]	[0.017]	[0.009]	[0.037]	[0.020]	[0.172]	[0.034]
$abundance_t$	**-0.132**	**-0.157**	**-0.369**	**-0.465**	**-0.424**	**-0.682**
[p-value]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[0.065]	[0.005]
rpo_t	-	**0.012**	-	**0.050**	-	**0.157**
[p-value]	-	[0.028]	-	[0.002]	-	[<0.001]
R^2	0.071	0.090	0.148	0.229	0.128	0.502

Note: Data are sampled monthly from April 1989 to June 2013. The variables $pc1_t$ and $pc2_t$ are, respectively, the first two principal components of the term structure of crude oil convenience yields. The variable $abundance_t = \log(stocks_t) - 0.3585 \times \log(production_t)$ is the (estimated) error-correction term between crude oil *stocks* (thousand barrels) and crude oil *production* (thousand of barrels per month) from PADD 2. The variable $rpo_t = s_t - p_t$ is the real price of spot oil. Bootstrap p-values computed using West (1997) HAC standard errors under the assumption that the error term ε_{t+h} follows a MA$(h-1)$ process are shown in square brackets. Figures in bold are statistically significant at the 10 per cent level.

Table 5
Convenience yields as predictors of the production of crude oil in PADD 2

$$\Delta^h \log(production_{t+h}) = \beta_0 + \beta_1 \times pc1_t + \beta_2 \times pc2_t + \beta_3 \times abundance_t + \beta_4 \times rpo_t + \varepsilon_{t+h}$$

	$h=1$		$h=3$		$h=12$	
$pc1_t$	**0.321**	**0.308**	**0.406**	**0.360**	**0.738**	**0.614**
[p-value]	[0.008]	[0.008]	[0.087]	[0.088]	[0.078]	[0.068]
$pc2_t$	0.064	0.018	-0.091	-0.253	-0.932	-1.653
[p-value]	[0.902]	[0.966]	[0.923]	[0.734]	[0.437]	[0.130]
$abundance_t$	**0.147**	**0.128**	**0.247**	**0.181**	**0.692**	**0.463**
[p-value]	[<0.001]	[0.005]	[0.007]	[0.023]	[<0.001]	[0.001]
rpo_t		0.009		**0.034**		**0.139**
[p-value]		[0.183]		[0.004]		[<0.001]
R^2	0.043	0.049	0.126	0.204	0.338	0.721

Note: Data are sampled monthly from April 1989 to June 2013. The variables $pc1_t$ and $pc2_t$ are, respectively, the first two principal components of the term structure of crude oil convenience yields. The variable $abundance_t = \log(stocks_t) - 0.3585 \times \log(production_t)$ is the (estimated) error-correction term between crude oil *stocks* (thousand barrels) and crude oil *production* (thousand of barrels per month) from PADD 2. The variable $rpo_t = s_t - p_t$ is the real price of spot oil. Bootstrap p-values computed using West (1997) HAC standard errors under the assumption that the error term ε_{t+h} follows a MA$(h-1)$ process are shown in square brackets. Figures in bold are statistically significant at the 10 per cent level.

Table 6
Convenience yields as predictors of the global demand of commodities

$$rea_{t+h} = \beta_0 + \beta_1 \times pc1_t + \beta_2 \times pc2_t + \beta_3 \times rea_t + \beta_4 \times rpo_t + \varepsilon_{t+h}$$

	$h=1$		$h=3$		$h=12$	
$pc1_t$	-0.027	-0.032	-0.095	-0.117	-0.873	-1.069
[p-value]	[0.869]	[0.854]	[0.773]	[0.742]	[0.280]	[0.179]
$pc2_t$	**2.009**	**2.009**	**3.487**	**3.485**	4.139	4.311
[p-value]	[0.010]	[0.012]	[0.092]	[0.099]	[0.251]	[0.256]
rea_t	**0.961**	**0.962**	**0.840**	**0.847**	**0.541**	**0.633**
[p-value]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[0.015]	[0.006]
rpo_t		-0.002		-0.007		-0.071
[p-value]		[0.896]		[0.860]		[0.503]
R^2	0.921	0.921	0.701	0.701	0.289	0.300

Note: Data are sampled monthly from April 1989 to June 2013. The variables $pc1_t$ and $pc2_t$ are, respectively, the first two principal components of the term structure of crude oil convenience yields. The variable rea_t is the index of global real economic activity constructed by Kilian (2009). The variable $rpo_t = s_t - p_t$ is the real price of spot oil. Bootstrap p-values computed using West (1997) HAC standard errors under the assumption that the error term ε_{t+h} follows a MA$(h-1)$ process are shown in square brackets. Figures in bold are statistically significant at the 10 per cent level.

Table 7
Convenience yields as predictors of the real price of crude oil

$$rpo_{t+h} = \beta_0 + \beta_1 \times pc1_t + \beta_2 \times pc2_t + \beta_3 \times rpo_t + \beta_4 \times rea_t + \varepsilon_{t+h}$$

	$h = 1$		$h = 3$		$h = 12$	
$pc1_t$	**-0.430**	**-0.481**	**-1.222**	**-1.348**	**-3.021**	**-3.246**
[p-value]	[0.013]	[0.007]	[0.015]	[0.008]	[0.002]	[0.002]
$pc2_t$	0.003	-0.001	-2.401	-2.418	-3.886	-3.582
[p-value]	[0.997]	[0.999]	[0.245]	[0.232]	[0.190]	[0.239]
rpo_t	**0.981**	**0.968**	**0.934**	**0.902**	**0.824**	**0.748**
[p-value]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
rea_t		**0.050**		**0.124**		0.249
[p-value]		[0.035]		[0.077]		[0.283]
R^2	0.968	0.968	0.895	0.897	0.733	0.741

Note: Data are sampled monthly from April 1989 to June 2013. The variables $pc1_t$ and $pc2_t$ are, respectively, the first two principal components of the term structure of crude oil convenience yields. The variable rea_t is the index of global real economic activity constructed by Kilian (2009). The variable $rpo_t = s_t - p_t$ is the real price of spot oil. Bootstrap p-values computed using West (1997) HAC standard errors under the assumption that the error term ε_{t+h} follows a MA($h-1$) process are shown in square brackets. Figures in bold are statistically significant at the 10 per cent level.

Table 8

Convenience yields as predictors of the spot price of crude oil

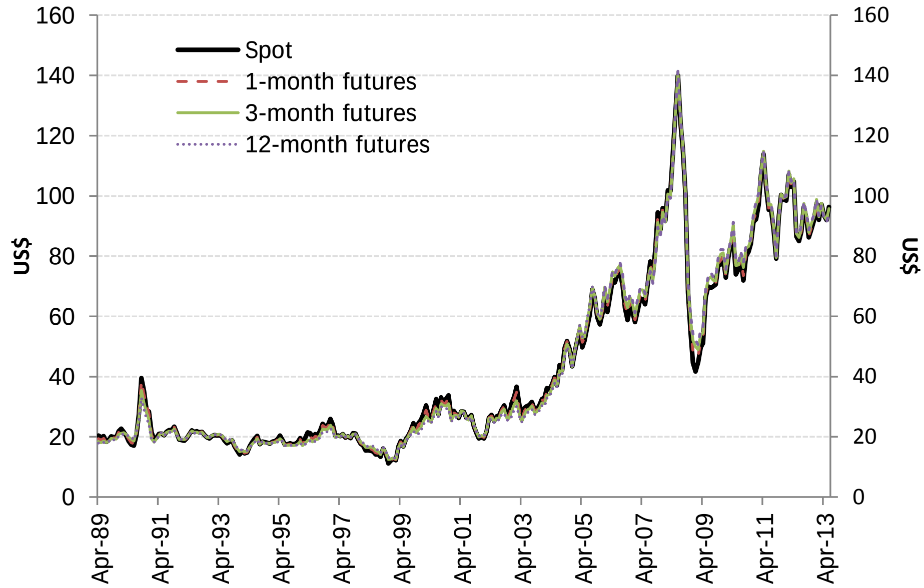
$$\Delta^h s_{t+h} = \beta_0 + \beta_1 \times pc1_t + \beta_2 \times pc2_t + \beta_3 \times rpo_t + \beta_4 \times rea_t + \varepsilon_{t+h}$$

	$h = 1$		$h = 3$		$h = 12$	
$pc1_t$	**-0.417**	**-0.469**	**-1.208**	**-1.337**	**-2.994**	**-3.229**
[p-value]	[0.030]	[0.010]	[0.029]	[0.013]	[0.018]	[0.016]
$pc2_t$	0.097	0.093	-2.243	-2.260	-3.630	-3.314
[p-value]	[0.924]	[0.927]	[0.376]	[0.359]	[0.375]	[0.417]
rpo_t	-0.019	**-0.032**	-0.067	**-0.100**	-0.181	-0.261
[p-value]	[0.343]	[0.042]	[0.290]	[0.040]	[0.371]	[0.107]
rea_t		**0.051**		**0.128**		0.260
[p-value]		[0.043]		[0.099]		[0.368]
R^2	0.029	0.042	0.083	0.106	0.180	0.205

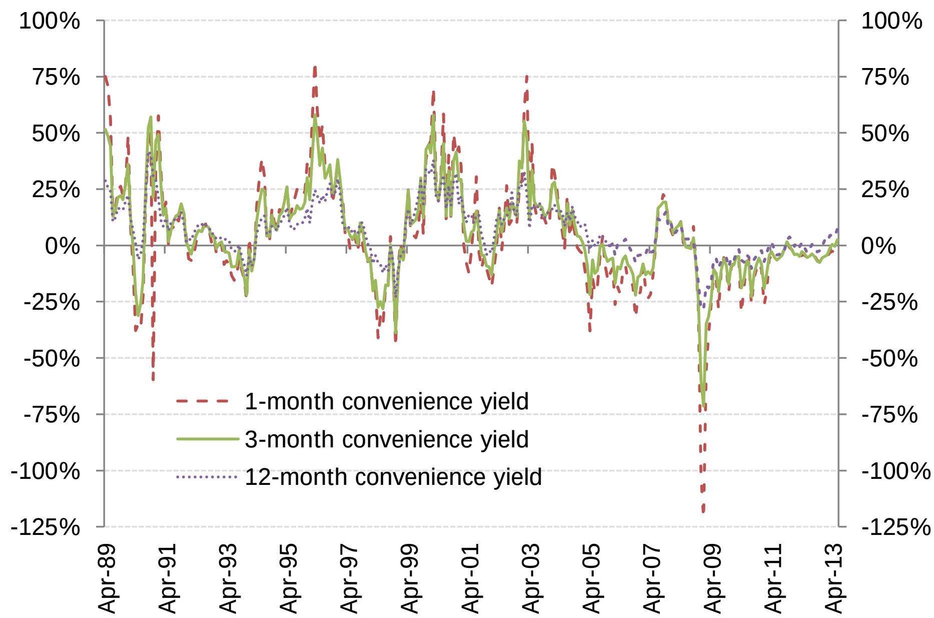
Note: Data are sampled monthly from April 1989 to June 2013. The variables $pc1_t$ and $pc2_t$ are, respectively, the first two principal components of the term structure of crude oil convenience yields. The variable rea_t is the index of global real economic activity constructed by Kilian (2009). The variable $rpo_t = s_t - p_t$ is the real price of spot oil. Bootstrap p-values computed using West (1997) HAC standard errors under the assumption that the error term ε_{t+h} follows a MA$(h-1)$ process are shown in square brackets. Figures in bold are statistically significant at the 10 per cent level.

Figure 1: Crude Oil Futures Prices and Convenience Yields

Panel a: Crude Oil Spot and Futures Prices

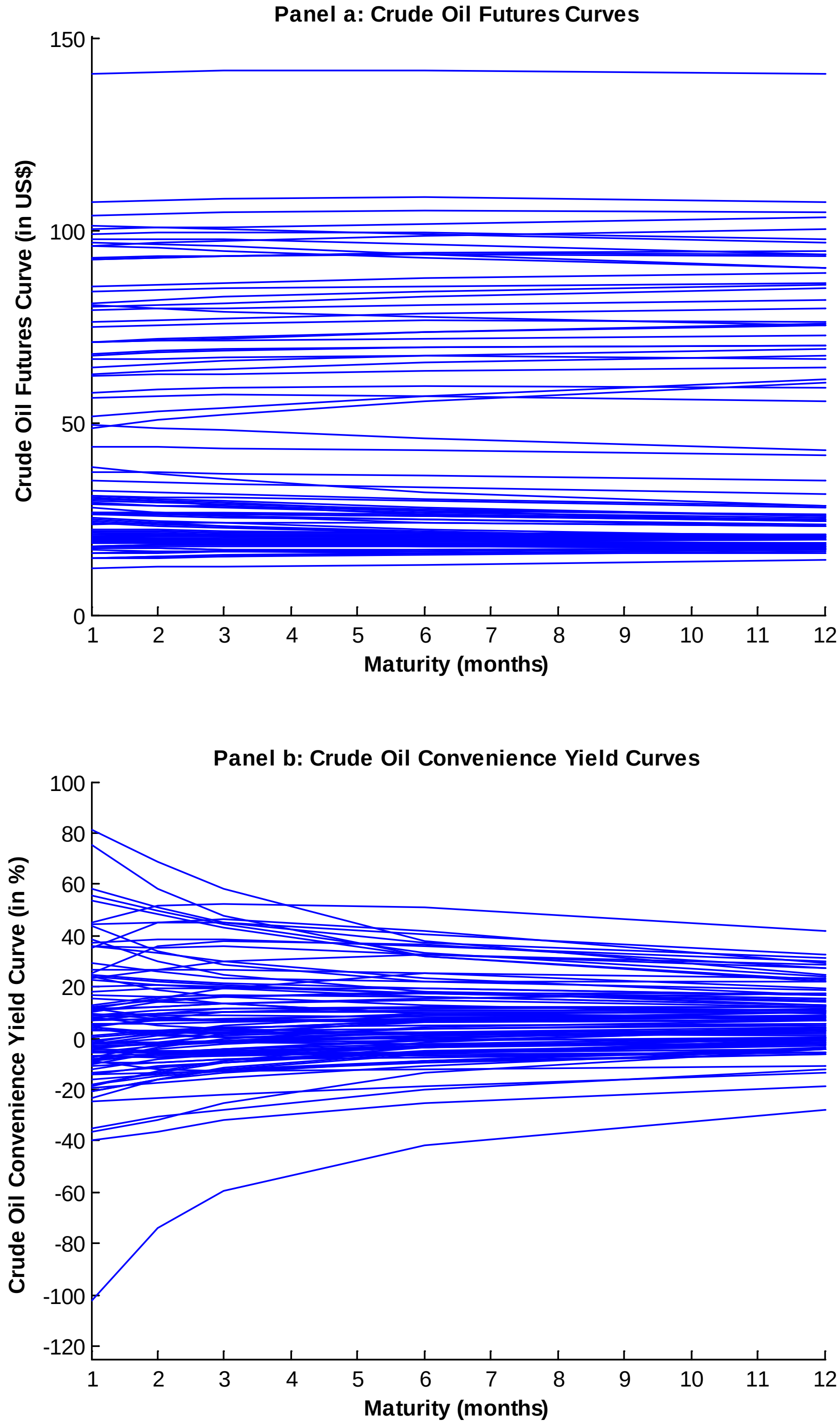


Panel b: Convenience Yields



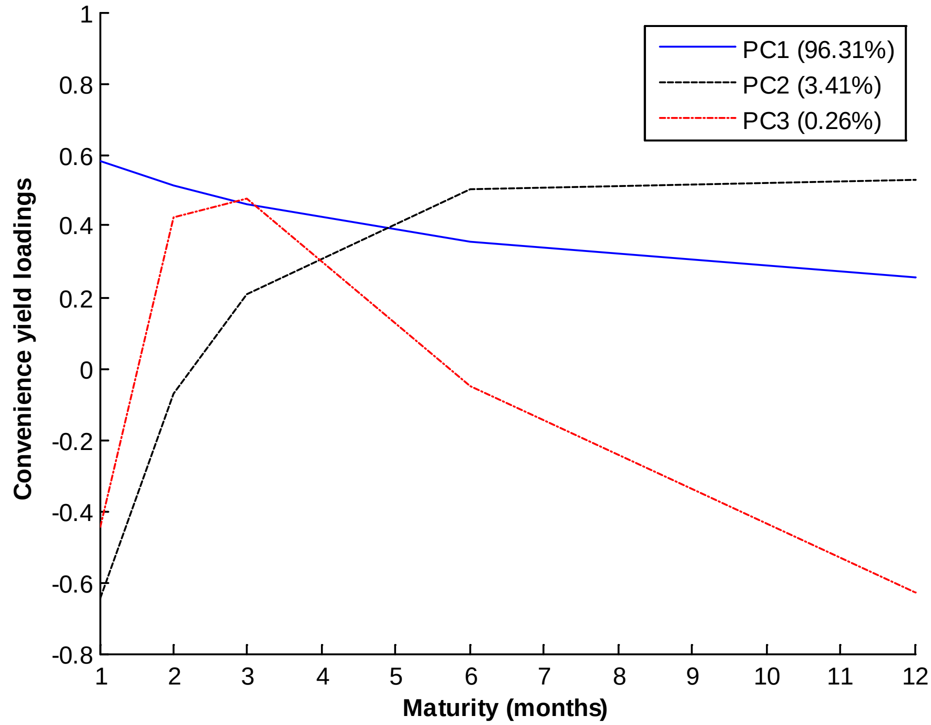
Note: Data are sampled monthly from April 1989 to June 2013. Panel a displays the temporal evolution of end-of-month prices of WTI futures contracts for different maturities. To compute the spot price, we select the futures contract that is closest to delivery. The convenience yields displayed in Panel b are computed using the no-arbitrage relationship (equation (2) in the main text of the paper) $f_t^{(n)} - s_t = ny_t^{(n)} - n\delta_t^{(n)}$, where $f_t^{(n)}$ is the (log) price at time *t* of a futures contract that matures at time *t+n*, s_t is the spot price of oil at time *t*, $y_t^{(n)}$ is the nominal interest rate at which investors can borrow between *t* and *t+n*, and $\delta_t^{(n)}$ is the *n*-period convenience yield.

Figure 2. Term Structures



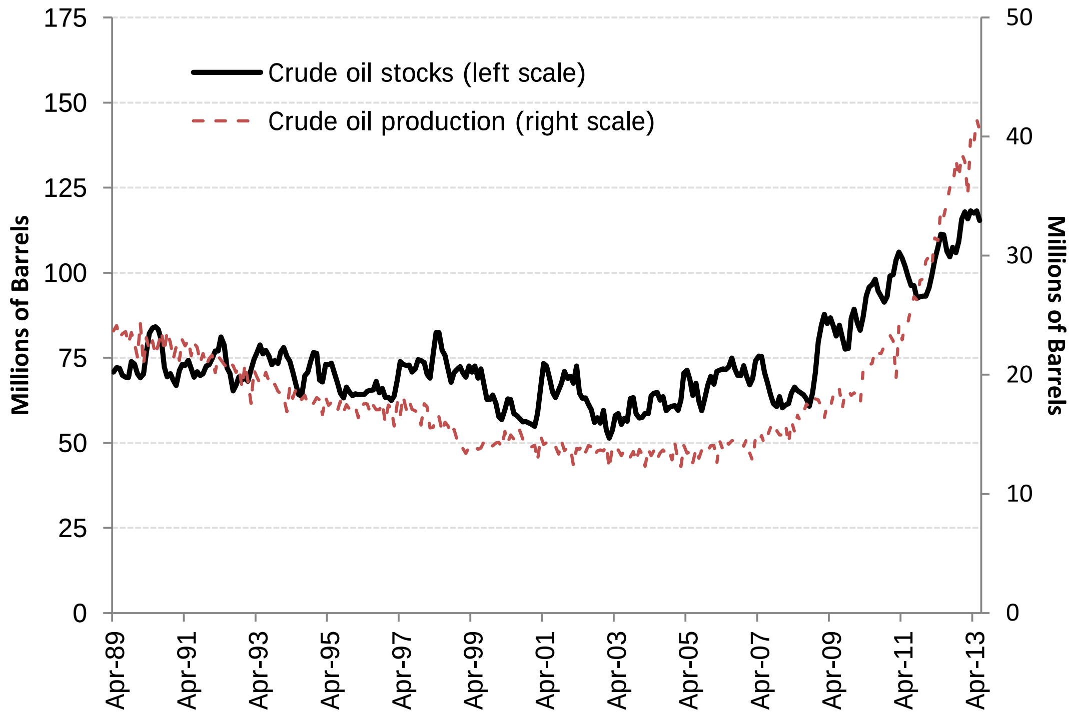
Note: Panel a displays the WTI crude oil futures curves drawn for the end-of-quarter observations. Panel b displays convenience yield curves for end-of-quarter observations, where the convenience yields are computed using the no-arbitrage relationship in equation (2) in the main text of the paper.

Figure 3. Factor Loadings: Convenience Yields



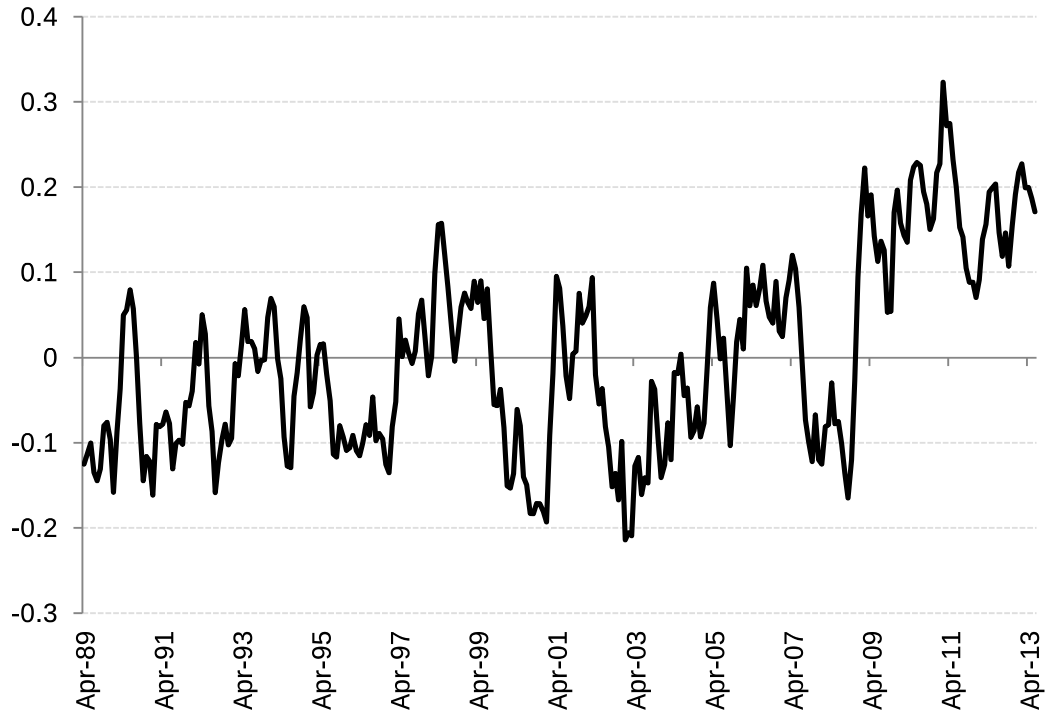
Note: Factor loadings computed using a principal components analysis of the cross-section of convenience yields. The percentage of the variation of the convenience yield curve explained by each of the first three components is reported in the legend of the figure in parentheses. Data are sampled monthly from April 1989 to June 2013.

Figure 4. Stocks and Production of Crude Oil in PADD2



Note: Crude oil stocks refer to stocks on the last day of the month from PADD 2, the administrative region in the United States oil distribution network where Cushing, Oklahoma (the delivery point for the WTI futures contract) is located. Crude oil production refers to the monthly crude oil production in the PADD 2 region. Data are sampled monthly from April 1989 to June 2013.

Figure 5. (Demeaned) Crude Oil Abundance



Note: The variable crude oil $abundance_t = \log(stocks_t) - 0.3585 \times \log(production_t)$ is the estimated error-correction term between crude oil stocks (measured in thousands of barrels) and crude oil production (measured in thousands of barrels per month) from PADD 2, the administrative region in the United States oil distribution network where Cushing, Oklahoma (the delivery point for the WTI futures contract) is located. Data are sampled monthly from April 1989 to June 2013.

Table A1
Convenience yields as predictors of U.S. crude oil stocks

$$\Delta^h \log(stocks^{US}_{t+h}) = \beta_0 + \beta_1 \times pc1_t + \beta_2 \times pc2_t + \beta_3 \times abundance_t + \beta_4 \times rpo_t + \varepsilon_{t+h}$$

	$h=1$		$h=3$		$h=12$	
$pc1_t$	**-0.174**	**-0.180**	**-0.451**	**-0.478**	-0.062	-0.132
[p-value]	[0.017]	[0.013]	[0.006]	[0.007]	[0.876]	[0.718]
$pc2_t$	-0.485	**-0.506**	**-1.410**	**-1.506**	-1.091	-1.502
[p-value]	[0.113]	[0.093]	[0.042]	[0.032]	[0.439]	[0.277]
$abundance_t$	**-0.056**	**-0.064**	**-0.197**	**-0.236**	-0.213	**-0.344**
[p-value]	[0.012]	[0.005]	[<0.001]	[<0.001]	[0.112]	[0.010]
rpo_t	-	0.004	-	**0.020**	-	**0.079**
[p-value]	-	[0.275]	-	[0.060]	-	[0.003]
R^2	0.029	0.034	0.090	0.120	0.083	0.324

Note: Data are sampled monthly from April 1989 to June 2013. The variables $pc1_t$ and $pc2_t$ are, respectively, the first two principal components of the term structure of crude oil convenience yields. The variable $abundance_t = \log(stocks_t) - 0.3585 \times \log(production_t)$ is the (estimated) error correction term between crude oil *stocks* (thousand barrels) and crude oil *production* (thousand of barrels per month) from PADD 2. The variable $rpo_t = s_t - p_t$ is the real price of spot oil. Bootstrap p-values computed using West (1997) HAC standard errors under the assumption that the error term ε_{t+h} follows a MA$(h-1)$ process are presented in square brackets. Figures in bold are statistically significant at the 10 per cent level.

Table A2
Convenience yields as predictors of global crude oil stocks

$$\Delta^h \log(stocks^{global}_{t+h}) = \beta_0 + \beta_1 \times pc1_t + \beta_2 \times pc2_t + \beta_3 \times abundance_t + \beta_4 \times rpo_t + \varepsilon_{t+h}$$

	$h = 1$		$h = 3$		$h = 12$	
$pc1_t$	**-0.142**	**-0.149**	**-0.356**	**-0.383**	0.120	0.061
[p-value]	[0.050]	[0.030]	[0.022]	[0.015]	[0.743]	[0.857]
$pc2_t$	-0.142	-0.342	-1.034	-1.129	-0.810	-1.150
[p-value]	[0.228]	[0.188]	[0.147]	[0.105]	[0.516]	[0.344]
$abundance_t$	**-0.056**	**-0.065**	**-0.187**	**-0.226**	-0.165	**-0.273**
[p-value]	[0.013]	[0.003]	[<0.001]	[<0.001]	[0.151]	[0.026]
rpo_t	-	0.005	-	**0.020**	-	**0.066**
[p-value]	-	[0.199]	-	[0.043]	-	[0.007]
R^2	0.025	0.030	0.089	0.122	0.103	0.320

Note: Data are sampled monthly from April 1989 to June 2013. The variables $pc1_t$ and $pc2_t$ are, respectively, the first two principal components of the term structure of crude oil convenience yields. The variable $abundance_t = \log(stocks_t) - 0.3585 \times \log(production_t)$ is the (estimated) error correction term between crude oil *stocks* (thousand barrels) and crude oil *production* (thousand of barrels per month) from PADD 2. The variable $rpo_t = s_t - p_t$ is the real price of spot oil. Bootstrap p-values computed using West (1997) HAC standard errors under the assumption that the error term ε_{t+h} follows a MA$(h-1)$ process are presented in square brackets. Figures in bold are statistically significant at the 10 per cent level.