Commodity Forward Curve Dynamics with Inventory Information

Marcel Prokopczuk,* and Sebastian Vicedom †

Abstract

In this paper we introduce a new two-factor commodity term structure model for which inventories serve as a second state variable. We derive a closed-form formula for futures prices and empirically analyze the model's properties. Besides being economically appealing, our model also outperforms the well-known Gibson and Schwartz (1990) model in terms of hedging abilities.

JEL classification: G13, C50, Q40

Keywords: Commodities, theory of storage, forward curve dynamics, crude oil, futures markets, hedging, inventory, convenience yield

*Marcel Prokopczuk, School of Economics and Management Leibniz University Hannover, Königsworther Platz 1, 30171 Hannover, Tel: +49-511-762-5225, E-mail: Marcel.Prokopczuk@fmt.uni-hannover.de

†Sebastian Vicedom, School of Economics and Management Leibniz University Hannover, Königsworther Platz 1, 30171 Hannover, Tel: +49-175-318-8950, E-mail: sebastian.vicedom@icloud.com

I Introduction

In almost every commodity market inventories play a central role as they act as buffer for demand and supply shocks. For example, in the U.S. crude oil market the average level in stocks in 2013 equals around 60% of the U.S. yearly production in 2013.[1] One can observe similar effects in the U.S. soybean market, where the amount of U.S. yearly soybean production in stocks in 2013 varies within the year between around 4% to 60%.[2]

Likewise, inventories similarly link futures prices with different maturities under the theory of storage. Gorton et al. (2013) and Symeonidis et al. (2012) find a strong negative dependency between the futures basis and inventory levels using real inventory data. Despite this strong impact, there have been only a very little attempts to consider inventory as a state variable in dynamic term structure models, exploiting inventory dynamics as an additional source of information for pricing these contracts. More frequently, a latent convenience yield process is used to describe the dynamics of commodity futures (e.g., Gibson and Schwartz, 1990).

In this paper, we introduce a new two-factor term structure model in which we explicitly consider inventory levels as a state variable. This is attractive for at least two reasons. First, inventories are observable, whereas convenience yields are not. Second, using inventories directly as a state variable makes the model more easily interpretable from an economic and practical point of view. This is the case because information, such as inventory forecasts is more easily obtained than similar information on the convenience yield.

The contribution of our paper to existing literature is threefold. First, we broaden the class of new models which uses inventory as state variable. Our new two-factor term structure model directly incorporates the empirical findings of Gorton et al.

[1]With ∼250 million barrels of oil produced in the U.S. on average each month and an average storage capacity in the U.S. of ∼1.8 billion barrels. (Source: American Energy Information Association, EIA.)

[2]With ∼3.3 million bushels of soybeans produced in total in the U.S. in 2013 and inventory stocks of ∼1.0 million bushels (March 1), ∼0.4 million bushels (June 1), ∼0.1 million bushels (September 1) and ∼2.1 million bushels (December 1) in U.S. stocks 2013. (Source: United States Department of Agriculture, National Agricultural Statistics Service.)

(2013) and Symeonidis et al. (2012) that the convenience yield is a function of inventory levels. Second, our results confirm the negative dependency between inventory levels and convenience yield for the crude oil market. Further, we find that using non-linear transformations of inventory data such as the log inventory or the inverse of inventory leads to a better pricing and hedging performance compared to using inventory levels directly. This result is in line with the findings of Gorton et al. (2013), who report a non-linear relationship between inventory levels and convenience yield. Third, and most relevant from a practical point of view, the new two-factor commodity term structure outperforms the well-known Gibson and Schwartz (1990) two-factor model in terms of hedging abilities.

Our paper is based on two strands of literature. First, the theory of storage, which provides a linkage between the futures term structure and inventories and, second, the literature which deals with the pricing of contingent claims in commodity markets using dynamic term structure models.

The theory of storage, as introduced, among others, by Kaldor (1939), Working (1948), Brennan (1958), and Telser (1958) links the futures price of a commodity to the level of inventories by introducing the concept of a net value or a yield an agent can earn by physically holding the commodity besides the time value of the commodity. Hence, the convenience yield can be seen as an option to sell the commodity from inventory in the market when prices are high, or to keep it in storage when prices are low. As a consequence, it is also a measure of the future (expected) scarcity of a commodity.

Fama and French (1987), Williams and Wright (1991), Ng and Pirrong (1994), Deaton and Laroque (1992), and Milonas and Thomadakis (1997) further deepen insights about how inventories can link with the futures term structure. Fama and French (1987) study the futures prices of precious metals and find empirical evidence for seasonality in the basis which supports the theory of storage. Williams and Wright (1991) show with a simple model of price determination the existence of a non-linear dependency between inventories and commodity price dynamics as inventory stock-outs lead to peaks in prices as inventory levels cannot be negative. Ng and Pirrong (1994) and Milonas and Thomadakis (1997) show that the convenience yield is inversely related to inventory levels. Deaton and Laroque (1992) link the level of inventory to the future spot price variance since inventories act as buffer stocks which can absorb shocks in demand and supply. They provide evidence that with low inventory levels future spot price volatility rises, as the risk of stock-out and exhaustion of inventories increases.

With increasing inventory data availability, the number of empirical studies on the theory of storage has increased. Dincerle et al. (2003), Geman and Ohana (2009), Symeonidis et al. (2012), Gorton et al. (2013), and Prokopczuk and Wu (2013) test and further detail the linkage between spot and futures prices by using real inventory data. Geman and Ohana (2009) find a negative correlation between spot price volatility and inventory when inventory levels are below their long run average. Symeonidis et al. (2012) confirm that low (high) inventories are associated with forward curves in backwardation (contango) and additionally show that price volatility is a decreasing function of inventory for the majority of commodities. Gorton et al. (2013) show that the convenience yield is a decreasing, non-linear function of inventories and argue that price measures such as the futures basis, past futures returns, and past spot returns reflect the state of the inventories. Prokopczuk and Wu (2013) show that convenience yields are dependent on both commodity specific factors such as inventory levels, but also on systematic factors such as expectations in the consumer price index.

Besides the strand of literature describing the economics of commodity markets, there are also many papers about the pricing of commodity contingent claims. The valuation of commodity derivatives is usually achieved by assuming a specific reduced-form stochastic price dynamics for the relevant state variable in continuous time, i.e. a system of stochastic differential equations (SDE). Perhaps the earliest work in this context is Brennan and Schwartz (1985). The authors develop a one-factor model of commodity futures prices that is capable of valuing the entire term structure in a dynamic and consistent way. To capture the price dynamics more correctly, one can include more stochastic risk factors than just the spot price. Gibson and Schwartz (1990) present such a two-factor model with the latent convenience yield as the second stochastic risk factor. Gabillon (1991) uses as second state variable the long term oil price, and introduces the convenience yield as a function of spot and long term price. Schwartz (1997) extends the model of Gibson and Schwartz (1990) by a third risk factor, the stochastic interest rate. Schwartz and Smith (2000) introduce a two-factor model in which they model the spot price as sum of a mean reverting short term spot price component and a long term equilibrium component. They show that their model is mathematically equivalent to the model of Gibson and Schwartz (1990). Paschke and Prokopczuk (2009) generalize the Schwartz and Smith model by using a continuous autoregressive moving average process (CARMA) instead of the typical Ornstein–Uhlenbeck process for the second latent factor. Geman and Nguyen (2005) present a three-factor model to describe soybean futures price dynamics. They are, to the best of our knowledge, the only authors in the literature explicitly considering inventory information in the context of a dynamic term structure model. In their model, the structural assumption is that the variance of the spot return can be represented as linear decomposition of a constant and the inverse of inventory. Our work differs from Geman and Nguyen (2005), as we model the convenience yield (and not the variance of spot returns), as function of inventory levels.

The outline of this paper is as follow. In Section II we provide an overview of our data set. Section III introduces our new two-factor term structure model and presents two estimation procedures to estimate it. In Section IV we discuss the empirical results and provide a comparison of our new model in- and out-of-sample with relevant term structure models from the literature, such as the Gibson and Schwartz (1990) two-factor model. Finally Section V concludes and provides a prospectus for further research.

II Data set

Our data set contains weekly settlement prices of light sweet crude oil futures contracts traded at NYMEX (symbol CL) and weekly inventory data freely available from the American Energy Information Administration (EIA).[3] The full data set covers the period from 01/01/1990 to 12/28/2012, yielding 1199 data points.

We obtain the futures data from the Commodity Research Bureau (CRB). We consider the 1-month, 3-month, 5-month, 7-month, and 9-month maturity contracts, as these contracts are among the most liquid ones in the WTI crude oil market. Prices are quoted in U.S. dollars. To construct our time series we roll contracts from the nearest to maturity to the next nearest to maturity contract on the last trading day of the month preceding delivery. This procedure avoids expiration effects, as discussed by Fama and French (1987) and low liquidity effects due to thin trading (Symeonidis et al., 2012). We take the daily settlement price of the last day of the week to obtain the weekly price. The contracts are settled in the U.S. as free-on-board (F.O.B.) delivery at any pipeline or storage facility in Cushing, Oklahoma with pipeline access to Enterprise, Cushing storage or Enbridge, Cushing storage. Figure 1 shows the 1-month and 9-month futures contract price series as well as the inventory levels over the entire sample period.

The inventory data are downloaded directly from the website of the EIA. For our analysis we use the weekly U.S. crude oil ending stocks data excluding petroleum products and excluding strategic petroleum reserves (SPR). Stocks in crude oil markets are defined as inventories of fuel stored for future use. The data is reported as of the last day of the period (here week) and expressed in thousands of barrels. Crude oil stocks include those domestic and customs-cleared foreign crude oil stocks held at refineries, in pipelines, in lease tanks, and in-transit to refineries. In particular, it includes the foreign crude oil stocks held in tank farms in Lincoln, Payne, and Creek counties in Oklahoma as well as Alaska in-transit crude oil stocks. Alaska in-transit crude oil stocks include crude oil stocks in-transit by water between Alaska and the other States, the District of Columbia, Puerto Rico, and the Virgin

[3] *http://www.eia.gov/dnav/pet/pet_stoc_wstk_dcu_nus_w.htm.*

Islands, as well as stocks held at transhipment terminals. Petroleum products are made out of already refined and processed crude oil and therefore excluded. SPR are petroleum stocks maintained by the Federal Government for use during periods of major supply interruption and therefore excluded in our data set as well. SPR crude oil stocks include non-U.S. stocks held under foreign or commercial storage agreements.

The focus on U.S. stocks (in contrast to global stocks) for our analysis is in line with the common approach in the existing literature (e.g., Gorton et al., 2013 and Symeonidis et al., 2012). Further, our futures are settled in the U.S. and Geman and Ohana (2009) provide evidence that using either U.S. or global petroleum inventories leads to very similar conclusions when analyzing the futures basis. For our analysis, we construct two further inventory data series to test the effect of non-linearity. We obtain a log inventory data series by taking the logarithm of our inventory series and a inverse inventory series by taking the inverse of inventory levels. We normalize all three data series in order to be of the same order of size.[4]

Table 1 shows summary statistics of the futures prices and inventory levels. Panel A is based on the full data period from 01/01/1990 to 12/28/2012. It covers major global events such as the first Gulf War, the bursting of the dotcom bubble, the attacks of 9/11, the second Gulf War, the rise in global growth until the deep financial crisis, as well as the post-financial crisis recovery period. Panel B is a subset of Panel A and covers the period from 01/01/1990 to 12/31/2004, which leaves us data for an out-of-sample data set Panel C, ranging from 01/01/2005 to 12/28/2012. The maturities of the futures contracts are for each of the time periods comparable with equal mean and standard deviation. Our inventory data process is stationary for all four time periods, as shown by the augmented Dickey–Fuller (ADF) test.

[4]We divide the inventory series by 1 million, the log inventory series by 10 and multiply the inverse inventory series by one 100,000 in order to have compareable magnitudes.

III Model and Estimation

A. Model

In the following we outline the new two-factor term structure model and present two different procedures for estimation. As benchmark we employ the well-known model of Gibson and Schwartz (1990).

As in Gibson and Schwartz (1990), we model the spot price as one factor, but assume the second factor, the convenience yield, to be a linear function of the inventory level. This modeling choice is motivated by the empirical evidence from Gorton et al. (2013) and Symeonidis et al. (2012), who confirm the negative dependency of the futures basis and inventory level with real inventory data. In this new term structure model the commodity is treated as an asset that pays a "dividend" yield decomposed from a constant and a stochastic part. The core assumption in the model is that the stochastic part of this yield is solely driven by the level of inventories. The dynamics of the level of inventories is described by an Ornstein–Uhlenbeck process as the inventory data display mean-reverting characteristics. Hence, we define the dynamics under the real probability measure, $\mathbb{P}$, as:

$$dS_t = (\mu - \delta_t)S_t dt + \sigma_1 S_t dz_1 \tag{1}$$

$$\delta_t = \alpha + \beta I_t \tag{2}$$

$$dI_t = a(m - I_t)dt + \sigma_2 dz_2 \tag{3}$$

with μ, α, β, a, m, σ_1 and σ_2 being constants. dz_1, dz_2 are the increments of correlated Brownian motions, i.e. $dz_1 dz_2 = \rho dt$.

Under the risk neutral measure, $\mathbb{Q}$, the dynamics change to:

$$dS_t = (r - \delta_t)S_t dt + \sigma_1 S_t dz_1^* \tag{4}$$

$$dI_t = a(m^* - I_t)dt + \sigma_2 dz_2^* \tag{5}$$

With z_1^* and z_2^* being again correlated Brownian motions, and m^* being the risk adjusted long term mean reversion level of the inventory process m, with $m^* = m - \frac{\lambda}{a}$.

As we show in the appendix, using the risk free dynamics, the futures price can be calculated as:[5]

$$F(t,T) = S(t)\exp[-\frac{\beta}{a}I(t)(1-\exp^{-a(T-t)}) + B(t,T)] \tag{6}$$

with

$$\begin{aligned} B(t,T) =& (r - \alpha - \beta m^* + \frac{1}{2}\frac{\beta^2}{a^2}\sigma_2^2 - \frac{\beta}{a}\rho\sigma_1\sigma_2)(T-t) \\ &+ (+\frac{\beta}{a}m^* - \frac{\beta^2}{a^3}\sigma_2^2 + \frac{\beta}{a^2}\rho\sigma_1\sigma_2)(1-\exp^{-a(T-t)}) \\ &+ (\frac{1}{4}\frac{\beta^2}{a^3}\sigma_2^2)(1-\exp^{-2a(T-t)}) \end{aligned} \tag{7}$$

Looking at the presented dynamics more closely, one recognizes that the new model nests popular term structure models from the commodity literature. Setting $\alpha = 0$ and $\beta = 1$ and not using the inventory data in estimation procedures, the model reduces to the familiar Gibson and Schwartz (1990) two-factor dynamics as the convenience yield is solely represented by the Ornstein–Uhlenbeck process. Further, when one sets $\beta = 0$, making δ_t constant, the model reduces to the well-known Brennan and Schwartz (1985) one-factor model.

[5]We use the common known results from Jamshidian and Fein (1990) and Bjerksund (1991) to solve the stochastic differential equation.

B. Estimation

As is standard, we estimate the model with a dynamic panel data set of futures prices and the Kalman filter combined with the method of maximum likelihood.[6] The advantages of the Kalman filter is that for a linear Gaussian model, it is optimal in the least-squared sense, exploring time-series and cross-sectional properties of the data simultaneously.

In the following we consider two different estimation procedures. The first procedure simultaneously estimates both state variables, the log spot price and the inventory levels, while including the inventory data together with the futures data in the multivariate time series of observable quantities y_t. We refer to this method as "Procedure I." The second approach does not assume a difference between observed real world inventories and the inventory level state variable. As the real world inventory levels are equal to the inventory state variable, the Kalman filter only provides an estimate for the state variable of the log spot price. In the following we refer to this estimation routine as "Procedure II."

For the first method, "Procedure I," the measurement equation is given by:[7]

$$\underbrace{\begin{bmatrix} lnf(T_1) \\ \vdots \\ lnf(T_N) \\ I_{obs} \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} 1 & -\frac{\beta}{a}(1-\exp^{-aT_1}) \\ \vdots & \\ 1 & -\frac{\beta}{a}(1-\exp^{-aT_N}) \\ 0 & 1 \end{bmatrix}}_{F_t} \underbrace{\begin{bmatrix} X_t \\ I_t \end{bmatrix}}_{Z_t} + \underbrace{\begin{bmatrix} B(T_1) \\ \vdots \\ B(T_N) \\ 0 \end{bmatrix}}_{d_t} + \epsilon_t \tag{8}$$

with the state vector Z_t and ϵ_t being a $(N+1) \times 1$ vector of serially uncorrelated disturbances with

$$E(\epsilon_t) = 0 \qquad Var(\epsilon_t) = H_t \tag{9}$$

[6]See the first application of the Kalman filter in finance Chow (1975), Pagan (1975), and Watson and Engle (1983); more recent studies, among others, include Schwartz (1997), Geman and Nguyen (2005), and Prokopczuk and Wu (2013).

[7]With F_t being a $(N+1) \times 2$ vector and d_t being a $(N+1) \times 2$ vector.

We include the inventory data in the measurement equation to be able to simultaneously estimate the two-factor term structure model. Thereby, the observed inventories I_{obs} are on the left-hand side of the equation, while the estimated inventory level model state variable I_t is on the right-hand side of the equation. This means the for pricing relevant state variable I_t is not forced to be exactly the same as the observed inventories I_{obs} but may differ from it, reflecting a measurement error (e.g., to reflect for possible effects of using U.S. stocks only instead of global stocks for pricing). The Kalman filter combines the information from both futures prices and inventory levels simultaneously to find an optimal set of parameters and state variables (X_t, I_t) that maximizes the log-likelihood score by calculating in each timestep the measurement error and the variance of the measurement error for a respective set of parameters.

The transition equation for "Procedure I" is given by:[8]

$$\underbrace{\begin{bmatrix} X_t \\ I_t \end{bmatrix}}_{Z_t} = \underbrace{\begin{bmatrix} (\mu - \frac{1}{2}\sigma_1^2 - \alpha)\Delta t \\ am\Delta t \end{bmatrix}}_{c_t} + \underbrace{\begin{bmatrix} 1 & -\beta\Delta t \\ 0 & 1 - a\Delta t \end{bmatrix}}_{G_t} \underbrace{\begin{bmatrix} X_{t-1} \\ I_{t-1} \end{bmatrix}}_{Z_{t-1}} + w_t \tag{10}$$

where Δt denotes the time interval between different observations, and w_t is a serially uncorrelated and normally distributed error term with:

$$E[w_t] = 0 \qquad Var[w_t] = W_t = \begin{bmatrix} \sigma_1^2\Delta t & \beta\sigma_1\sigma_2\rho\Delta t \\ \beta\sigma_1\sigma_2\rho\Delta t & \beta^2\sigma_2^2\Delta t \end{bmatrix} \tag{11}$$

The parameter β appears in $Var[w_t]$ due to Equation (2), where we model the dependency between inventory levels and convenience yield.

For the second estimation method "Procedure II" we assume the observed level of inventories I_{obs} to be identical to the inventory state variable. Hence, for the Kalman filter there remains only the spot price state variable, X_t, in the state

[8]With the transition matrice G_t being a 2×2 vector and c_t being a 2×1 vector.

vector Z_t, and we define the measurement equation as:

$$\underbrace{\begin{bmatrix} lnF(T_1) \\ \vdots \\ lnF(T_N) \end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}}_{F_t} \underbrace{\begin{bmatrix} X_t \end{bmatrix}}_{Z_t} + \underbrace{\begin{bmatrix} B(T_1) - \frac{\beta}{a}(1 - \exp^{-aT_1}) I_{obs,t} \\ \vdots \\ B(T_N) - \frac{\beta}{a}(1 - \exp^{-aT_N}) I_{obs,t} \end{bmatrix}}_{d_t} + \epsilon_t \tag{12}$$

The transition equation is qiven by:

$$\underbrace{\begin{bmatrix} X_t \end{bmatrix}}_{Z_t} = \underbrace{\begin{bmatrix} (\mu - \frac{1}{2}\sigma_1^2 - \alpha - \beta I_{obs,t})\Delta t \end{bmatrix}}_{c_t} + \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{G_t} \underbrace{\begin{bmatrix} X_{t-1} \end{bmatrix}}_{Z_{t-1}} + w_t \tag{13}$$

with w_t being a serially uncorrelated and normally distributed error term with:

$$E[w_t] = 0 \qquad Var[w_t] = W_t = \begin{bmatrix} \sigma_1^2 \Delta t \end{bmatrix} \tag{14}$$

We include in the variance of the distributed error term $Var[w_t]$ only components from the spot price transition dynamics, as our state vector contains only X_t.

IV Empirical Results

A. Parameter Estimates

In this section, we report the results when estimating the two-factor term-structure model with both estimation methods, "Procedure I" and "Procedure II," for both in-sample data Panels A and B. To compare results, we use the Gibson and Schwartz (1990) two-factor model as a benchmark model.[9] To explore the effect of non-linearity in the pricing model we estimate the inventory model with three

[9]For the Gibson and Schwartz (1990) two-factor model we use the following dynamics under the risk neutral measure $\mathbb{Q}$:

$$dS_t = (\mu - \delta_t)dtS_t dt + \sigma_1 S_t dz_1^*$$
$$d\delta_t = a(m - \delta_t)dt + \sigma_2 dz_2^*$$

different inventory data series, untransformed inventory levels, log inventories, and inverse inventory. We assume the risk free rate to be constant.[10] For the initial inventory level state variable I_0 of the inventory model we use the observed real inventory level. When estimating the inventory model with the simultaneous estimation method “Procedure I” we observe a sixth standard error, as we include the inventory level data in the Kalman filter estimation procedure. The additional standard error reflects the measurement error between the real world inventory and the inventory state variable estimated by the filter as we estimate the parameter of the processes of the two state variables simultaneously.

Table 2 provides the parameter estimates and the respective value of the log-likelihood score for the Gibson–Schwartz two-factor model and the new two-factor inventory model for both estimation methods and for the different inventory level series data.

There are three main observations. First, in the new two-factor inventory model the parameters are estimated with similar or even higher precision than in the Gibson–Schwartz model. Especially, for the parameter m, the mean reversion level of the convenience yield, or the mean reversion of the inventory level, respectively, one observes in the Gibson–Schwartz model high standard errors and a low degree of precision, while the new inventory model provides significantly sharper estimates, for nearly all cases at the 1% level. The same holds for the drift parameter μ, which traditionally tends to have high standard errors in the Gibson–Schwartz environment; here we observe with the inventory model a slightly higher precision as we find on average higher t-statistics. When comparing the different estimation methods one can conclude that the simultaneous estimation method “Procedure I” leads on average to slightly sharper estimates for the short-term mean m and the volatility parameter σ_2 of the inventory levels dynamics, and to a slightly sharper drift rate μ in comparison to when using the more direct estimation method “Procedure II”.

[10] For the constant risk free rate we use an approximately average interest rate following Schwartz (1997), and to be comparable with Schwartz (1997). Alternatively, one may estimate the constant interest rate r directly with the Kalman filter estimation procedure; results are not reported.

Second, for the two-factor inventory model the parameters α and β, which relate inventories to the convenience yield, are highly significant and estimated for all models significantly at the 1% level. We observe a negative relationship between the inventory level and the convenience yield as the scaling parameter β of the inventory term structure model is negative when using inventory level data or log inventory level data and positive when using the inverse of inventories. This result is in line with the theory of storage and confirms the model free results of Gorton et al. (2013) and Symeonidis et al. (2012). Looking more closely at α and β one sees that the estimates vary across the different inventory data series employed, but also across the different panels. This observation provides a hint that the proportion of the convenience yield described by the inventory process is changing over time, hence the influence of inventories on convenience yield might vary over time.

Third, parameter estimates for a, σ_1, ρ, and μ are comparable in the magnitude of size across the different models, which is in line with what one would expect as the inventory model is an extension of the Gibson–Schwartz model.

Figures 2-5 show the estimated state variables X_t and δ_t for the in-sample data Panel A. One can observe that the log spot price is almost identical between the different two-factor term structure models. For the estimated convenience yield, δ_t, this picture changes. The Gibson–Schwartz model and the simultaneously estimated inventory model are almost identical, while one observes deviating convenience yields when estimating the inventory model with the second estimation method, where the inventory levels are not directly included in the Kalman filter. This is due to the fact, that in the simultaneous estimation method, the inventory state variable for pricing may not exactly equal real world inventory levels, but allow for an error term. Hence, it seems that in the simultaneous estimation method only a small part of the additional information coming from inventories is in the end reflected and that the information coming from futures prices remains, containing the main important information for pricing. Using the direct estimation method "Procedure II" the filter does not adjust the observed real world inventory levels, which is the reason why this model provides a more "smoothed" convenience yield in contrast to the more “spiky” convenience yield structure in the Gibson–Schwartz model. The observation of the “smoothed” convenience yield is in line with Prokopczuk and Wu (2013), who show that inventory levels explain about around one third of the convenience yield extracted with the Gibson–Schwartz model.

B. Comparison of model performance

In the following we compare the performance of the different term structure models in- and out-of-sample. For the in-sample comparison we focus on the information critera and pricing errors. Out-of-sample we focus on comparing the hedging performance of the different models, as this is very important in practice. All in all, we find three main results throughout the different analyses.

First, the new two-factor inventory model is significantly superior in hedging with a significant reduction of the mean and median root mean squared hedging errors compared to the benchmark model of Gibson and Schwartz (1990). Second, non-linear transformations of inventory levels mainly lead to lower hedging errors compared to using inventory level data directly. Especially, using the inverse of inventory data significantly outperforms in all cases compared to using untransformed inventory data. The direct estimation approach “Procedure II” with inverse inventory data produces the best hedging results. Third, the new two-factor model shows a slightly lower in-sample fit compared to the Gibson–Schwartz benchmark model.

In-Sample performance: Akaike and Bayes Information criterion

A first method to compare the models in-sample fit is to compare the Akaike information criterion (AIC) and the Bayes Information criterion (BIC), respectively.[11] Both criteria use the likelihood scores from Table 2 as input and favor the model which needs for the same prediction results the lower amount of parameters.

[11] The Akaike and Bayesian information criteria (AIC and BIC) are defined as:

$$AIC_i = -2ln(L(O_i)) + 2n_i$$

$$BIC_i = -2ln(L(O_i)) + n_i lnT_i$$

Table 3 reports the AIC and BIC for the different models. We observe that the Gibson–Schwartz and the simultaneously estimated inventory two-factor term structure models are similar, which provides a first indication that including two additional parameters does not improve the in-sample properties. Results are comparable across linear or non-linear transformations of the inventory data series.

In-Sample performance: Pricing errors

Table 4 reports the in-sample pricing errors for all five futures contracts in terms of the root mean squared error (RMSE) in monetary terms, the relative root mean squared error (RRMSE) in percentage terms and the mean error for the in-sample error.[12] The observed in-sample pricing errors across the term structure models are in line with literature (e.g., Schwartz, 1997; Schwartz and Smith, 2000) as we observe higher errors on the short end of the futures curve, the 1-month maturity contract, than for the higher maturity contracts across the different models. Typically, pricing errors are higher for short maturity contracts as those contracts are closest to expiration and market participants take additional subjective information into account, which has not been reflected yet in historical prices such as, for example, news events.

Overall, the results of the in-sample pricing errors presented in Table 4 show that the Gibson–Schwartz model performs slightly better, in terms of RMSE, RRMSE, and mean error compared to the two-factor inventory model.

Out-of-Sample performance: Hedging errors

While statistical model fit and in-sample pricing provides a first indication of the usefulness of a commodity term structure model, its out-of-sample hedging ability is clearly a much more relevant aspect for real world applications. To calculate hedging errors we employ the following methodology. At each time step in the

---

with n_i being the number of parameters to be estimated, $ln(L(O_i))$ being the log-likelihood score, and T_i denoting the sample size.

[12]Reported error terms based on log prices. Analyzing models performance based on log prices is most appropriate, as these were used for estimation (see also Schwartz, 1997; Prokopczuk, 2011).

out-of-sample period we construct a portfolio out of two futures contracts, which serves as a hedging portfolio for a third contract. For the portfolio composition we compute the hedge ratios and calculate the number of futures contracts needed for the delta neutral portfolio at each point in time.[13] We can observe the hedge error as the change in the value of the futures portfolios, as a perfect hedge should result in zero change in the portfolio value (Schwartz, 1997). We analyze three different hedging settings. In the first setting, we hedge the long maturity 9-month contract using the 3-month and 7-month month maturity contracts. In the second setting we hedge the 7-month contract with the 1-month and 5-month month maturity contracts. Finally, we construct a set of only short term contracts where we hedge the 1-month contract with the 3-month and 5-month contracts.

Table 5 presents the out-of sample mean hedging error, the standard deviation in the hedging error, the root mean squared hedging error (RMSE), and the mean absolute hedging error (MAE). Tables 6 and 7 report the statistical significance of results reported in Table 5 for the hedging of the 7-month maturity contract on a weekly, respectively 5-weekly basis and shows the reduction in mean and median RMSE of the absolute hedging error.

Summarizing, we find three main results relating to hedging performance. First, the new two-factor inventory model estimated with inverse inventory level data provides significantly superior hedging results in terms of the reduction of root mean squared hedging errors compared to the benchmark Gibson–Schwartz model, but also compared to the new model when estimated with single or log inventory data. This result holds for both the 1-week hedging time horizon and the 5-week hedging time horizon, as well as for both estimation procedures. We can see that the direct estimation "Procedure II" (where real inventory levels remain unchanged) produces

[13]For the two-factor term structures the hedge ratios h_1 and h_2 can be obtained by solving the following set of equations:

$$h_1 \frac{\partial F_1(t,T)}{\partial Xi(t)} + h_2 \frac{\partial F_2(t,T)}{\partial Xi(t)} = \frac{\partial F(t,T)}{\partial Xi(t)}$$

$$h_1 \frac{\partial F_1(t,T)}{\partial I(t)} + h_2 \frac{\partial F_2(t,T)}{\partial I(t)} = \frac{\partial F(t,T)}{\partial I(t)}$$

around 14.81% lower root mean squared hedging errors for the 1-week horizon (15.94% for the 5-week horizon, respectively) compared to the Gibson–Schwartz model. For estimation "Procedure I" the reduction of the mean squared hedging error is slightly less (0.17%) for 1-week (0.21%, for 5-week, respectively) but still significant. Second, the hedging superiority of our new model estimated with inverse inventory level data and estimation "Procedure II" holds as well for the reduction of the median root mean squared error compared to the Gibson–Schwartz model. We document a significant reduction of the median root mean squared error of 6.14% for the 1-week time horizon (4.49% for the 5-week horizon, respectively). Third, a non-linear transformation of inventory levels by using the inverse of inventory leads to significantly lower hedging errors compared to using inventory level data directly. This holds for both the 1-week and 5-week hedging horizon, but also for both estimation procedures.

Figures 6-8 show the hedging error across the different term structure models. Figure 6 focuses on the simultaneous estimation routine and Figure 7 on the direct estimation routine. In Figure 7 the hedging error of the inventory model has less strong peaks than the Gibson–Schwartz benchmark model. Figure 8 explores this in more detail by plotting the hedging error during the peak weeks of the financial crisis from October 2008 to Mai 2009. One can clearly observe that the inventory model has clear advantages and produces fewer hedging errors across all inventory data used than the Gibson–Schwartz benchmark model.

V Conclusion

In this paper, we introduce a new two-factor commodity term structure model which employs inventory levels as a second state variable. This approach is motivated by the findings of Gorton et al. (2013) and Symeonidis et al. (2012), who empirically show that there is a close relationship between the futures basis and inventory levels using real inventory data.

In our empirical study we can confirm the negative dependency between inventory levels and convenience yield for the crude oil market and therefore provide additional evidence in support of the theory of storage. Further, we show that using non-linearly transformed inventory data leads to lower pricing and hedging errors compared to when using inventory data directly. Second, we find that the newly introduced two-factor inventory term structure model exhibits superior hedging characteristics compared to the model of Gibson and Schwartz (1990). This is evidenced by significantly lower mean and median root mean squared errors. Third, from a practical point of view, the two-factor inventory model has the advantage that views on future inventory levels can easily be integrated into the model.

Further research can focus on applying the model to other commodities where inventory levels are of great importance, such as soybean oil, orange juice, coffee, silver, heating oil, and natural gas. To explore the efficiency of futures term structure models for these commodities, one would need to develop the inventory model further so that it incorporates seasonality. Besides this, it would be interesting to investigate the impact of inventory levels and their volatility on options.

Appendix

Futures price $F(t,T)$ of our two-factor model

The stochastic differential equation to price any contingent claim for the given dynamics is given by :

$$-\frac{\partial F}{\partial t} + (r - \alpha - \beta I_t) S_t \frac{\partial F}{\partial S} + a(m^* - I_t)\frac{\partial F}{\partial I} + \frac{1}{2}\sigma_1^2 S_t^2 \frac{\partial^2 F}{\partial S^2} + \frac{1}{2}\sigma_2^2 \frac{\partial^2 F}{\partial I^2} + \rho_1 \sigma_1 \sigma_2 S_t \frac{\partial^2 F}{\partial S \partial \delta} = 0 \tag{15}$$

The terminal conditions for a futures contract are

$$F(S_t, I_t, T) = S(T) \tag{16}$$

The solution of this is known (see for instance Jamshidian and Fein (1990) and Bjerksund (1991)) to be of the form

$$F(t,T) = S(t) \exp^{-A(t,T)I_t + B(t,T)} \tag{17}$$

Substituing this solution into (6) leads to the following set of differential equations

$$A^{'} - aA + \beta = 0 \tag{18}$$

$$B^{'} + (\frac{1}{2}\sigma_2^2)A^2 + (-am^* - \rho\sigma_1\sigma_2)A + (r - \alpha) = 0 \tag{19}$$

with initital conditions $A(T,T) = 0$, and $B(T,T) = 0$.
Equation (9) is elementary and its solution is

$$A(t,T) = \frac{\beta}{a}(1 - \exp^{-a(T-t)}) \tag{20}$$

Plugging (11) into (10) leads to the expression of $B(t,T)$.

$$\begin{aligned} B(t,T) = &-(r - \alpha - \beta m^* + \frac{1}{2}\frac{\beta^2}{a^2}\sigma_2^2 - \frac{\beta}{a}\rho\sigma_1\sigma_2)(T-t) \\ &+ (-\frac{\beta}{a}m^* + \frac{\beta^2}{a^3}\sigma_2^2 - \frac{\beta}{a^2}\rho\sigma_1\sigma_2)(1 - \exp^{-a(T-t)}) \\ &+ (-\frac{1}{4}\frac{\beta^2}{a^3}\sigma_2^2)(1 - \exp^{-2a(T-t)}) \end{aligned} \tag{21}$$

■

References

P. Bjerksund. Contingent claims evaluation when the convenience yield is stochastic: Analytical results. *Working Paper, Norwegian School of Economics and Business Administration*, 1991.

M. Brennan. The supply of storage. *American Economic Review*, 48(1):50–72, 1958.

M. Brennan and E. Schwartz. Evaluating natural resource investments. *Journal of Business*, 58(2):135–157, 1985.

G. C. Chow. *Analysis and control of dynamic economic systems*. John Wiley, New York, 1975.

A. Deaton and G. Laroque. On the behaviour of commodity prices. *Review of Economic Studies*, 59:1–23, 1992.

C. Dincerle, Z. Khokher, and S. Titman. Futures premia and storage. *Working Paper, University of Western Ontario*, 2003.

E. F. Fama and K. R. French. Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. *Journal of Business*, 60:55–74, 1987.

Jacques Gabillon. The term structure of oil futures prices. *Working Paper, Oxford Institute for Energy Studies*, 1991.

H. Geman and V. Nguyen. Soybean inventory and forward curve dynamics. *Management Science*, 51(7):1076–1091, 2005.

H. Geman and S. Ohana. Forward curves, scarcity and price volatility in oil and natural gas markets. *Energy Economics*, 31:576–585, 2009.

R. Gibson and E. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 45(3):959–976, 1990.

G. Gorton, F. Hayashi, and K. G. Rouwenhorst. The fundamentals of commodity futures returns. *Review of Finance*, 17(1):35–105, 2013.

F. Jamshidian and M. Fein. Closed-form solutions for oil futures and European options in the Gibson–Schwartz model: A note. *Working Paper, Merrill Lynch Capital Markets*, 1990.

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

N. T. Milonas and S. Thomadakis. Convenience yields as a call options: An empirical analysis. *Journal of Futures Markets*, 17(1):1–15, 1997.

V. Ng and S. Pirrong. Fundamentals and volatility: Storage, spreads, and the dynamics of metal prices. *Journal of Business*, 67(2):203–230, 1994.

A.R. Pagan. A note on the extraction of components from time series. *Econometrica*, 43:163–168, 1975.

R. Paschke and M. Prokopczuk. Integrating multiple commodities in a model of stochastic price dynamics. *Journal of Energy Markets*, 2(3):47–82, 2009.

M. Prokopczuk. Pricing and hedging in the freight futures market. *Journal of Futures Markets*, 31:440–464, 2011.

M. Prokopczuk and Y. Wu. The determinants of convenience yields. *Working paper*, 2013.

E. Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance*, 52(3):923–973, 1997.

E. Schwartz and J. E. Smith. Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7):893–911, 2000.

L. Symeonidis, M. Prokopczuk, C. Brooks, and E. Lazar. Futures basis, inventory and commodity price volatility: An empirical analysis. *Economic Modelling*, 29 (6):2651–2663, 2012.

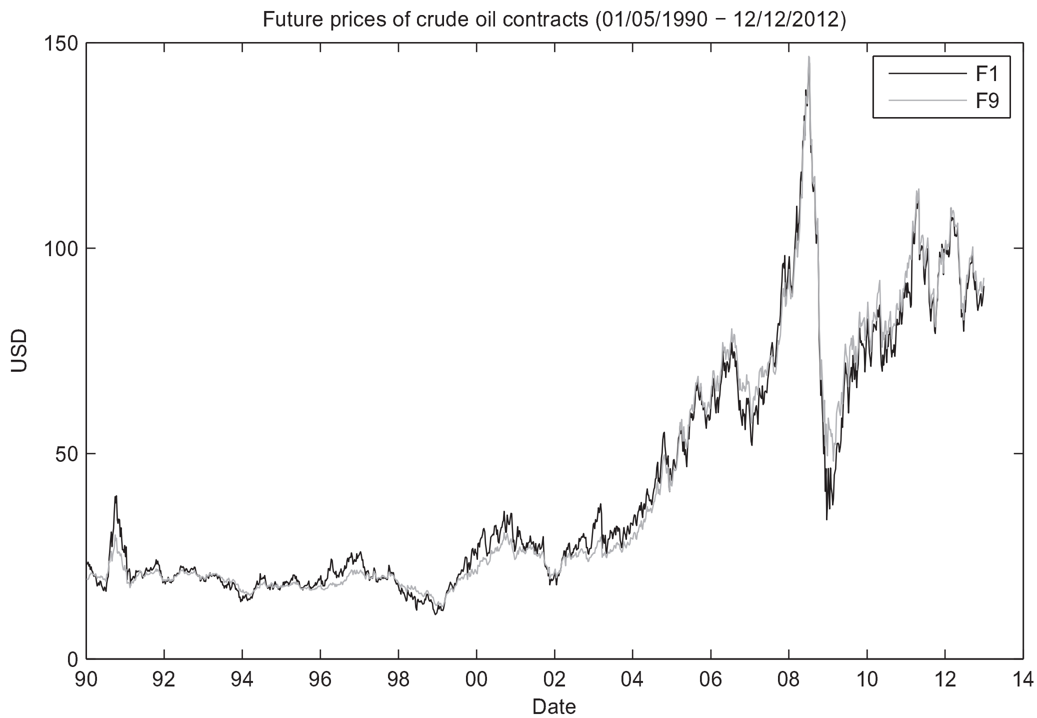
L. Telser. Futures trading and the storage of cotton and wheat. *Journal of Political Economy*, 66(3):233–255, 1958.

W. Watson and F. Engle. Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. *Journal of Econometrics*, 23:385–400, 1983.

J. Williams and B. Wright. *Storage and commodity markets*. Cambridge University Press, New York., 1991.

H. Working. Theory of the inverse carrying charge in futures markets. *Journal of Farm Economics*, 30(1):1–28, 1948.

Figure 1: **Futures prices and inventories**
The first figure shows the prices of the 1-month and 9-month maturity crude oil futures contracts in terms of USD for the entire data period of Panel A (1990–12 with 1199 data points). The second figure shows the U.S. crude oil stocks excluding petroleum products and excluding strategic petroleum reserves (SPR) in terms of billion barrels for the same time period.



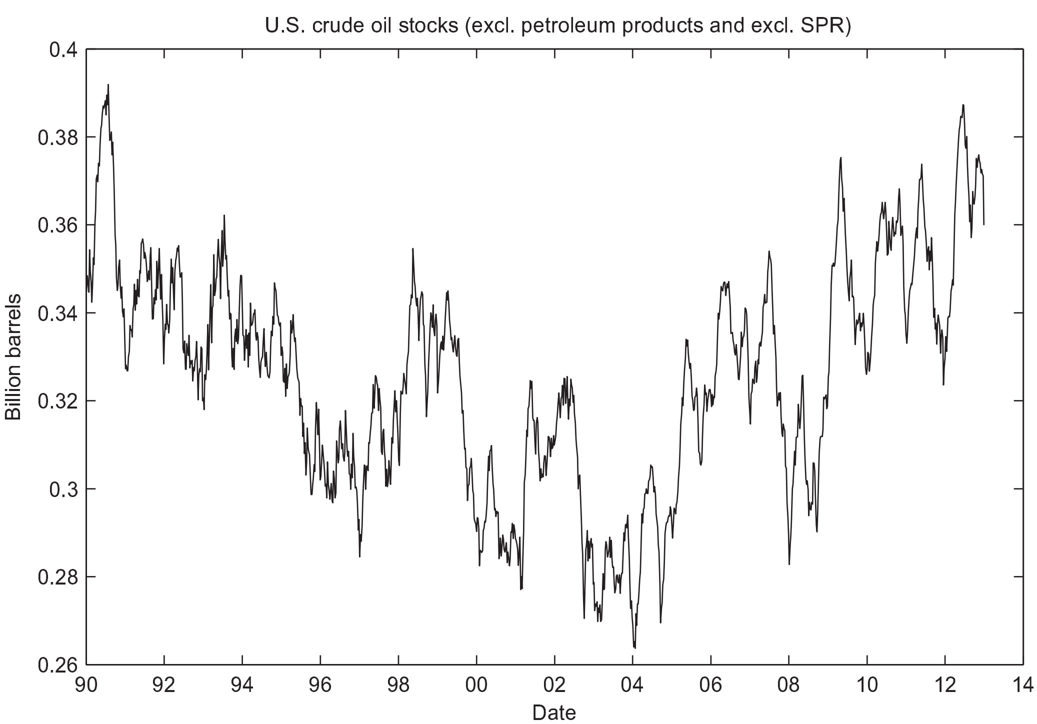


Figure 2: **Estimated state variables: ln X (inventory model with simultaneous estimation method "Procedure I")**

The figure shows the log spot price for the Gibson–Schwartz two-factor model and the two-factor inventory model for the entire in-sample period 1990–12. We estimate our inventory model with three different inventory level data series, the untransformed inventories, log inventories, and inverse inventories. The inventory model is estimated with the simultaneous estimation method "Procedure I," where both the inventory data and the future price data is included in the multivariate time series of observed quantities in the measurement equation of the Kalman filter.

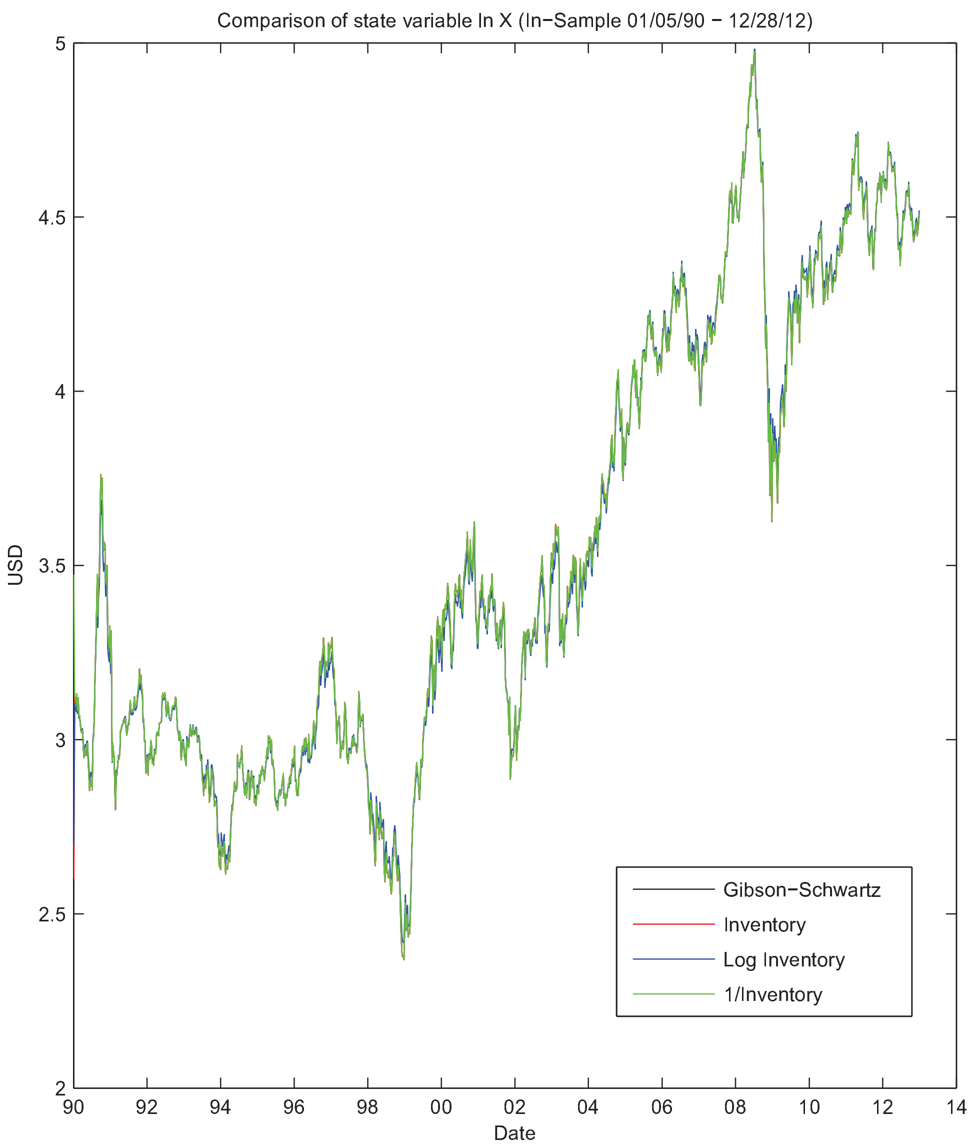


Figure 3: **Estimated state variables: ln X (inventory model with direct estimation method "Procedure II")**

The figure shows the log spot price for the Gibson–Schwartz two-factor model and the two-factor inventory model for the entire in-sample period 1990–12. We estimate our inventory model with three different inventory level data series, the untransformed inventories, log inventories, and inverse inventories. The inventory model is estimated with the direct estimation method "Procedure II," where the inventory state variable is not estimated in the Kalman filter routine as we use for this state variable the real world observed inventory data.

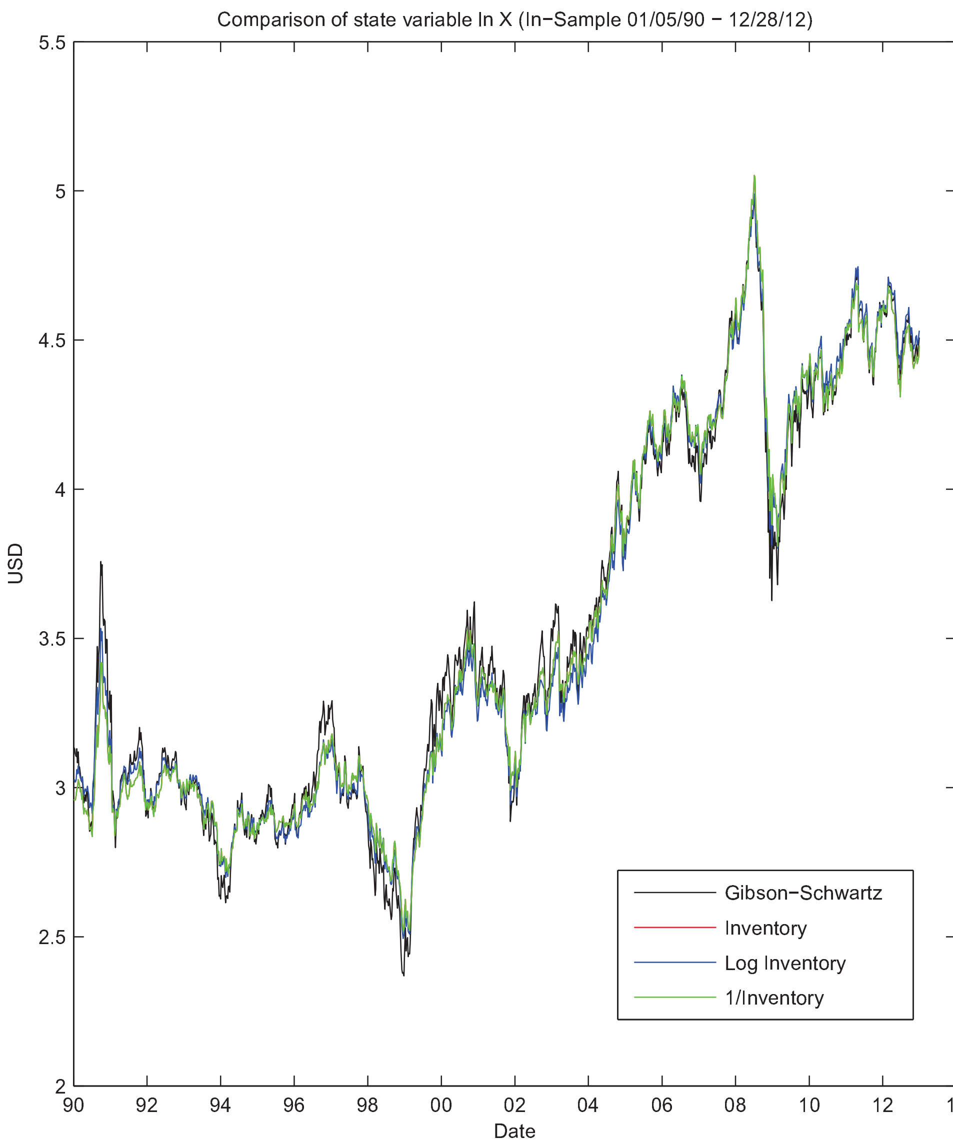


Figure 4: **Estimated state variables: Convenience yield δ_t (inventory model with simultaneous estimation method "Procedure I")**
The figure shows the convenience yield for the Gibson–Schwartz two-factor model and the two-factor inventory model for the entire in-sample period 1990–12. We estimate our inventory model with three different inventory level data series, the untransformed inventories, log inventories, and inverse inventories. The inventory model is estimated with the simultaneous estimation method "Procedure I," where both the inventory data and the future price data is included in the multivariate time series of observed quantities in the measurement equation of the Kalman filter.

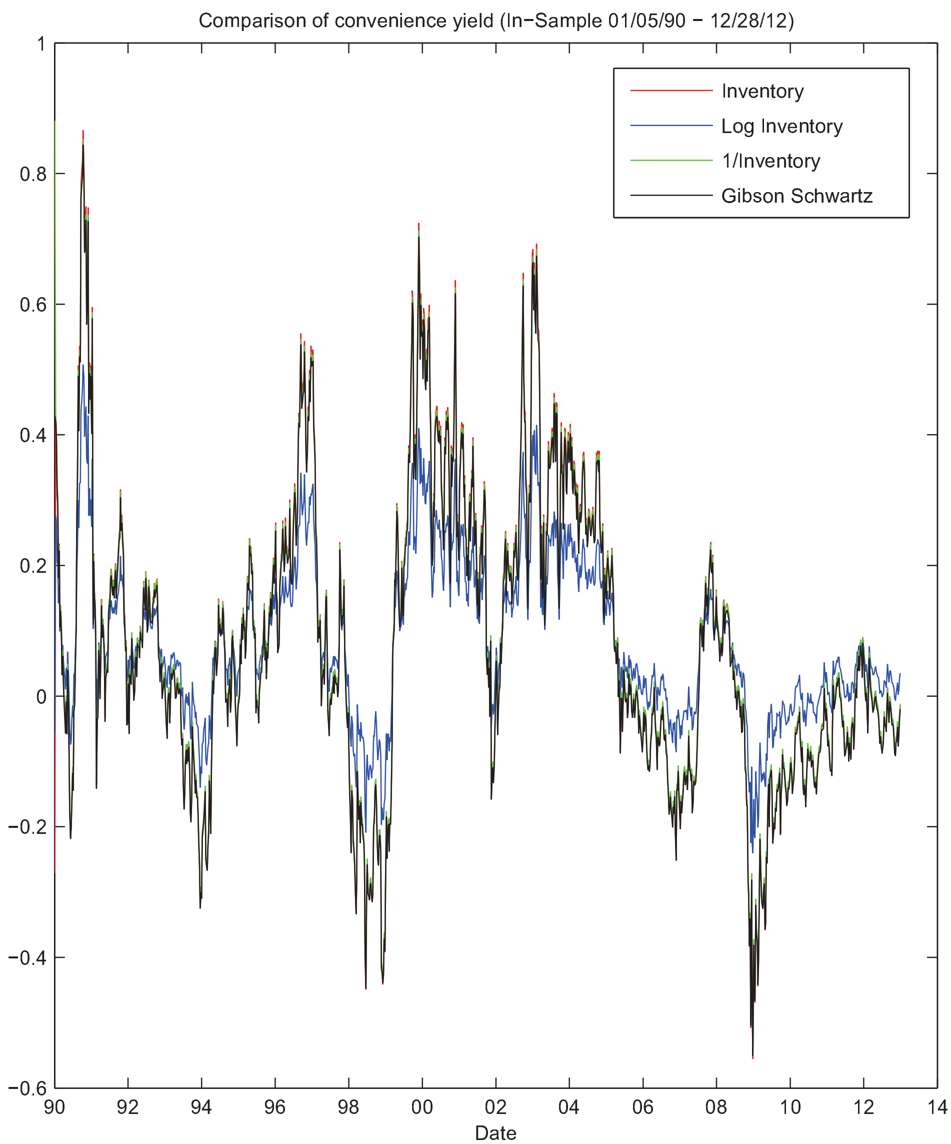


Figure 5: **Estimated state variables: Convenience yield δ_t (inventory model with direct estimation method "Procedure II")**
The figure shows the convenience yield for the Gibson–Schwartz two-factor model and the two-factor inventory model for the entire in-sample period 1990–12. We estimate our inventory model with three different inventory level data series, the untransformed inventories, log inventories, and inverse inventories. The inventory model is estimated with the direct estimation method "Procedure II," where the inventory state variable is not estimated in the Kalman filter routine as we use for this state variable the real world observed inventory data.

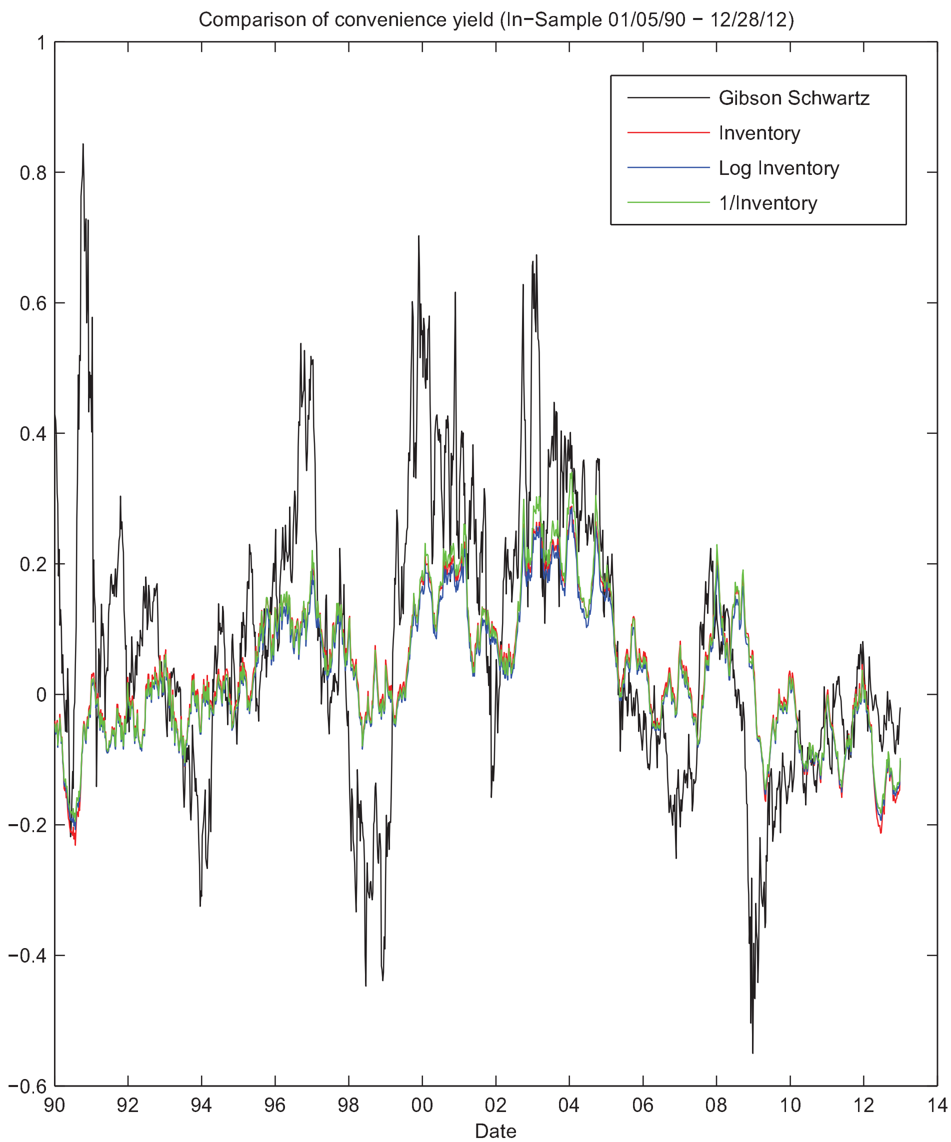


Figure 6: **Hedging error (inventory model with simultaneous estimation method "Procedure I")**
The plot shows the 1-week hedging error for the replication of the 7-month future contract for the out-of-sample period 2005–12. We plot the Gibson–Schwartz two-factor model and the new two-factor inventory model. We hedge the 7-month contract with the 1-month and 5-month maturity contracts. We calibrate the inventory model with three different inventory level data series, the untransformed inventories, log inventories, and inverse inventories. The inventory model is estimated with the simultaneous estimation method "Procedure I," where both the inventory data and the future price data is included in the multivariate time series of observed quantities in the measurement equation of the Kalman filter. We construct the out-of-sample environment by keeping the calibrated parameters constant and employ one-week ahead forecasts based on the Kalman filter.

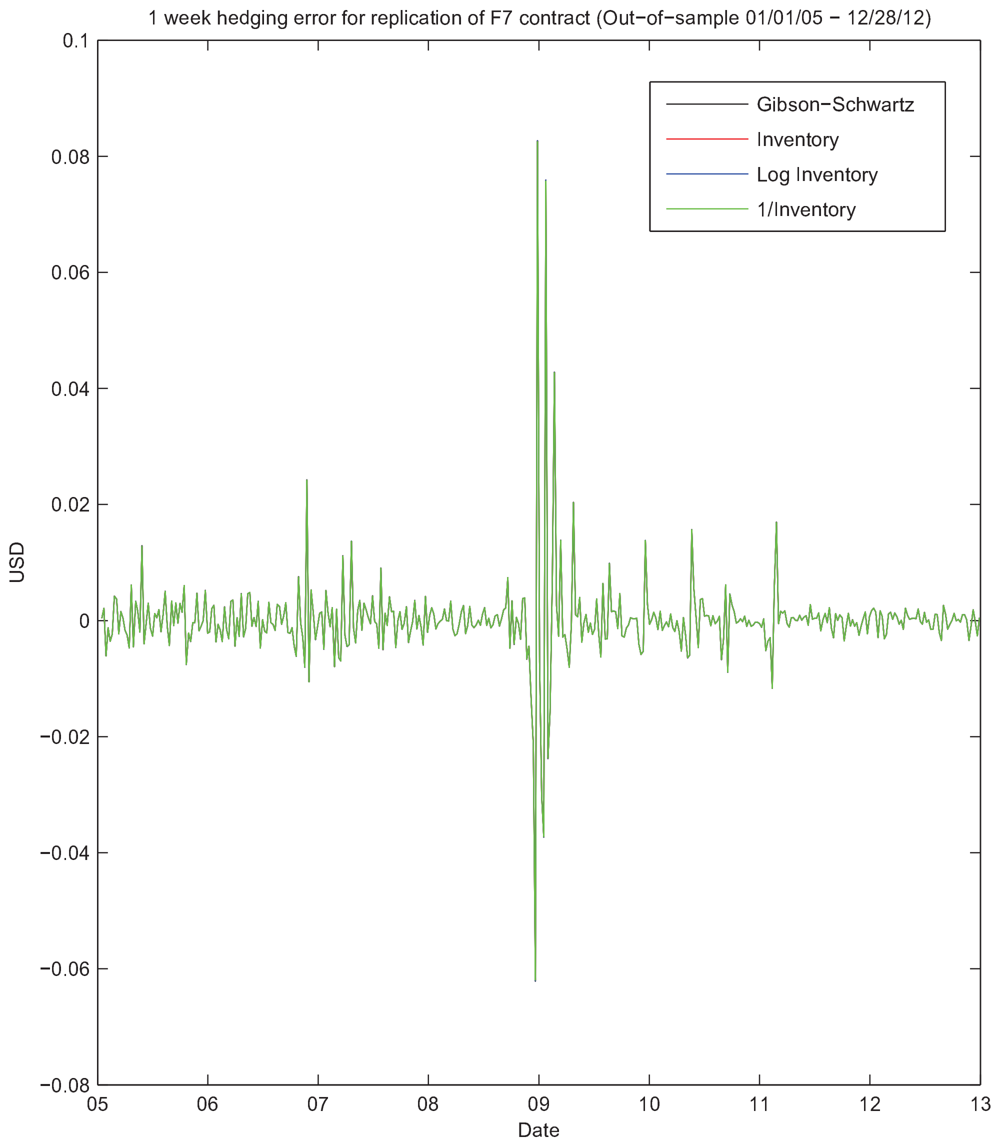


Figure 7: **Hedging error (inventory model with more direct estimation method "Procedure II")**

The plot shows the 1-week hedging error for the replication of the 7-month future contract for the out-of-sample period 2005–12. We plot the three models the Gibson–Schwartz two-factor model and the new two-factor inventory model. We hedge the 7-month contract with the 1-month and 5-month maturity contracts. We calibrate the inventory model with three different inventory level data series, the untransformed inventories, log inventories, and inverse inventories. The inventory model is estimated with the direct estimation method "Procedure II," where the inventory state variable is not estimated in the Kalman filter routine as we use for this state variable the real world observed inventory data. We construct the out-of-sample environment by keeping the calibrated parameters constant and employ one-week ahead forecasts based on the Kalman filter.

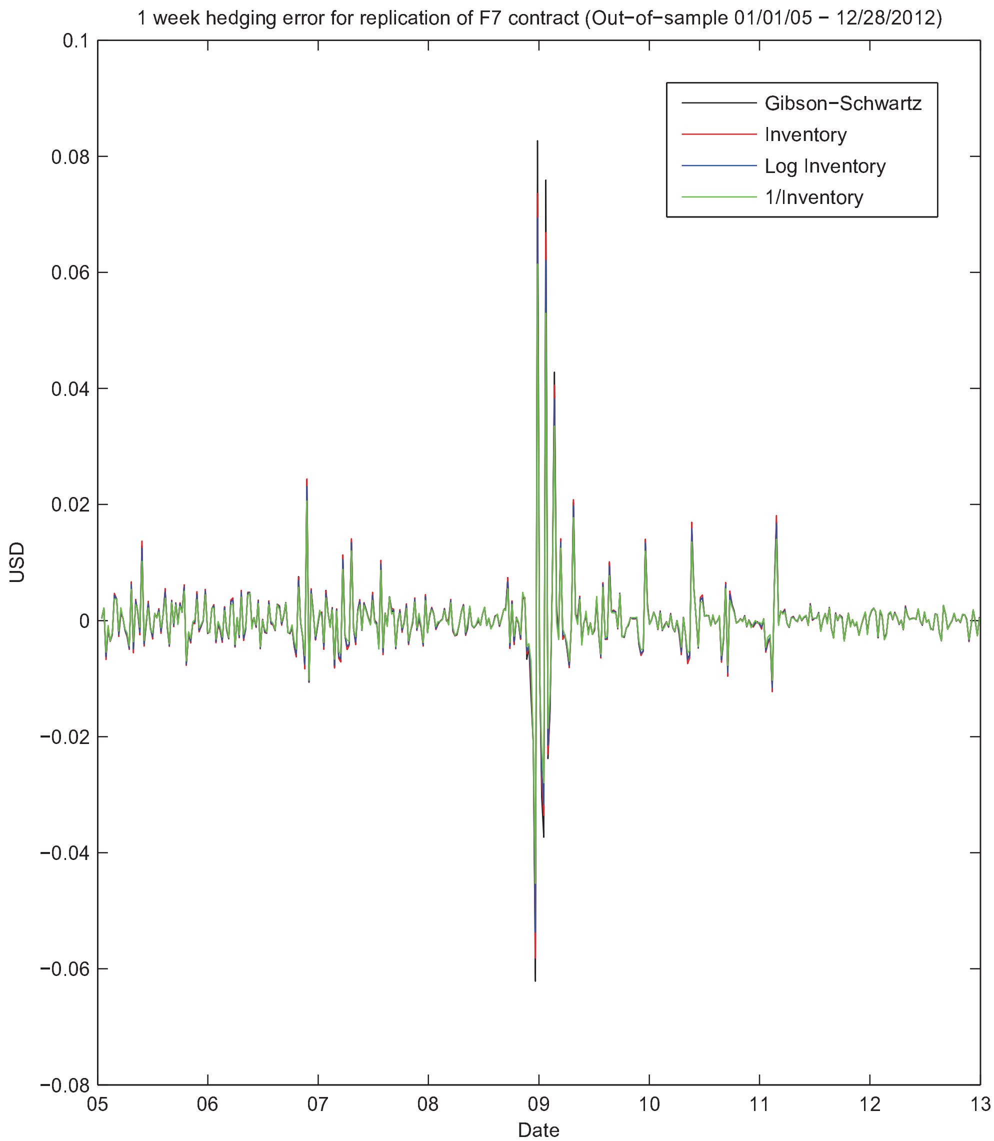


Figure 8: **Hedging error during the financial the crisis (inventory model with more direct estimation method "Procedure II")**
The plot shows the 1-week hedging error for the replication of the 7-month future contract for the peak weeks of the financial crisis from October 2008 – May 2009. We plot the Gibson–Schwartz two-factor model and the new two-factor inventory model. We hedge the 7-month contract with the 1-month and 5-month maturity contracts. We calibrate the inventory model with three different inventory level data series, the untransformed inventories, log inventories, and inverse inventories. The inventory model is estimated with the direct estimation method "Procedure II," where the inventory state variable is not estimated in the Kalman filter routine as we use for this state variable the real world observed inventory data. We construct the out-of-sample environment by keeping the calibrated parameters constant and employ one-week ahead forecasts based on the Kalman filter.

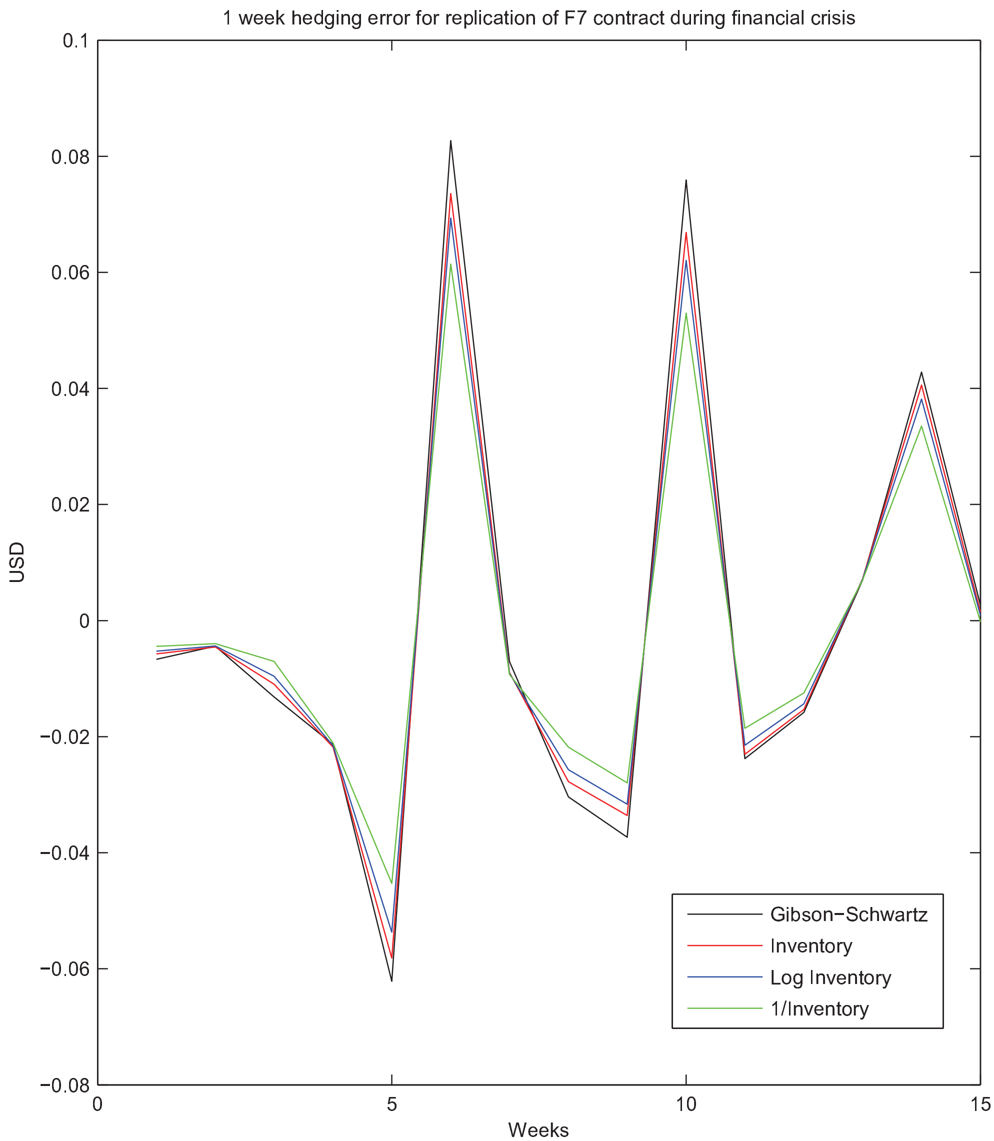


Table 1: **Overview inventory and future data**

This table reports the mean and the standard deviation of the futures prices (in USD) and maturities (in days) for data Panels A (1990–12 with 1199 data points), B (1990–04 with 783 data points) and C (2005–12 with 416 data points). Further, we report the mean and the standard deviation for the three inventory level series: the inventory, the log inventory, and the inverse inventory series. We calculate the three inventory series as follows to make them of comparable magnitude: the inventory series is divided by 1 million, the log inventory series is divided by 10 and the inverse inventory series is multiplied by 100,000. In addition we report the test statistics and the number of lags for the augmented Dickey-Fuller (ADF) to test the inventory data for stationarity. For the ADF test we use a lag length that provides the highest ADF t-statistics.

		Mean price/ inventory level (Standard Deviation)	Mean maturity in days (Standard Deviation)	ADF t-stat (Number of lags)
Panel A				
1990–12 (N=1199)	Inventory	0.3252 (0.0259)	-	-3.616 (8)
	Log Inventory	0.5511 (0.0035)	-	-3.550 (8)
	1/Inventory	0.3095 (0.0250)	-	-3.560 (8)
	F1	42.52 (29.36)	15.22 (8.92)	-
	F3	42.72 (29.84)	76.01 (8.94)	-
	F5	42.73 (30.21)	136.90 (8.98)	-
	F7	42.67 (30.46)	197.83 (8.96)	-
	F9	42.59 (30.63)	258.66 (8.96)	-
Panel B				
1990–04 (N=783)	Inventory	0.3187 (0.0253)	-	-2.749 (8)
	Log Inventory	0.5502 (0.0034)	-	-2.720 (8)
	1/Inventory	0.3157 (0.0251)	-	-2.756 (8)
	F1	23.48 (7.39)	15.33 (8.91)	-
	F3	23.08 (6.95)	76.13 (8.94)	-
	F5	22.68 (6.53)	137.05 (8.96)	-
	F7	22.34 (6.16)	197.94 (8.95)	-
	F9	22.05 (5.86)	258.80 (8.96)	-
Panel C				
2005–12 (N=416)	Inventory	0.3375 (0.0224)	-	-3.2138 (5)
	Log Inventory	0.5527 (0.0029)	-	-3.1724 (5)
	1/Inventory	0.2976 (0.0200)	-	-3.1489 (8)
	F1	78.36 (20.31)	15.00 (8.95)	-
	F3	79.69 (19.51)	75.78 (8.96)	-
	F5	80.47 (19.16)	136.63 (9.01)	-
	F7	80.95 (18.91)	197.62 (9.00)	-
	F9	81.25 (18.69)	258.41 (8.97)	-

Table 2: **Parameter estimates**

This table reports the estimated parameters for the new term structure model and different inventory series. Results are presented for both data Panels A (1990–12 with 1199 data points) and B (1990–04 with 783 data points). All parameters are estimated by the Kalman filter maximum likelihood approach. LL denotes the negative log-likelihood score. Adj. LL denotes the adjusted log-likelihood score when including in the LL calculation only the error and error variance resulting from the future contracts and neglecting the error and variance part resulting from the inventory data. Standard Errors are shown in italic.

			Est. Procedure I						Est. Procedure II					
	Gibson–Schwartz		Inv		LogInv		1/Inv		Inv		LogInv		1/Inv	
	A	B	A	B	A	B	A	B	A	B	A	B	A	B
a	1,3658***	1,3061***	1,4088***	1,3078***	0,0009	1,3056***	1,3656***	1,3077***	1,0828***	0,9076***	1,1668***	1,1456***	1,0575***	1,6256***
	0,0227	*0,0327*	*0,0235*	*0,0295*	*0,0008*	*0,0309*	*0,0227*	*0,0299*	*0,1413*	*0,2843*	*0,0244*	*0,0769*	*0,1378*	*0,0729*
m	0,0648	0,1405*	0,3264***	0,3189***	1,4116***	0,5502***	0,3049***	0,3157***	0,3806***	0,5247***	-3,4400***	-0,4748	2,1496***	-1,2441***
	0,0595	*0,0788*	*0,0030*	*0,0049*	*0,0015*	*0,0008*	*0,0014*	*0,0046*	*0,0135*	*0,0320*	*0,0164*	*0,0061*	*0,0313*	*0,0769*
σ_1	0,3606***	0,3512***	0,3762***	0,3682***	0,3420***	0,3675***	0,3655***	0,3677***	0,2594***	0,2291***	0,2785***	0,2299***	0,2597***	0,2556***
	0,0070	*0,0088*	*0,0077*	*0,0062*	*0,0065*	*0,0084*	*0,0072*	*0,0091*	*0,0054*	*0,0072*	*0,0054*	*0,0051*	*0,0049*	*0,0064*
S_0	3,1352***	3,1323***	2,5999***	2,6716***	2,6994***	2,6766***	3,4737***	2,6780***	2,9574***	2,9811***	2,9965***	2,9773***	2,9560***	3,0058***
	0,0466	*0,0484*	*0,0321*	*0,0270*	*0,0302*	*0,0244*	*0,0411*	*0,0266*	*0,0341*	*0,0956*	*0,0379*	*0,0292*	*0,0333*	*0,0327*
λ_2	0,0310	0,1478	0,0003	-0,0075	0,0010	-0,0011	0,0009	0,0090	0,0982***	0,1697***	-4,6350***	-1,1778***	1,8659***	-2,5623***
	0,0812	*0,1030*	*0,0042*	*0,0064*	*0,0007*	*0,0010*	*0,0015*	*0,0061*	*0,0101*	*0,0339*	*0,0049*	*0,0802*	*0,2449*	*0,0333*
σ_2	0,3850***	0,4146***	0,0010***	0,0021***	0,0001***	0,0000***	0,0001	0,0021***	0,2054***	0,0000	0,0350***	0,0114**	0,1917***	0,1409
	0,0082	*0,0126*	*0,0000*	*0,0001*	*0,0000*	*0,0000*	*0,0000*	*0,0002*	*0,0547*	*0,0660*	*0,0043*	*0,0057*	*0,0478*	*0,1191*
r	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600	0,0600
μ	0,2041***	0,2601***	0,2437***	0,3062***	0,1644**	0,3041***	0,2008***	0,3054***	0,1391***	0,2364	0,0767	0,2322***	0,1395**	0,2319***
	0,0736	*0,0859*	*0,0769*	*0,0833*	*0,0699*	*0,0875*	*0,0695*	*0,0673*	*0,0536*	*0,1508*	*0,0820*	*0,0575*	*0,0555*	*0,0690*
ρ	0,8187***	0,8817***	0,8353***	0,8948***	0,7848***	0,8944***	0,8236***	0,8944***	-0,5482***	-0,5632***	0,9296***	-1,0000	-0,3431	-0,2898***
	0,0087	*0,0088*	*0,0084*	*0,0070*	*0,0096*	*0,0064*	*0,0089*	*0,0082*	*0,0001*	*0,0021*	*0,0005*	*2,2493*	*1,0126*	*0,0057*
δ_0	0,4291***	0,4193***	0,3448	0,3448	0,5538	0,5538	0,2900	0,2900	0,3448	0,3448	0,5538	0,5538	0,2900	0,2900
	0,0527	*0,0588*												
α	-	-	6,6132***	4,8027***	31,2705***	57,4459***	16,8090***	-4,3334***	1,3571***	1,3286***	15,8303***	17,1644***	-1,2957***	-1,5438***
	-	-	*0,0110*	*0,0404*	*0,7264*	*1,6139*	*1,0385*	*0,1653*	*0,0886*	*0,1568*	*0,0356*	*0,2944*	*0,0898*	*0,0660*
β	-	-	-19,965***	-14,493***	-56,573***	-104,08***	-54,925***	14,3010***	-4,0530***	-3,6888***	-28,674***	-30,927***	4,3135***	5,3087***
	-	-	*0,0054*	*0,1222*	*1,3177*	*2,9317*	*3,5119*	*0,5266*	*0,2654*	*0,4999*	*0,0040*	*0,5476*	*0,2674*	*0,1170*
LL	-23.643	-15.240	-27.529	-17.781	-28.918	-19.346	-27.290	-17.800	-17.110	-11.166	-17.037	-11.184	-17.186	-11.120
Adj. LL	-	-	-23.554	-15.183	-22.520	-15.184	-23.601	-15.184	-	-	-	-	-	-

Table 3: **Comparison of models**

This table reports the Akaike Information Criterion (AIC) and Bayes Information Criteria (BIC) values for the Gibson–Schwartz two-factor model and the new inventory two-factor model for both data Panels A (1990–12 with 1199 data points) and B (1990–04 with 783 data points). We estimate the inventory model with three different inventory level data series, the untransformed inventories, log inventories, and inverse inventories.

			AIC	BIC
Panel A	Gibson–Schwartz		-47.256	-47.179
(N=1199)	Est. Procedure I	Inv	-47.073	-46.981
		LogInv	-45.005	-44.913
		1/Inv	-47.166	-47.075
	Est. Procedure II	Inv	-34.187	-34.100
		LogInv	-34.041	-33.954
		1/Inv	-34.338	-34.252
Panel B	Gibson–Schwartz		-30.450	-30.380
(N=783)	Est. Procedure I	Inv	-30.330	-30.246
		LogInv	-30.332	-30.248
		1/Inv	-30.333	-30.249
	Est. Procedure II	Inv	-22.298	-22.219
		LogInv	-22.334	-22.255
		1/Inv	-22.205	-22.126

Table 4: **In-sample pricing error**

This table reports the in-sample pricing errors for all five futures contracts as well as the total error when considering all futures contracts. We report the root mean squared error (RMSE), the relative root mean squared error (RRMSE), and the mean error (ME).

		Panel A (1990–12 with N=1199)							Panel B (1990–04 with N=783)						
			Est. Procedure I			Est. Procedure II				Est. Procedure I			Est. Procedure II		
		Gibson–Schwartz	Inv	LogInv	1/Inv	Inv	LogInv	1/Inv	Gibson–Schwartz	Inv	LogInv	1/Inv	Inv	LogInv	1/Inv
RMSE	F1	3.06%	3.42%	3.68%	3.20%	6.88%	6.29%	6.80%	3.22%	3.59%	3.58%	3.58%	6.83%	6.79%	5.31
	F3	0.66%	1.35%	1.29%	1.00%	3.77%	2.54%	3.72%	0.72%	1.43%	1.42%	1.42%	3.93%	3.91%	2.22%
	F5	0.01%	0.95%	0.81%	0.59%	1.66%	0.00%	1.64%	0.04%	1.00%	0.99%	0.98%	1.75%	1.74%	0.03%
	F7	0.01%	0.78%	0.66%	0.48%	0.03%	1.98%	0.03%	0.01%	0.80%	0.79%	0.79%	0.04%	0.04%	1.72%
	F9	0.31%	0.73%	0.75%	0.47%	1.33%	3.55%	1.31%	0.35%	0.75%	0.74%	0.74%	1.39%	1.38%	3.07%
	All	**1,41%**	**1,77%**	**1,97%**	**1,71%**	**3,77%**	**3,71%**	**3,73%**	**1,48%**	**1,85%**	**1,84%**	**1,84%**	**3,66%**	**3,64%**	**3,01%**
RRMSE	F1	0.85%	0.95%	1.02%	0.89%	1.92%	1.75%	1.89%	1.03%	1.15%	1.14%	1.14%	2.19%	2.17%	1.70%
	F3	0.18%	0.38%	0.36%	0.28%	1.05%	0.71%	1.03%	0.23%	0.46%	0.46%	0.46%	1.26%	1.26%	0.71%
	F5	0.00%	0.27%	0.23%	0.17%	0.46%	0.00%	0.46%	0.01%	0.32%	0.32%	0.32%	0.57%	0.56%	0.01%
	F7	0.00%	0.22%	0.18%	0.13%	0.01%	0.55%	0,01%	0.00%	0.26%	0.26%	0.26%	0.01%	0.01%	0.56%
	F9	0.09%	0.20%	0.21%	0.13%	0.37%	0.99%	0.36%	0.11%	0.24%	0.24%	0.24%	0.45%	0.45%	1.00%
	All	**0,39%**	**0,49%**	**0,64%**	**0,55%**	**1,22%**	**1,20%**	**1,20%**	**0,48%**	**0,60%**	**0,59%**	**0,59%**	**1,18%**	**1,18%**	**0,97%**
ME	F1	-0.0037	-0.0040	-0.0039	-0.0047	0.0005	0.0008	0.0005	-0.0025	-0.0028	-0.0028	-0.0028	-0.0010	-0.0008	-0.0001
	F3	-0.0004	-0.0001	-0.0001	-0.0007	0.0004	-0.0001	0.0004	-0.0003	0.0001	0.0001	0.0001	0.0007	0.0006	0.0003
	F5	0.0000	0.0003	0.0002	-0.0002	0.0000	0.0000	0.0000	0.0000	0.0004	0.0004	0.0004	0.0004	0.0003	0.0000
	F7	0.0000	0.0002	0.0002	-0.0001	0.0000	0.0000	0.0000	0.0000	0.0003	0.0003	0.0003	0.0000	0.0000	-0.0001
	F9	0.0003	0.0005	0.0007	0.0002	0.0001	0.0002	0.0001	0.0003	0.0005	0.0005	0.0005	0.0002	0.0002	0.0002
	All	**-0.0007**	**-0.0006**	**-0.0006**	**-0.0011**	**0.0002**	**0.0002**	**0.0002**	**-0.0005**	**-0.0003**	**-0.0003**	**-0.0003**	**0.0001**	**0.0001**	**0.0001**

Table 5: **Hedging error**

This table reports the 1-week and 5-week hedging errors for the out-of-sample period Panel C (2005–12 with 416 data points). We report the mean error, the standard deviation of the error (STD), the mean absolute error (MAE), and the root mean squared error (RMSE). We construct the out-of-sample environment by keeping the calibrated parameters constant and employ one-week ahead forecasts based on the Kalman filter.

			Est. Procedure I			Est. Procedure II		
		Gibson–Schwartz	Inv	LogInv	1/Inv	Inv	LogInv	1/Inv
			F9 contract hedged with F7 and F3					
1-week	ME	0.000009	0.000009	0.000009	0.000009	-0.000008	-0.000005	0.000001
	MAE	0.001149	0.001148	0.001149	0.001148	0.001183	0.001119	0.001028
	STD	0.001894	0.001894	0.001895	0.001894	0.001895	0.001778	0.001595
	RMSE	0.001892	0.001892	0.001893	0.001892	0.001893	0.001776	0.001593
5-weeks	ME	0.000056	0.000056	0.000056	0.000056	0.000010	0.000025	0.000054
	MAE	0.002011	0.002009	0.002010	0.002009	0.002046	0.001966	0.001861
	STD	0.002866	0.002864	0.002866	0.002864	0.002874	0.002719	0.002513
	RMSE	0.002863	0.002861	0.002863	0.002861	0.002871	0.002716	0.002510
			F7 contract hedged with F5 and F1					
1-week	ME	0.000072	0.000072	0.000072	0.000072	0.000035	0.000038	0.000044
	MAE	0.003399	0.003398	0.003400	0.003398	0.003423	0.003235	0.002896
	STD	0.008210	0.008204	0.008209	0.008204	0.007699	0.007215	0.006318
	RMSE	0.008200	0.008194	0.008199	0.008194	0.007690	0.007206	0.006310
5-weeks	ME	0.000258	0.000256	0.000256	0.000256	0.000202	0.000215	0.000240
	MAE	0.005948	0.005944	0.005947	0.005944	0.006002	0.005633	0.005000
	STD	0.012014	0.011997	0.012004	0.011997	0.011445	0.010688	0.009300
	RMSE	0.012002	0.011985	0.011992	0.011985	0.011433	0.010678	0.009292
			F1 contract hedged with F3 and F5					
1-week	ME	0.000095	0.000094	0.000094	0.000094	0.000066	0.000308	0.000086
	MAE	0.006196	0.006199	0.006199	0.006199	0.006171	0.010150	0.006130
	STD	0.015512	0.015509	0.015510	0.015509	0.015364	0.021931	0.015046
	RMSE	0.015494	0.015491	0.015492	0.015491	0.015346	0.021907	0.015028
5-weeks	ME	0.000358	0.000355	0.000355	0.000355	0.000282	0.000072	0.000367
	MAE	0.010291	0.010291	0.010291	0.010291	0.010245	0.006155	0.009963
	STD	0.022419	0.022404	0.022406	0.022404	0.022150	0.015260	0.021483
	RMSE	0.022395	0.022380	0.022381	0.022380	0.022125	0.015242	0.021460

Table 6: **Reduction of hedging error (1-week horizon)**
This table reports the mean and median reduction of the mean absolute 1-week hedging error between the different term-structure models for hedging the 7-month contract for the out-of-sample period Panel C (2005–12 with 416 data points). We test if mean difference is significant different from zero by employing a standard t-test using Newey and West (1987) corrected standard errors; for the median we test significance by employing the test developed by Wilcoxon (1945). We construct the out-of-sample environment by keeping the calibrated parameters constant and employ one-week ahead forecasts based on the Kalman filter.

	Est. Procedure I			Est. Procedure II		
	Inv	LogInv	1/Inv	Inv	LogInv	1/Inv
			Reduction mean			
Gibson–Schwartz	0.04%	-0.01%	0.17%***	-0.70%	4.83%*	14.81%***
Inv	-	-0.05%***	0.13%***	-0.74%	4.79%*	14.78%***
LogInv	-	-	0.18%***	-0.69%	4.84%*	14.82%***
1/Inv	-	-	-	-0.87%	4.67%*	14.67%***
Inv	-	-	-	-	5.49%***	15.40%***
LogInv	-	-	-	-	-	10.49%***
			Reduction median			
Gibson–Schwartz	0.51%	0.39%	0.82%***	-5.75%***	-1.96%	6.14%***
Inv	-	-0.11%***	0.31%***	-6.29%***	-2.48%***	5.67%***
LogInv	-	-	0.20%***	-6.24%***	-2.43%***	5.69%***
1/Inv	-	-	-	-6.62%***	-2.80%***	5.37%***
Inv	-	-	-	-	3.58%***	11.25%***
LogInv	-	-	-	-	-	7.95%***

Table 7: **Reduction of hedging error (5-week horizon)**
This table reports the mean and median reduction of the mean absolute 5-week hedging error between the different term-structure models for hedging the 7-month contract for the out-of-sample period Panel C (2005–12 with 416 data points). We test if mean difference is significant different from zero by employing a standard t-test using Newey and West (1987) corrected standard errors; for the median we test significance by employing the test developed by Wilcoxon (1945). We construct the out-of-sample environment by keeping the calibrated parameters constant and employ one-week ahead forecasts based on the Kalman filter.

	Est. Procedure I			Est. Procedure II		
	Inv	LogInv	1/Inv	Inv	LogInv	1/Inv
			Reduction mean			
Gibson–Schwartz	0.06%	0.01%	0.21%**	-0.91%	5.29%**	15.94%***
Inv	-	-0.05%***	0.15%***	-0.98%	5.23%**	15.89%***
LogInv	-	-	0.20%	-0.92%	5.28%**	15.93%***
1/Inv	-	-	-	-1.12%	5.09%**	15.76%***
Inv	-	-	-	-	6.14%***	16.70%***
LogInv	-	-	-	-	-	11.25%***
			Reduction median			
Gibson–Schwartz	-0.35%	-0.27%	-0.22%***	-4.23%***	-0.47%***	4.49%***
Inv	-	0.07%***	0.13%***	-3.86%***	-0.12%***	4.82%***
LogInv	-	-	0.06%***	-3.94%***	-0.19%***	4.75%***
1/Inv	-	-	-	-4.00%***	-0.25%***	4.70%***
Inv	-	-	-	-	3.61%***	8.37%***
LogInv	-	-	-	-	-	4.93%***