VOLATILITY AND COMMODITY PRICE DYNAMICS*

by

Robert S. Pindyck
Massachusetts Institute of Technology
Cambridge, MA 02142

This draft: August 19, 2002

Abstract: Commodity prices tend to be volatile, and volatility itself varies over time. Changes in volatility can affect market variables by directly affecting the marginal value of storage, and by affecting a component of the total marginal cost of production: the opportunity cost of producing the commodity now rather than waiting for more price information. I examine the role of volatility in short-run commodity market dynamics, as well as the determinants of volatility itself. Specifically, I develop a model describing the joint dynamics of inventories, spot and futures prices, and volatility, and estimate it using daily and weekly data for the petroleum complex: crude oil, heating oil, and gasoline.

JEL Classification Numbers: G13; L71, Q40

Keywords: Commodity prices, futures markets, price volatility, inventories.

---

1 Introduction.

Most commodity markets are characterized by periods of sharp changes in prices and inventory levels. In addition, the level of volatility itself fluctuates over time. This paper examines the short-run dynamics of commodity prices and inventories, with a particular focus on the role of volatility. A goal is to determine how changes in spot prices, futures prices, and inventories are affected by changes in volatility, and to elucidate the channels through which such effects occur. Another objective is to examine the behavior of volatility itself.

Understanding the behavior and role of volatility is important for several reasons. First, as I show here, including volatility as a market variable can help us better understand short-run commodity market dynamics. Also, price volatility is a key determinant of the value of commodity-based contingent claims, including opportunities to invest in production facilities. Thus understanding its behavior is important for derivative valuation, hedging decisions, and decisions to invest in production facilities.

Changes in volatility affect prices, production, and inventories in two main ways. First, volatility directly affects the marginal value of storage, or, as it is commonly called, *marginal convenience yield*, i.e., the flow of benefits from an extra unit of inventory held by producers and/or consumers of the commodity. When prices — and hence production and demand — are more volatile, consumers and producers have a greater demand for inventories, which are needed to smooth production and deliveries, and reduce marketing costs. Thus an increase in volatility can lead to inventory build-ups and thereby raise prices in the short run.

Second, volatility affects the total marginal cost of production by affecting the size of the "option premium." Commodity producers hold operating options, with an exercise price equal to direct marginal production cost and a payoff equal to the price of the commodity. The *total* marginal cost of production equals the direct marginal cost plus the opportunity cost of exercising the firm's operating option rather than waiting for new price information. An increase in price volatility will raise the value of this option and the associated opportunity cost, and can thus result in a decrease in production.

Litzenberger and Rabinowitz (1995) used a two-period model to show that this option

1

premium can cause backwardation in futures markets. Using data for crude oil, they demonstrated that consistent with the theory, production is negatively correlated with price volatility, and the extent of futures market backwardation is positively correlated with price volatility. Also, Schwartz (1997) and Schwartz and Smith (2000) have shown how futures and spot prices can be used to estimate the parameters of a mean-reverting price process and derive values of commodity-based options. I go further and show how volatility and option value can be incorporated in a complete equilibrium model of a commodity market.

In this paper, I develop a weekly model that relates the dynamics of inventories, spot and futures prices, and the level of volatility. I estimate the model using data for the three commodities that make up the petroleum complex: crude oil, heating oil, and gasoline. To estimate volatility, I use sample standard deviations of adjusted daily log changes in spot and futures prices. As Campbell et. al. (2001) point out in their study of stock price volatility, in addition to its simplicity, this approach has the advantage that it does not require a parametric model describing the evolution of volatility over time.

As shown in this paper, at least for the petroleum complex, changes in price volatility are not predicted by market variables such as spot prices, inventory levels, or convenience yields, and can be viewed as largely exogenous. The volatility of, say, crude oil prices can be forecasted by past levels of volatility, but the marginal forecasting power of market variables is very low. However, changes in volatility directly affect market variables, by affecting the marginal value of storage, and by affecting price and production through the option premium. In addition, changes in the value of storage affect production, inventory holdings, and spot prices, so these variables are indirectly affected by changes in volatility.

This paper also provides evidence on how inventory holdings affect short-run price movements. In a competitive commodity market, inventories can be used to reduce costs of varying production (when marginal cost is increasing), and to reduce marketing costs by facilitating production and delivery scheduling and avoiding stockouts. These latter factors make it costly for firms to reduce inventories beyond some minimal level, even if marginal production cost is constant. The extent to which price will move in the short run depends on the cost of varying production as well as the cost of drawing down inventories.

Equilibrium inventory behavior is the solution to a stochastic dynamic optimization problem. Early studies of manufacturing inventories, as well as Eckstein and Eichenbaum's (1985) study of crude oil inventories, rely on a linear-quadratic specification to obtain an analytical solution to this problem. This is unrealistic for commodity markets because the cost of drawing down inventory is highly convex in the stock of inventory, rising rapidly as the stock falls toward zero, and remaining very small as the stock varies across moderate to high levels. I therefore adopt a more general specification and estimate the Euler equations that follow from intertemporal optimization, using futures market data to directly measure the marginal value of storage.[1]

This paper differs from my earlier (1994) study of commodity markets in several respects. First, I explicitly account for price volatility as a determinant of both the marginal value of storage and the full marginal cost of production. I can thereby estimate the extent to which changes in volatility affect prices and inventories, and obtain evidence on the channels through which these effects occur. I also examine the determinants of volatility. Finally, my earlier work used monthly data, but commodity market fluctuations occur on a shorter time scale. The weekly model estimated here yields a clearer picture of market dynamics.

In the next section, I lay out a model of short-run commodity market dynamics that links prices, inventories, convenience yield, and volatility. The model includes a set of Euler equations (first-order conditions) and cannot be solved analytically. However, in Section 3 phase diagrams are used to trace through the effects of various shocks. In Section 4, I discuss the data set, and examine the behavior of price volatility and market variables for each commodity. I show that these volatilities are rapidly mean reverting, but can be viewed as largely exogenous with respect to market variables such as inventory changes and price. Section 4 also presents GMM estimates of the full model. In Section 5, I use the model to examine the impact of shocks to volatility on price, inventories, and convenience yield. Section 6 concludes.

---

[1]See Pindyck (1994). This approach has also been used in studies of manufacturing inventories; see, e.g., Miron and Zeldes (1988) and Ramey (1991). Considine (1997) and Considine and Heo (2000) have estimated Euler equation models of inventory behavior for various petroleum products, focusing on the joint production characteristics of petroleum refining.

2 A Model of Prices, Inventories, and Volatility.

This section lays out a structural model that describes equilibrium in two competitive markets: the cash market for spot purchase and sale of the commodity, and the market for storage, in which an equilibrium level of inventories is held at a "price" equal to marginal value, i.e., marginal convenience yield. Together, these markets determine the spot price, the inventory level, and the convenience yield (and hence, implicitly, the futures price).

2.1 Cash Markets and Storage Markets.

In a competitive commodity market subject to stochastic fluctuations in production and/or consumption, producers (and to a lesser extent, consumers and third parties) will hold inventories. Producers hold them to reduce costs of adjusting production over time, and also to reduce marketing costs by facilitating production and delivery scheduling and avoiding stockouts. If marginal production costs are increasing with the rate of output and if demand is fluctuating, producers can reduce costs over time by selling out of inventory during high-demand periods, and replenishing inventories during low-demand periods. Inventories also serve as a "lubricant" to facilitate scheduling and reduce marketing costs. Industrial consumers of a commodity also hold inventories to facilitate their own production processes.

When inventory holdings can change, the market-clearing price is determined not only by current production and consumption, but also by inventories. Thus, we must account for equilibrium in both the cash and storage markets. In the cash market, purchases and sales of the commodity for immediate delivery occur at the *spot price*. Equilibrium in this market defines a relationship between the spot price and net demand, i.e., the difference between production and consumption. To see this, write consumption demand as $Q = Q(P, z_1)$, where P is the spot price and z_1 is a vector of demand-shifting variables. Likewise, write the supply function as $x = x(P, z_2)$, where z_2 is a vector of supply-shifting variables. Letting N_t denote the inventory level, the change in inventories at time t is:

$$\Delta N_t = x(P, z_2) - Q(P, z_1).$$

This just says that the cash market is in equilibrium when *net demand* (the demand for

production in excess of consumption) equals net supply. We can rewrite this in terms of the following inverse net demand function:

$$P_t = f(\Delta N_t; z_1, z_2).$$ (1)

Market clearing in the cash market therefore implies a relationship between the spot price and the change in inventories.

Now consider the market for storage. At any point in time, the supply of storage is the total quantity of inventories, N_t. In equilibrium, this must equal the quantity demanded, which is a function of price. The price of storage is the "payment" by inventory holders for the privilege of holding a unit of inventory, and has three components: the cost of physical storage (e.g., tanks to hold heating oil); the opportunity cost of forgone interest; and any expected change in the spot price. The price of storage will equal the value of the flow of services from the marginal unit of inventory, which is usually referred to as *marginal convenience yield*. Denoting the price of storage by ψ_t, the demand for storage function can be written as $N(\psi_t, z_3)$, where z_3 is a vector of demand-shifting variables, such as temperature. One important component of z_3 is the volatility of price, which is a good proxy for market volatility in general.[2] Writing this as an inverse demand function, we have:

$$\psi_t = g(N, z_3).$$ (2)

Thus market clearing in the storage market implies a relationship between marginal convenience yield (the price of storage) and the demand for storage.

Market equilibrium is determined from eqns. (1) and (2), and an additional equation (derived below) describing the dynamic tradeoff between producing and selling out of inventory. Given values for the exogenous variables z_1, z_2, and z_3, these three equations determine the values at each point in time of the three endogenous variables P_t, N_t, and ψ_t.

---

[2]The marginal value of storage is small when the total stock of inventories is large (because one more unit of inventory is of little extra benefit), but can rise sharply when the stock becomes small. Thus the demand for storage function should be downward sloping and convex, i.e., $\partial N/\partial \psi < 0$ and $\partial^2 N/\partial \psi^2 > 0$.

2.2 Operating Options and Convenience Yield.

Consider the incremental production decision for a firm that produces a commodity from a fixed quantity of reserves or other raw material, has a constant marginal production cost c, and faces a market price that fluctuates stochastically. The firm has an option to produce a unit now (and receive incremental net revenue $P - c$), or wait and possibly produce the unit in the future. At any point in the future, the net payoff from exercising this option is $V = \max[0, P_t - c]$. The greater the volatility of price, the greater is the expected value of this payoff, and thus the greater is the opportunity cost of exercising the option now rather than waiting. Thus price will exceed marginal cost by a premium, which I denote by ω_t. (See Dixit and Pindyck (1994) for a detailed discussion.) I discuss the determination of ω_t later; here, simply note that it is an increasing function of volatility. Hence an increase in volatility raises full marginal cost.

Next, consider the net (of storage costs) marginal convenience yield. This can be measured from spot and futures prices:

$$\psi_t - k = (1 + r)P_t - F_{1t}, \tag{3}$$

where F_{1t} is the futures price at time t for a contract maturing at time $t + 1$, r is the one-period interest rate, and k is the one-period cost of storage.[3] As discussed above, ψ_t is the value of the flow of production- and delivery-facilitating services from the marginal unit of inventory, a value that should be greater the greater is the volatility of price. However, ψ_t also includes operating options, such as the value of keeping oil in the ground rather than producing it now. To see this, suppose inventories yield no other services and $k = 0$. We would still need $\psi_t > 0$ for any oil to be produced. If $\psi_t = 0$, producers would have no incentive to exercise their options to produce (just as a call option on a non-dividend paying

---

[3]To see why eqn. (3) must hold, note that the (stochastic) return from holding a unit of the commodity for one period is $\psi_t + (P_{t+1} - P_t) - k$. Suppose that one also shorts a futures contract. The return on this futures contract is $F_{1t} - F_{1,t+1} = F_{1t} - P_{t+1}$, so one would receive a total return equal to $\psi_t + (F_{1t} - P_t) - k$. No outlay is required for the futures contract, and this total return is non-stochastic, so it must equal the risk-free rate times the cash outlay for the commodity, i.e., rP_t, from which eqn. (3) follows. Because futures contracts are marked to market, strictly speaking, F_{1t} should be a forward price. For most commodities, however, the difference between the futures and forward prices is negligible.

stock is optimally exercised only at expiration). Put differently, oil in the ground provides a price-protection service, the value of which is positive and is included in ψ_t. Thus even if inventory levels are large, we should observe at least weak backwardation in the futures market.[4] This simply reflects the fact that there is some value to delaying production and waiting for more information, even if the expected future spot price is less than the current spot price. Furthermore, this value is greater the greater is the volatility of price.

In summary, volatility should affect convenience yield in two ways, and in both cases positively. First, it should affect the value of the flow of production- and delivery-facilitating services that inventories provide. Second, it should affect the price-protection service that is part of convenience yield. As an empirical matter, it will not be possible to measure these effects separately. Instead, we can only measure the combined effect, and test whether convenience yield depends positively on volatility.

2.3 Costs.

I now turn to the specification of the model that will be estimated. The total economic cost of commodity production, marketing, and storage is given by:

$$\text{TC} = C(x) + \Omega(x; \sigma, r) + \Phi(N, P, \sigma) + kN, \tag{4}$$

and has four components: (1) $C(x)$ is direct production cost. (2) $\Omega(x; \sigma, r)$ is the opportunity cost of producing x now, rather than waiting. As explained below, it depends on the level of price volatility, σ, and the risk-free rate r. (3) $\Phi(N, P, \sigma)$ is total marketing cost, i.e., the cost of production and delivery scheduling and avoidance of stockouts, and is decreasing in the level of inventories N. (4) k is the per-unit cost of storage, which I assume is constant.

Two other variables must be defined. First, $\psi = -\partial\Phi/\partial N$ is the marginal value of storage, i.e., marginal convenience yield: $\psi_t = (1 + r)P_t - F_{1t} + k$. Second, $\omega = \partial\Omega/\partial x$ is the marginal option premium, i.e., the opportunity cost of exercising the option to produce an incremental unit of the commodity, given a total production level x.

---

[4]For a detailed discussion of this point and derivation of ψ_t in the context of a two-period model, see Litzenberger and Rabinowitz (1995). McDonald and Shimko (1998) also address this point, and measure ψ_t and its components for the gold market.

The components of cost are modelled as follows. I assume that the direct cost of production is quadratic. For crude oil, direct cost is:

$$C(x) = (c_0 + \eta_t)x_t + \tfrac{1}{2}c_1 x_t^2 \,, \tag{5}$$

where η_t is a random shock. Note that there are no input cost variables (such as wage rates) in (5); such variables cannot be measured — and are unlikely to vary much — on a weekly basis. For heating oil and gasoline, however, the cost of the crude oil input is a large component of direct production cost, and must be accounted for:

$$C(x) = (c_0 + \eta_t)x_t + \tfrac{1}{2}c_1 x_t^2 + c_2 P_{C,t} x_t \,, \tag{6}$$

where $P_{C,t}$ is the price of crude oil.

Marketing cost should be roughly proportional to the price of the commodity. It should also be increasing in the level of price volatility, which I use as a proxy for market volatility in general.[5] Greater volatility makes scheduling and stockout avoidance more difficult, increasing the demand for storage. Ideally, the total marketing cost function Φ should be derived from a dynamic optimizing model that accounts for stockout costs and costs of scheduling and managing production and shipments, but that is beyond the scope of this work. Instead, I assume that this function is isoelastic in price, the variance of log price changes, and the total inventory level:

$$\Phi(N, P, \sigma) = \frac{1}{\alpha_3 - 1} \exp(b_0 + \sum_{j=1}^{11} b_j \mathrm{DUM}_{jt}) P_t^{\alpha_1} (\sigma_t^2)^{\alpha_2} N_t^{1-\alpha_3} \,, \tag{7}$$

where DUM_{jt} are monthly time dummies and $\alpha_3 > 1$. This implies that the marginal value of storage (marginal convenience yield), $-\partial\Phi/\partial N$, can be written as:

$$\log \psi_t = b_0 + \sum_{j=1}^{11} b_j \mathrm{DUM}_{jt} + \alpha_1 \log P_t + \alpha_2 \log \sigma_t^2 - \alpha_3 \log N_t \,. \tag{8}$$

To model the marginal opportunity cost $\omega_t = \partial\Omega_t/\partial x_t$, we need an expression for the value of the option to produce a marginal unit of the commodity, and the optimal price P^* at

---

[5]Price volatility is highly correlated with the volatility of consumption and production. Also, price fluctuations themselves (whether caused by fluctuations in net demand, or something else, such as speculative buying and selling) cause consumption and production to fluctuate.

which that option should be exercised. The difference between P^* and the direct marginal cost $C'(x)$ is the opportunity cost of exercising the option to produce the marginal unit. Valuing this option requires assumptions about the stochastic dynamics of price. Because commodity prices tend to be mean-reverting, I assume that random shocks to demand and supply lead to a reduced form price process that can be written in continuous time as:

$$dP/P = \lambda(\mu - P)dt + \sigma dz ,\tag{9}$$

or equivalently:

$$d\log P = [\lambda(\mu - P) - \tfrac{1}{2}\sigma^2]dt + \sigma dz .\tag{10}$$

Here, μ is the "normal" price to which P_t tends to revert and λ is the speed of reversion. I treat σ as a constant because allowing for stochastic volatility precludes a closed-form solution for the option value. Furthermore, it should not affect the way in which the option value depends on volatility, although it will affect its magnitude (overstating it). To account for this, I include a scaling coefficient that is estimated as part of the model.[6]

In the Appendix, I show that if the price process follows eqn. (10) and direct marginal cost is non-stochastic, a series solution can be found for the value of the option to produce. For estimation purposes, I use a quadratic approximation to this solution. As shown in the Appendix, letting r denote the risk-free interest rate and ρ the risk-adjusted expected return on the commodity, the opportunity cost ω_t can be written as:

$$\omega_t = \frac{1}{\sqrt{\gamma_1\gamma_2}} - \mu ,\tag{11}$$

where

$$\gamma_1 = \frac{\lambda\theta}{\lambda\mu + \theta\sigma^2} , \qquad \gamma_2 = \frac{\lambda(\theta + 1)}{2\lambda\mu + (2\theta + 1)\sigma^2} ,\tag{12}$$

and

$$\theta = \frac{1}{2} + \frac{(\rho - r - \lambda\mu)}{\sigma^2} + \sqrt{\left[\frac{(r - \rho + \lambda\mu)}{\sigma^2} - \frac{1}{2}\right]^2 + \frac{2r}{\sigma^2}} .\tag{13}$$

I include a scaling coefficient, so that $c_3\omega_t$ is the marginal opportunity cost. Note that the estimated value of c_3 should be close to 1.

---

[6]The numerical analyses of Hull and White (1987) suggest that treating σ as non-stochastic makes little quantitative difference in any case. See, also, Franks and Schwartz (1991).

2.4 Euler Equations.

With expressions for the components of cost, we can solve the intertemporal profit maximization problem, making use of the fact that in the U.S. markets for crude oil and oil products are reasonably competitive, so that producers can be treated as price takers. Of course much of the crude oil and some of the gasoline consumed in the U.S. is imported, but the presence of imports will simply make the *domestic* net demand function (which I estimate) more elastic than it would be otherwise. (If the supply of imports is highly elastic, the spot price will have little or no dependence on the change in domestic inventories.)

Taking prices as given, firms choose production and inventory levels to maximize the present value of the expected flow of profits:

$$\max \mathcal{E}_t \sum_{\tau=0}^{\infty} R_{\tau,t}(P_{t+\tau}Q_{t+\tau} - \text{TC}_{t+\tau}) , \qquad (14)$$

where $R_{\tau,t}$ is the τ-period discount factor, Q is sales, TC is given by eqn. (4), and the maximization is subject to the accounting identity

$$\Delta N_t = x_t - Q_t . \qquad (15)$$

(The maximization is subject to the additional constraint that $N_{t+\tau} \geq 0$ for all τ, but because $\Phi \to \infty$ as $N \to 0$, this constraint will never be binding.)

To obtain first-order conditions, first maximize with respect to x_t, holding N_t fixed so that $\Delta x_t = \Delta Q_t$:

$$P_t = c_0' + c_1 x_t + c_2 P_{C,t} + c_3 \omega_t + \eta_t . \qquad (16)$$

(I have included the term for the crude oil input price, $c_2 P_{C,t}$, so this equation would apply to heating oil and gasoline; for crude oil this term is dropped.)

It will be convenient to eliminate production and write the model in terms of prices and inventories. In the short run (a period of one week), consumption should be very inelastic with respect to price, so I model it as:

$$Q_t = \overline{Q} + \sum_{j=1}^{11} d_j \text{DUM}_{jt} + c_4 \text{HDD}_t + c_5 \text{CDD}_t + c_6 T_t + \epsilon_t , \qquad (17)$$

where the DUM_{jt} are monthly dummies, HDD and CDD are, respectively, heating and cool-ing degree days, and T is a time trend. Thus I assume that consumption fluctuates seasonally and in response to changes in temperature, is subject to (possibly serially correlated) random shocks (ϵ_t), but is insensitive to price. Substituting for Q_t in eqn. (15) and rearranging:

$$x_t = \Delta N_t + \overline{Q} + \sum_{j=1}^{11} d_j \text{DUM}_{jt} + c_4 \text{HDD}_t + c_5 \text{CDD}_t + c_6 T_t + \epsilon_t \qquad (18)$$

Thus eqn. (16) can be rewritten as:

$$P_t = c_0 + c_1 \Delta N_t + c_2 P_{C,t} + c_3 \omega_t + c_4' \text{HDD}_t + c_5' \text{CDD}_t + c_6' T_t + \sum_{j=1}^{11} d_j' \text{DUM}_{jt} + c_1 \epsilon_t + \eta_t \,, \quad (19)$$

where $c_0 = c_0' + c_1 \overline{Q}$, $c_4' = c_1 c_4$, $c_5' = c_1 c_5$, $c_6' = c_1 c_6$, and $d_j' = c_1 d_j$. Eqn. (19) describes market clearing in the cash market.[7]

Next, maximize eqn. (14) with respect to N_t, holding Q_t and N_{t+1} fixed:

$$0 = \mathcal{E}_t[c_0'(1-R_{1t}) + \psi_t - k - \eta_t + R_{1t}\eta_{t+1} - c_1(x_t - R_{1t}x_{t+1}) - c_2(P_{C,t} - R_{1t}P_{C,t+1}) - c_3(\omega_t - R_{1t}\omega_{t+1})] \,. \qquad (20)$$

Over a one-week time period, $R_{1t} \approx 1$. Making this substitution, substituting eqn. (18) for x_t, and denoting $\Delta^2 N_{t+1} \equiv \Delta N_{t+1} - \Delta N_t$ yields the second first-order condition:

$$\begin{aligned} 0 = \mathcal{E}_t[c_1 \Delta^2 N_{t+1} \quad + \quad & \psi_t - k + c_2 \Delta P_{C,t+1} + c_3 \Delta\omega_{t+1} + \sum_{j=1}^{11} d_j' \Delta\text{DUM}_{j,t+1} \\ + \quad & c_4' \Delta\text{HDD}_{t+1} + c_5' \Delta\text{CDD}_{t+1} + c_6' + \Delta\eta_{t+1} + c_1 \Delta\epsilon_{t+1}] \,. \qquad (21) \end{aligned}$$

Eqn. (19) simply equates price with full marginal cost, where the latter includes the opportunity cost of exercising the marginal operating option. It contains error terms repre-senting the unexplained part of marginal cost (η_t) and unanticipated shocks to demand (ϵ_t). Eqn. (21) describes the tradeoff between selling out of inventory versus producing. To see this, rearrange the equation so that $\psi_t - k$ is on the left-hand side. The equation then says that net marginal convenience yield (the savings in marketing costs over the coming period from having another unit of inventory, net of storage costs) should equal the expected change

---

[7]Eqn. (19) is an expanded version of a model that has been used by a number of other authors. See, for example, Williams and Wright (1991), Routledge, Seppi, and Spatt (2000), and Schwartz and Smith (2000).

in production cost (the increase this period minus the decrease next period) from producing a unit now rather than selling it from inventory and then replenishing inventory by producing it next period. The expected change in production cost can come from expected changes in input prices, expected changes in opportunity costs, and expected increases in cost due to convexity of the cost function.

To estimate the model, I substitute eqn. (8) for ψ_t in eqn. (21). Also, because estimation is by GMM, I drop the expectation operator and use actual values of variables dated at $t+1$:

$$
\begin{aligned}
0 \;=\; & c_1 \Delta^2 N_{t+1} + \exp\!\left(b_0 + \sum_{j=1}^{11} b_j \mathrm{DUM}_{jt}\right) P_t^{\alpha_1} (\sigma_t^2)^{\alpha_2} N_t^{-\alpha_3} - k + c_2 \Delta P_{C,t+1} + c_3 \Delta \omega_{t+1} + \\
& \sum_{j=1}^{11} d'_j \Delta \mathrm{DUM}_{j,t+1} + c'_4 \Delta \mathrm{HDD}_{t+1} + c'_5 \Delta \mathrm{CDD}_{t+1} + c'_6 + \Delta \eta_{t+1} + c_1 \Delta \epsilon_{t+1} \,.
\end{aligned}
\tag{22}
$$

The model is closed by including eqn. (8) for the marginal convenience yield. Together, eqns. (19), (22), and (8) describe the evolution of the state variables P_t, N_t, and ψ_t. As I show later, it is reasonable to treat the fourth state variable, the volatility σ_t, as exogenous.

3   Market Dynamics.

Before discussing the estimation of this model, it is useful to examine its theoretical implications for market dynamics and the effects of shocks. I have not actually solved the firm's stochastic dynamic optimization problem (beyond deriving first-order conditions), so I cannot calculate optimal trajectories for market variables that correspond to particular stochastic processes for demand, cost, and volatility shocks. However, I can analyze deterministically optimal trajectories for market variables, consistent with firms choosing output and inventory levels that solve the corresponding deterministic optimization problem. I examine such trajectories as a way of characterizing the market behavior implied by the theory. Later I use this approach to estimate the response of prices and inventories to various shocks.

First, consider the (deterministic) steady-state equilibrium in which there are no seasonal variations in cost or demand, σ is constant, and there are no other shocks so that $\Delta N = \Delta^2 N = \Delta P = 0$. Replacing expectations with actual values in eqn. (21), treating HDD and CDD as constants and setting the time dummies and time trend parameter c_6 to zero, and

using overbars to denote equilibrium values, we have $\overline{\psi} = k$. Also, $\overline{P} = c_0 + c_3\overline{\omega}$,[8] and

$$\overline{N} = b\overline{P}^{\alpha_1/\alpha_3}\sigma^{2\alpha_2/\alpha_3}k^{-1/\alpha_3} \ ,$$

where $b = e^{b_0/\alpha_3}$.

We can now draw a phase diagram for the two state variables N and ΔN. When estimating the model, I find that for both crude oil and heating oil, the parameters α_1 and α_3 in eqn. (8) are close to 1 and 2 respectively. Using these values, eqns. (19) and (21) yield the following isocline for $\Delta^2 N = 0$:

$$N = g(\sigma)\sqrt{c_0 + c_1\Delta N} \ , \tag{23}$$

where $g(\sigma) = b\sigma^{2\alpha_2/\alpha_3}k^{-1/\alpha_3}$, and $g'(\sigma) > 0$. In Figure 1, this isocline is the curve labelled $\Delta^2 N = 0$. The second isocline, for $\Delta N = 0$, is simply the vertical axis. Note that the optimization problem implies a unique approach path to the equilibrium value for N.

Figure 2 traces through the impact of a permanent but unanticipated increase in σ. Note from eqn. (23) that the $\Delta^2 N = 0$ isocline shifts up, so that the steady state equilibrium level of inventories increases from N_0 to N_1. The process through which this occurs must follow the optimal trajectory, labelled A in the top panel of the figure. Thus ΔN_t jumps to an initially high level (ΔN_1), and then declines toward zero as N_t approaches N_1.

The bottom half of Figure 2 shows the movements of price, inventory, and convenience yield in the cash and storage markets. In the cash market, the increase in σ increases the opportunity cost of producing, ω, so the net demand curve shifts upward. As ΔN_t jumps to its initially high level, the spot price jumps to P_1. Price and ΔN_t then move down the new net demand curve, until in the new equilibrium, with $\Delta N_t = 0$, the price is $P_2 > P_0$. In the storage market, the increase in σ causes the marginal value of storage to increase, so that the demand for storage curve shifts up from ψ_0^D to ψ_1^D. In the short run, the inventory level (the supply of storage) is fixed at N_0, so convenience yield jumps from ψ_0 to ψ_1. Over time the inventory level increases to N_1, so that convenience yield falls to $\psi_2 > \psi_0$.

---

[8]This applies to crude oil. For heating oil or gasoline, $\overline{P} = c_0 + c_2\overline{P_C} + c_3\overline{\omega}$.

4   Data and Estimation.

This section discusses the construction of the dataset, the method of estimation, the modelling of volatility, and the estimation results.

4.1   The Data.

The model is estimated using weekly data covering the period January 1, 1984 through January 31, 2001 for crude oil and heating oil. This start date was chosen because it is about three months after the beginning of trading of crude oil futures. The data for gasoline begin in January 1985, reflecting the later start of futures trading for that commodity.

For each commodity, daily futures settlement price data were compiled for the nearest contract (often the spot contract), the second-nearest contract, and the third-nearest. These prices are denoted by $F1$, $F2$, and $F3$. The spot price can be measured in three alternative ways. First, one can use data on *cash prices*, purportedly reflecting actual transactions. One problem with this approach is that daily cash price data are usually not available. A second and more serious problem is that a cash price can include discounts and premiums that result from relationships between buyers and sellers, and need not even reflect precisely the same product that is specified in the futures contract. A second approach, which avoids these problems, is to use the price on the spot futures contract, i.e., the contract expiring in month t. But this also has problems, because the spot contract often expires before the end of the month. In addition, active spot contracts do not always exist for each month.

The third approach, which I use here, is to infer a spot price from the nearest and the next-to-nearest active futures contracts. This is done for each day by extrapolating the spread between these contracts backwards to the spot month as follows:

$$P_t = F1_t (F1_t/F2_t)^{n_{0t}/n_1} , \qquad (24)$$

where P_t is the spot price on day t, $F1_t$ and $F2_t$ are the prices on the nearest and next-to-nearest futures contracts, and n_{0t} and n_1 are the number of days from t to the expiration of the first contract, and the number of days between the first and second contracts.

14

Given these daily estimates of spot prices, I compute weekly estimates of volatility. To do this, one must take into account weekends and other non-trading days. If the spot price of the commodity followed a geometric Brownian motion, then this could be done simply by dividing log price changes by the square root of the number of intervening days (e.g., three days in the case of a week-end), and then calculating the sample variance. However, as is well known, on average the standard deviation of n-day log price changes is significantly less than $\sqrt{n}$ times the standard deviation of 1-day log price changes, when n includes non-trading days. To deal with this, I sort the daily price data by intervals, according to the number of days since the last trading day. For example, if there were no holidays in a particular period, prices for Tuesday, Wednesday, Thursday, and Friday would all be classified as having an interval of one day, since there was always trading the day before. Monday, on the other hand, would be classified as an interval of three days, because of the 2-day weekend. Because of holidays, some prices could also be assigned to intervals of two, four, or even five days (the latter occurring when a weekend was followed by a 2-day holiday).

For each interval set, I calculate the sample standard deviation of log price changes for the entire 16- or 17-year sample for each commodity. Let $\hat{s}_n$ denote this sample standard deviation for log price changes over an interval of n days. I then compute the "effective" daily log price change for each trading day as follows:

$$r_\tau = \frac{(\log P_\tau - \log P_{\tau-n})}{\hat{s}_n/\hat{s}_1} \ . \tag{25}$$

For each week, I then compute a sample variance and corresponding sample standard deviation using these daily log price changes for that week and the preceding four weeks:

$$\hat{\sigma}_t = \sqrt{\frac{1}{N-1}\sum_{\tau-1}^{N}(r_{t\tau} - \overline{r}_t)^2} \ , \tag{26}$$

where N is the number of "effective" days in the five-week interval. Eqn. (26) gives the sample standard deviation of *daily* percentage price changes; to put it in weekly terms, I multiply by $\sqrt{30/4} = \sqrt{7.5}$. The resulting weekly series is my measure of volatility, σ_t.

The T-period net marginal convenience yield, $\psi'_{T,t} = \psi_{T,t} - k_T$, is computed weekly from

eqn. (3) using the futures price and estimated spot price for the Wednesday of each week:[9]

$$\psi'_{T,t} = (1 + R_{T,t})P_t - F_{T,t} \,, \tag{27}$$

where $R_{T,t}$ is a risk-free T-period interest rate. I use the futures price corresponding as closely as possible to a 3-month interval from the spot price, and I use the 3-month Treasury bill rate for the interest rate. These net marginal convenience yields are then converted to weekly terms, i.e., dollars per unit of commodity per week.

For each commodity, there are periods when $\psi'_{T,t}$ is negative. By definition, *gross* marginal convenience yield must always be positive, so I estimate k for each commodity as $\hat{k} = |\min \psi'_t|$, and then compute gross marginal convenience yield as $\psi_t = \psi'_t + \hat{k}$.

To calculate the opportunity cost ω_t from eqn. (11), I need estimates of μ and λ, and the average value of σ, for each commodity. I estimate these parameters from an OLS regression of the discrete-time version of eqn. (10):

$$\Delta \log P_t = \alpha - \lambda P_{t-1} + \sigma \epsilon_t \,, \tag{28}$$

so that $\hat{\mu} = \hat{\alpha}/\hat{\lambda} + \hat{\sigma}^2/2\hat{\lambda}$. The resulting estimates of μ, λ, and σ for crude oil, heating oil, and gasoline respectively are: $\hat{\mu} = \$20.44$, 57.2 cents, and 58.6 cents; $\hat{\lambda} = .00114$, .00050, and .00071; and $\hat{\sigma} = .050$, .052, and .059. Using these estimates, a weekly series for ω_t was computed from eqn. (11) for each commodity.

Finally, from the U.S. Department of Energy, *Monthly Energy Review*, I obtained data on weekly production and inventory levels for crude oil, heating oil (distillate fuel oil), and gasoline, measured in millions of barrels. These numbers are announced on the Tuesday evening of each week, so that the information is incorporated in the prices and convenience yields I use for the Wednesday of each week.

Summary statistics are presented in Table 1. For crude oil, spot and futures prices are in dollars per barrel; for heating oil and gasoline, they are in cents per gallon. For crude oil, convenience yield is measured in dollars per barrel per week, and for heating oil and gasoline, in cents per gallon per week. For all three commodities, volatility (σ) is the standard

---

[9]If Wednesday is a holiday, I use Thursday's price.

deviation of weekly log price changes (computed from daily data, as described above). Stocks and production levels are measured in millions of *barrels* for all three commodities. Observe that for all three commodities, on average the spot price is higher than the futures price, i.e., on average there is strong backwardation.

I ran augmented Dickey-Fuller unit root tests on P_t, N_t, ψ_t, and σ_t, with six lags included. The tests were run including a constant, and then a constant and trend, in the equation. In all cases the results implied a rejection of a unit root. Thus in much of the empirical analysis that follows, I work with variables in levels.

Figures 3–5 show, for each commodity, the four key variables analyzed in this paper: weekly inventories (N_t), convenience yield (ψ_t), the spot price (P_t), and spot price volatility (σ_t). Observe that for crude oil and heating oil, volatility had major spikes in 1986 (when Saudi Arabia flooded the oil market, causing prices to fall sharply) and in 1991 (following the Iraqi invasion of Kuwait and the Gulf War). For the remainder of the sample, changes in crude oil and heating oil volatility were much more subdued. Spot price volatility for gasoline, however, varied throughout the period, and changes were much more persistent. Also note that there are strong cyclical patterns to inventory holdings, and, consequently, to convenience yield. Convenience yield fluctuates considerably for all three of the commodities, and as we will see, much of this can be explained by changes in volatility.

4.2   Estimation Method.

I estimate the model defined by equations (19), (22), and (8) using Generalized Method of Moments (GMM). GMM is an instrumental variables procedure that minimizes the correlation between variables known at time t and the equation residuals, and is thus a natural estimator for an Euler equation model such as this one. I also estimate a separate equation for the volatility variable, σ_t. As discussed below, σ_t is best forecasted by its own past values and past values of exogenous variables such as exchange rates and interest rates.

Eqns. (19), (22), and (8) include the "structural" error terms η_t and ϵ_t, which represent unobserved shocks to cost and demand. These errors may be serially correlated, and appear in differenced form in eqn. (22). In addition, when estimating the model, actual values for

variables at time $t+1$ are used in place of expectations, which introduces expectational errors. Thus the equations will have composite error terms with a possibly complex autocorrelation structure. The GMM procedure uses an autocorrelation-robust weighting matrix and yields autocorrelation-robust standard errors. However, the error structure has implications for the choice of instruments.

By definition, the expectational errors are uncorrelated with any variable known at time t. The structural errors, however, may be correlated with endogenous variables. Hence I use as instruments only variables that can reasonably be viewed as exogenous. The instrument list includes the seasonal dummy variables, the time trend, heating and cooling degree days, and the following variables unlagged, lagged once, and lagged twice: the exchange-weighted value of the U.S dollar (EXVUS), the New York Stock Exchange Index (NYSE), the three-month Treasury bill rate (TBILL), the rate on Baa corporate bonds (BAA), and the Commodity Research Bureau's commodity price index (CRB). I also include the following endogenous variables lagged two periods: the spot price, production, inventory, and convenience yield. With the constant term, this gives a total of 34 instruments.

The minimized value of the objective function from the GMM procedure times the number of observations provides a statistic, J, which is distributed as χ^2 with degrees of freedom equal to the number of instruments times the number of equations minus the number of parameters. This statistic is used to test the model's overidentifying restrictions, and hence the hypothesis that agents are optimizing with rational expectations.

4.3   Volatility.

In Sections 2 and 3, I set forth a structural model in which price, inventories, and convenience yield are determined endogenously, and can depend directly or indirectly on volatility, as well as exogenous variables such as heating and cooling degree days. Given a time series for volatility and the other exogenous variables, the model can be solved forward through time to yield trajectories for the three endogenous variables.

A natural question is whether these market variables, or other exogenous variables, can predict volatility. To investigate this, I estimate a simple vector autoregression (VAR) relat-

ing the three market variables and volatility to each other and to a set of exogenous variables, and then examine the predictive power of each variable. Specifically, I estimate a VAR using six lags of each of the four variables and six lags of the following exogenous variables: the 3-month Treasury bill rate, the Baa corporate bond rate, the exchange-weighted value of the dollar, and monthly dummy variables to account for seasonal variation. For each equation, I then test the hypothesis that all lags of a particular endogenous variable can be excluded as explanators of the dependent variable.

The results are shown in Table 2. Each entry shows the marginal significance level (based on an F-test) for omitting the six lags of the variable in the column heading from the unrestricted ordinary least squares (OLS) prediction equation that includes a constant and six lags of each of the four variables, along with the exogenous variables mentioned above.

Observe that the spot price, inventories, and convenience yield all have virtually no predictive power with respect to volatility in the case of crude oil and heating oil, which is consistent with the view that volatility is exogenous. However, both the spot price and convenience yield are significant predictors of volatility for gasoline. Of course this can simply reflect the fact that past values of the spot price affect past values of volatility, which in turn affect current values of volatility. In terms of predicting the other variables, the results in Table 2 differ substantially across the three commodities. Most notably, for crude oil, volatility is a significant predictor of the spot price, but for the other commodities it is not a significant predictor of any of the other three market variables.

Table 3 shows estimates of linear forecasting equations for volatility, based on a sixth-order autoregression and including six lags of the 3-month Treasury bill rate, the Baa corporate bond rate, the exchange-weighted value of the dollar, the CRB commodity price index, and monthly dummy variables. These equations are estimated both by GMM and OLS. Note that the only significant explanators of volatility are its own past values; the other explanatory variables are largely insignificant. These results are unchanged by adding lagged values of the three market variables to the regressions.

Thus volatility is not explained by market variables, or by economy-wide variables such as interest rates or exchange rates. I therefore treat volatility as exogenous. For simulation

19

purposes, I use the GMM estimates in Table 3 to generate forecasts of volatility.

4.4 Euler Equation Estimates.

Table 4 shows the results of estimating eqns. (19), (22), and (8) as a system by GMM. In addition to the 10 coefficients shown in the table, there are an additional 22 coefficients (not shown) associated with the 11 monthly time dummies: the b_js in the marginal convenience yield eqn. (8), and the d_js in the demand eqn. (17) and thus in eqns. (19) and (22). The table shows t-statistics based on autocorrelation-consistent standard errors. The J-statistics are distributed as $\chi^2(71)$ for crude oil and $\chi^2(70)$ for heating oil and gasoline, and are all insignificant at the 5 percent level. Thus we fail to reject the overidentifying restrictions.

For each commodity, the marginal cost of production is increasing (c_1 is positive and significant), so the net demand curve, $P(\Delta N_t)$, is upwards sloping. As expected, for heating oil and gasoline, the price of crude oil is the most important determinant of marginal cost. The estimate of c_2 is between 2.7 and 2.9; thus a $1 increase in the per-barrel price of crude oil, i.e., a $1/42 = 2.4$ cents per gallon increase, leads to a roughly commensurate increase in the per gallon price of heating oil or gasoline. Also as expected, an increase in heating degree days increases the demand for heating oil (shifting the net demand curve upwards), and reduces the demand for gasoline. Cooling degree days, however, is insignificant for all three commodities. Finally, the opportunity cost ω_t affects total marginal cost as predicted only for heating oil: c_3 is close to 1 and significant for heating oil, but negative for crude oil and gasoline.

Apart from the constant term and time dummies, the marginal value of storage, ψ, is characterized by the three coefficients α_1, α_2, and α_3, which appear in eqns. (8) and (22). For crude oil and heating oil, the estimates of these coefficients are all positive and significant, and consistent with a well-behaved marginal value of storage function: $\hat{\alpha}_3 > 1$, $\hat{\alpha}_1$, the elasticity of ψ with respect to price, is close to 1, and $\hat{\alpha}_2 > 0$, i.e., ψ is increasing with volatility. For gasoline, however, $\hat{\alpha}_3$ was negative, so the model was re-estimated with α_3 constrained to equal 1.1. The resulting estimates of $c_1, ..., c_6$ are largely unchanged, but $\hat{\alpha}_1$ drops from .84 to .63, and $\hat{\alpha}_2$ becomes insignificant.

Thus the model fits the theory very well for heating oil, but less so for crude oil and gasoline. For both crude oil and gasoline, the net demand function is upward sloping, but does not depend on the marginal opportunity cost as predicted. Also, the unconstrained marginal value of storage function for gasoline is increasing in the level of inventories N_t, and when α_3 is constrained to equal 1.1 so that the function is slightly decreasing in N_t, the elasticity with respect to volatility becomes zero.

5   Simulations.

Dynamic simulations, in which eqns. (19), (22), and (8) solved as a system, can be used to evaluate the ability of the model to replicate the behavior of the endogenous variables, and to study the effects over time of a shock to volatility (or some other variable). As explained earlier, I cannot calculate optimal trajectories for market variables that correspond to particular stochastic processes for demand, cost, and volatility shocks. However, it is still useful to calculate deterministically optimal trajectories for market variables, consistent with output and inventory levels that solve the corresponding deterministic optimization problem.

Because the coefficient estimates for heating oil are consistent with all of the predictions of the theory, I focus on this commodity. The first simulation covers the 10-week period Aug. 8, 1990 to Oct. 3, 1990, a tumultuous time for the heating oil market: Iraq invaded Kuwait, and the spot price of heating oil increased over the period from about 60 cents to about \$1 per gallon. The second simulation covers the last 10 weeks of the sample: Nov. 29, 2000 to Jan. 31, 2001. In both simulations, the GMM-estimated equation for volatility is used to forecast that variable forward from the starting date. Also, both simulations are dynamic; actual values of P_t, N_t, ψ_t, and σ_t are used only prior to the starting date.

Simulated and actual values of the spot price, inventory level, and convenience yield are shown in Figures 6, 7, and 8. (Each figure shows both simulations.) Note from Figure 6 that for the 1990 period, the simulated spot price tracks the actual sharp increase during August 22 to 29, but not the temporary decline on September 5. The simulated price is again close to the actual over the last three weeks of the period. Actual inventories rose

by about 11% over this period, but the simulated series increases by much more. Finally, simulated convenience yield closely tracks the actual series throughout the 10-week period.

I turn now to the second simulation. From November 22, 2001 to January 31, 2001, the actual spot price fell from about $1.10 per gallon to about 83 cents, and the simulated series tracks this decline quite closely. The model again over-predicts inventories; actual inventories were fairly flat over this period, but simulated inventories increase by about 13% and then fall. Finally, the actual convenience yield fluctuated widely, and the simulated series replicates the directional movements but not the magnitude of the fluctuations.

Overall, the model replicates the dynamics of the heating oil spot price and convenience yield well, given the high volatility of these variables over weekly intervals and the sharp movements that occurred during the two simulation periods. The model does not, however, capture the dynamics of inventories very well. This is not surprising given that eqns. (19) and (22) explain, respectively, first- and second-differences of inventories, so that prediction errors in the level of inventories will accumulate over time.

I also used the model to examine the impact of a change in volatility. To do this, I repeated the second simulation, adding a shock to volatility. Specifically, I increased the entire trajectory of volatility from November 29, 2000 to January 31, 2001, by .0448, which is one standard deviation of the level of volatility over the entire 1984–2001 sample, and then re-solved eqns. (8), (19) and (22). The results are included in Figures 6, 7, and 8.

Note that this shock to volatility has a substantial effect on convenience yield, but only a small effect on the spot price and inventories. Convenience yield increases because the value of storage depends directly on volatility. The increase in convenience yield leads to a small increase in inventories. The increase in volatility also increases the marginal opportunity cost of production and thus the spot price, but the effect is small.

6   Conclusions.

This paper examines the role of volatility as a determinant of commodity market dynamics. In principle, volatility should affect market variables through the marginal value of storage

and through the opportunity cost component of marginal cost. For the petroleum complex, and in particular for heating oil, changes in volatility do influence market variables, although the effects are not large. As for volatility itself, market variables do little to explain its behavior. Volatility can be forecasted, but based largely on its own past values.

The estimation results presented here give partial support to the theory of commodity price dynamics presented in the beginning of the paper. For heating oil, the results fit the theory very well — all estimated coefficients have the predicted signs and are significant. For crude oil, the opportunity cost variable has the wrong sign, and for gasoline, both volatility and the opportunity cost variable are either insignificant or have the wrong sign.

The partial failure of the model to fit the theory for crude oil and gasoline may have several causes. First, my calculation of the marginal opportunity cost may be over-simplified: I assumed that the spot price is mean-reverting with constant volatility (even though volatility in fact fluctuates), and I used a quadratic approximation to the exact series solution for the option value. Second, these results may simply reflect the high-frequency nature of the data. I estimate a net demand curve based on weekly changes in inventories, and eqn. (22) includes the second differences of inventories and first differences of variables that drive net demand, such as heating degree days, the marginal opportunity cost, and (for gasoline) the price of crude oil. The effects of changes in the opportunity cost on actual production decisions, for example, may occur more slowly than can be captured by the weekly differences that appear in the estimating equations.[10] Finally, it is unclear how much of a commodity's short-run price movements can be explained by a model based on rational optimizing behavior and corresponding shifts of supply and demand in each of two markets. We might expect that some portion of commodity price variation is not based on such "fundamentals," but is instead the result of speculative noise trading or herd behavior, and there is some evidence that this is indeed the case.[11]

---

[10] Also, as Borenstein and Shepard (2002) show, gasoline prices adjust slowly to crude oil price shocks, which may reflect market power in gasoline markets.

[11] For example, Roll (1984) found that only a small fraction of price variation for frozen orange juice can be explained by fundamentals such as the weather, and Pindyck and Rotemberg (1990) found high levels of price correlation across commodities that are inconsistent with prices driven solely by fundamentals.

Appendix: Derivation of Opportunity Cost.

In this Appendix I derive eqn. (11) for the opportunity cost of production, assuming that the spot price P follows the mean-reverting process given by eqn. (10).

Let $V(P)$ be the value of the option to produce a unit of the commodity. It is easily shown that $V(P)$ must satisfy the following equation (see Dixit and Pindyck, 1994):

$$\tfrac{1}{2}\sigma^2 P^2 V''(P) + [r - \rho + \lambda(\mu - P)]PV'(P) - rP = 0 \,, \tag{29}$$

where r is the risk-free rate and ρ is the risk-adjusted return on the commodity. Thus the expected return "shortfall" is $\delta = \rho - \lambda(\mu - P)$. Also, the solution must satisfy the boundary conditions

$$V(P^*) = P^* - c \,, \tag{30}$$

$$V'(P^*) = 1 \,, \tag{31}$$

where P^* is the critical price that triggers production of an incremental unit, and c is marginal cost. The solution to eqn. (29) is:

$$V(P) = AP^\theta h(P) \,, \tag{32}$$

where θ is given by eqn. (13), and, letting $b = 2\theta + 2(r - \rho + \lambda\mu)/\sigma^2$, $h(P) = H(\frac{2\lambda}{\sigma^2}P; \theta, b)$. Here, $H()$ is the confluent hypergeometric function:

$$H(x; \theta, b) = 1 + \frac{\theta}{b}x + \frac{\theta(\theta+1)}{b(b+1)2!}x^2 + \frac{\theta(\theta+1)(\theta+2)}{b(b+1)(b+2)3!}x^3 + \cdots \,.$$

I use a quadratic approximation to $h(P)$:

$$h(P) \approx 1 + \gamma_1 P + \gamma_2 P^2 \,,$$

where γ_1 and γ_2 are given by eqn. (12). Thus $V(P) \approx AP^\theta(1 + \gamma_1 P + \gamma_1\gamma_2 P^2)$. Substituting into boundary conditions (eq:boundary1) and (eq:boundary2) gives two equations in P^* and the constant A. Divide one by the other to eliminate A and rearrange, yielding:

$$P - c = \frac{P(1 + \gamma_1 P + \gamma_1\gamma_2 P^2)}{\theta + \gamma_1(\theta + 1)P + \gamma_1(\gamma_2\theta + 2\gamma_2 + \gamma_1)P^2 + 3\gamma_1^2\gamma_2 P^3 + 2\gamma_1^2\gamma_2^2 P^4} \,.$$

Next, expand the right-hand side of this equation in a Taylor series around $P = \mu$, take a quadratic approximation, and set $c = \mu$ to obtain eqn. (11).

References

Ball, Clifford A., and Walter N. Torous, "The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence," *The Journal of Finance*, December 1999, **54**, 2339–2359.

Borenstein, Severin, and Andrea Shepard, "Sticky Prices, Inventories, and Market Power in Wholesale Gasoline Markets," *RAND Journal of Economics*, Spring 2002, bf 33, 116–139.

Campbell, John Y., "A Variance Decomposition for Stock Returns," *Economic Journal*, 1991, **101**, 157–179.

Campbell, John Y., "Asset Prices, Consumption, and the Business Cycle," National Bureau of Economic Research Working Paper No. 6485, March 1998.

Campbell, John Y., Burton Malkiel, Martin Lettau, and Yexiao Xu, "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, Feb. 2001, **56**, 1–43.

Chambers, Marcus J., and Roy E. Bailey, "A Theory of Commodity Price Fluctuations," *Journal of Political Economy*, 1996, **104**, 924–957.

Considine, Timothy J., "Inventories under Joint Production: An Empirical Analysis of Petroleum Refining," *Review of Economics and Statistics*, 1997, 493–502.

Considine, Timothy J., and Eunnyeong Heo, "Price and Inventory Dynamics in Petroleum Product Markets," *Energy Economics*, 2000, **22**, 527–547.

Dixit, Avinash K., and Robert S. Pindyck, *Investment Under Uncertainty*, Princeton University Press, 1994.

Eckstein, Zvi, and Martin S. Eichenbaum, "Inventories and Quantity-Constrained Equilibria in Regulated Markets: The U.S. Petroleum Industry, 1947–1972," in T. Sargent, Ed., *Energy, Foresight, and Strategy*, Washington, D.C.: Resources for the Future, 1985.

Fleming, Jeff, and Barbara Ostdiek, "The Impact of Energy Derivatives on the Crude Oil Market," *Energy Economics*, 1999, **21**, 135–167.

Franks, Julian R., and Eduardo Schwartz, "The Stochastic Behaviour of Market Variance Implied in the Prices of Index Options," *The Economic Journal*, Nov. 1991, **101**, 1460–1475.

Gibson, R., and Eduardo S. Schwartz, "Stochastic Convenience Yield and the Pricing of Oil Contingent Claims," *Journal of Finance*, 1990, **45**, 959–976.

Hull, John, and Alan White, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, June 1987, **42**, 281–300.

Litzenberger, Robert H., and Nir Rabinowitz, "Backwardation in Oil Futures Markets: Theory and Empirical Evidence," *The Journal of Finance*, December 1995, **50**, 1517–1545.

McDonald, Robert, and David Shimko "The Convenience Yield of Gold," unpublished working paper, Northwestern University, June 1998.

Miron, Jeffrey A., and Stephen P. Zeldes, "Seasonality, Cost Shocks, and the Production Smoothing Model of Inventories," *Econometrica*, 1988, **56**, 877–908.

Pindyck, Robert S., "The Present Value Model of Rational Commodity Pricing," *The Economic Journal*, 1993, **103**, 511–530.

Pindyck, Robert S., "Inventories and the Short-Run Dynamics of Commodity Prices," *RAND Journal of Economics*, Spring 1994, **25**, 141–159.

Pindyck, Robert S., and Julio J. Rotemberg, "Dynamic Factor Demands and the Effects of Energy Price Shocks," *American Economic Review*, December 1983, **73**, 1066–1079.

Ramey, Valerie A., "Nonconvex Costs and the Behavior of Inventories," *Journal of Political Economy*, 1991, **99**, 306–334.

Routledge, Bryan R., Duane J. Seppi, and Chester S. Spatt, "Equilibrium Forward Curves for Commodities," *Journal of Finance*, June 2000, **55**, 1297–1338.

Schwartz, Eduardo S., "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging," *The Journal of Finance*, July 1997, **52**, 923–973.

Schwartz, Eduardo S., and James E. Smith, "Short-Term Variations and Long-Term Dynamics in Commodity Prices," *Management Science*, 2000.

Williams, Jeffrey C., and Brian D. Wright, *Storage and Commodity Markets*, Cambridge University Press, Cambridge England, 1991.

Table 1: Summary Statistics

Variable	Mean	Min	Max	St.dev
A. Crude Oil				
Spot Price, P	21.18	11.04	41.74	5.407
First Future Price, F_1	21.06	11.12	41.15	5.260
Third Future Price, F_3	20.65	11.19	38.20	4.808
1-Month Conv. Yield, ψ_1	.0557	$-.1642$	.7345	.0829
3-Month Conv. Yield, ψ_3	.0694	$-.1661$	.5691	.0976
Volatility, σ	.0572	.0063	.2396	.0381
Stock, N	330.7	282.5	391.9	20.17
Production, X	7.254	5.662	8.969	.9879
Opportunity Cost, ω	67.07	42.07	96.24	9.457
B. Heating Oil				
Spot Price, P	58.09	29.60	114.5	15.02
First Future Price, F_1	57.85	29.84	109.4	14.57
Third Future Price, F_3	57.12	31.12	104.8	13.42
1-Month Conv. Yield, ψ_1	.1263	$-.3612$	3.740	.3211
3-Month Conv. Yield, ψ_3	.1471	$-.0352$	3.602	.3888
Volatility, σ	.0637	.0180	.3143	.0448
Stock, N	120.7	87.20	162.8	16.90
Production, X	3.095	2.119	3.982	.3464
Opportunity Cost, ω	253.2	174.0	332.0	32.80
C. Gasoline				
Spot Price, P	59.90	30.79	112.6	14.38
First Future Price, F_1	59.53	32.58	108.1	13.77
Third Future Price, F_3	58.21	32.40	94.93	11.93
1-Month Conv. Yield, ψ_1	.1621	$-.8439$	1.884	.2918
3-Month Conv. Yield, ψ_3	.2112	-1.001	1.649	.3925
Volatility, σ	.0562	.0179	.2367	.0266
Stock, N	175.7	145.2	212.9	14.65
Production, X	7.288	5.653	8.650	.6092
Opportunity Cost, ω	322.9	236.7	411.1	34.91

Table 2: Marginal Forecasting Significance Levels

Forecasted Variable	P_t	N_t	ψ_t	σ_t
A. Crude Oil				
Spot Price, P_t	—	.003	.000	.000
Inventory, N_t	.001	—	.002	.798
Conv. Yield, ψ_t	.000	.035	—	.196
Volatility, σ_t	.502	.924	.150	—
B. Heating Oil				
Spot Price, P_t	—	.000	.000	.565
Inventory, N_t	.078	—	.000	.541
Conv. Yield, ψ_t	.033	.0000	—	.855
Volatility, σ_t	.867	.651	.453	—
C. Gasoline				
Spot Price, P_t	—	.000	.000	.948
Inventory, N_t	.005	—	.315	.482
Conv. Yield, ψ_t	.087	.000	—	.542
Volatility, σ_t	.001	.433	.000	—

NOTE: For each variable, entries are significance levels for omitting six lags of the variable in the column heading from an unrestricted OLS prediction equation that includes a constant and six lags of each of the four variables, along with the following exogenous variables: monthly dummy variables, and six lags each of the Treasury bill rate, the Baa corporate bond rate, and the exchange-weighted value of the dollar.

Table 3: Forecasting Equation for Volatility

	Crude Oil		Heating Oil		Gasoline	
	GMM	OLS	GMM	OLS	GMM	OLS
σ_{-1}	0.9041	0.9792	1.0195	1.0405	1.0417	1.0231
	(8.30)	(31.21)	(7.73)	(33.18)	(8.71)	(31.08)
σ_{-2}	0.2093	0.0191	−0.0284	−0.0465	0.0457	0.0032
	(1.04)	(0.44)	(−0.15)	(−1.04)	(0.35)	(0.07)
σ_{-3}	−0.1812	−0.0245	−0.0082	−0.0138	−0.1559	−0.0607
	(−0.69)	(−0.56)	(−0.12)	(−0.31)	(−1.18)	(−1.30)
σ_{-4}	0.0541	−0.0030	0.0578	0.0130	0.0878	0.0114
	(0.30)	(−0.07)	(0.50)	(0.29)	(0.69)	(0.24)
σ_{-5}	−0.5559	−0.4583	−0.4191	−0.4463	−0.5364	−0.4373
	(−4.99)	(−10.71)	(−2.45)	(−10.00)	(−3.93)	(−9.40)
σ_{-6}	0.5049	0.4320	0.2892	0.3736	0.4439	0.3802
	(5.89)	(13.82)	(2.40)	(11.9)	(3.28)	(11.46)
$\text{TBILL}_{-1 \text{ to } -6}$	4.16×10^{-4}	3.96×10^{-4}	7.72×10^{-4}	9.88×10^{-4}	7.08×10^{-4}	6.08×10^{-4}
	(0.39)	(1.10)	(0.35)	(1.75)	(0.88)	(1.75)
$\text{NYSE}_{-1 \text{ to } -6}$	3.80×10^{-6}	3.30×10^{-6}	9.80×10^{-6}	4.40×10^{-6}	7.00×10^{-7}	2.90×10^{-6}
	(0.46)	(0.58)	(0.59)	(2.60*)	(0.76)	(1.45)
$\text{EXVUS}_{-1 \text{ to } -6}$	-3.00×10^{-6}	3.10×10^{-6}	-2.30×10^{-5}	-1.74×10^{-5}	-2.41×10^{-5}	-3.08×10^{-5}
	(0.31)	(0.54)	(0.44)	(2.81*)	(0.32)	(2.81*)
$\text{CRB}_{-1 \text{ to } -6}$	1.80×10^{-5}	4.50×10^{-6}	-9.90×10^{-6}	-1.42×10^{-4}	-1.78×10^{-5}	-1.30×10^{-6}
	(0.58)	(1.29)	(0.08)	(0.35)	(0.39)	(0.99)
R^2	.707	.897	.779	.894	.695	.890

Note: Equation is

$$\sigma_t = c_0 + \sum_{j=1}^{11} c_j \text{DUM}_{j,t} + \sum_{j=1}^{6} \beta_{1j}\sigma_{t-j} + \sum_{j=1}^{6} \beta_{2j}\text{TBILL}_{t-j}$$

$$+ \sum_{j=1}^{6} \beta_{3j}\text{NYSE}_{t-j} + \sum_{j=1}^{6} \beta_{4j}\text{EXVUS}_{t-j} + \sum_{j=1}^{6} \beta_{5j}\text{CBR}_{t-j}$$

For lags of volatility, i.e., σ_{t-j}, table shows estimates of β_{1j} and t-statistics. For TBILL, NYSE, EXVUS, and CBR, table shows $\sum_j \beta_{kj}$ and F-statistics for groupwise significance, with an asterisk indicating significance at the 5% level. The DUM_{jt} are monthly time dummies; estimates of the c_j are not shown.

Table 4: Estimates of Three-Equation System

	Crude Oil NOB = 881	Heating Oil NOB = 881	Gasoline NOB = 829 (1)	(2)
c_0	53.63 (20.69)	−30.48 (−5.33)	5.409 (0.99)	10.98 (2.03)
c_1	0.426 (4.76)	1.478 (10.16)	1.568 (11.00)	1.541 (11.79)
c_2	—	2.927 (43.35)	2.780 (46.67)	2.728 (45.34)
c_3	−0.594 (−12.53)	0.946 (4.31)	−0.018 (−1.06)	−0.035 (−2.15)
c_4'	0.001 (0.97)	0.016 (8.74)	−0.007 (−4.75)	−0.007 (−5.13)
c_5'	0.001 (0.20)	0.001 (0.32)	0.006 (1.67)	0.006 (1.69)
c_6'	0.015 (7.97)	−0.012 (−4.38)	0.007 (2.72)	0.009 (4.10)
α_1	0.932 (22.63)	0.951 (15.4)	0.841 (21.61)	0.633 (12.7)
α_2	0.097 (8.56)	0.092 (7.10)	0.036 (3.85)	0.003 (0.25)
α_3	1.946 (9.11)	2.196 (12.03)	−0.341 (−2.73)	1.1
J	.0905	.0882	.0992	.1069

Note: Table shows GMM estimates of eqns. (19), (22), and (8), with t-statistics in parentheses. Estimates of 22 parameters for monthly dummy variables are not shown. For gasoline, the estimate of α_3 is negative, so the model is re-estimated with α_3 constrained to equal 1.1. The J-statistics are distributed as $\chi^2(71)$ for crude oil and $\chi^2(70)$ for heating oil and gasoline; the critical 5% values are 91.40 and 90.32 respectively.
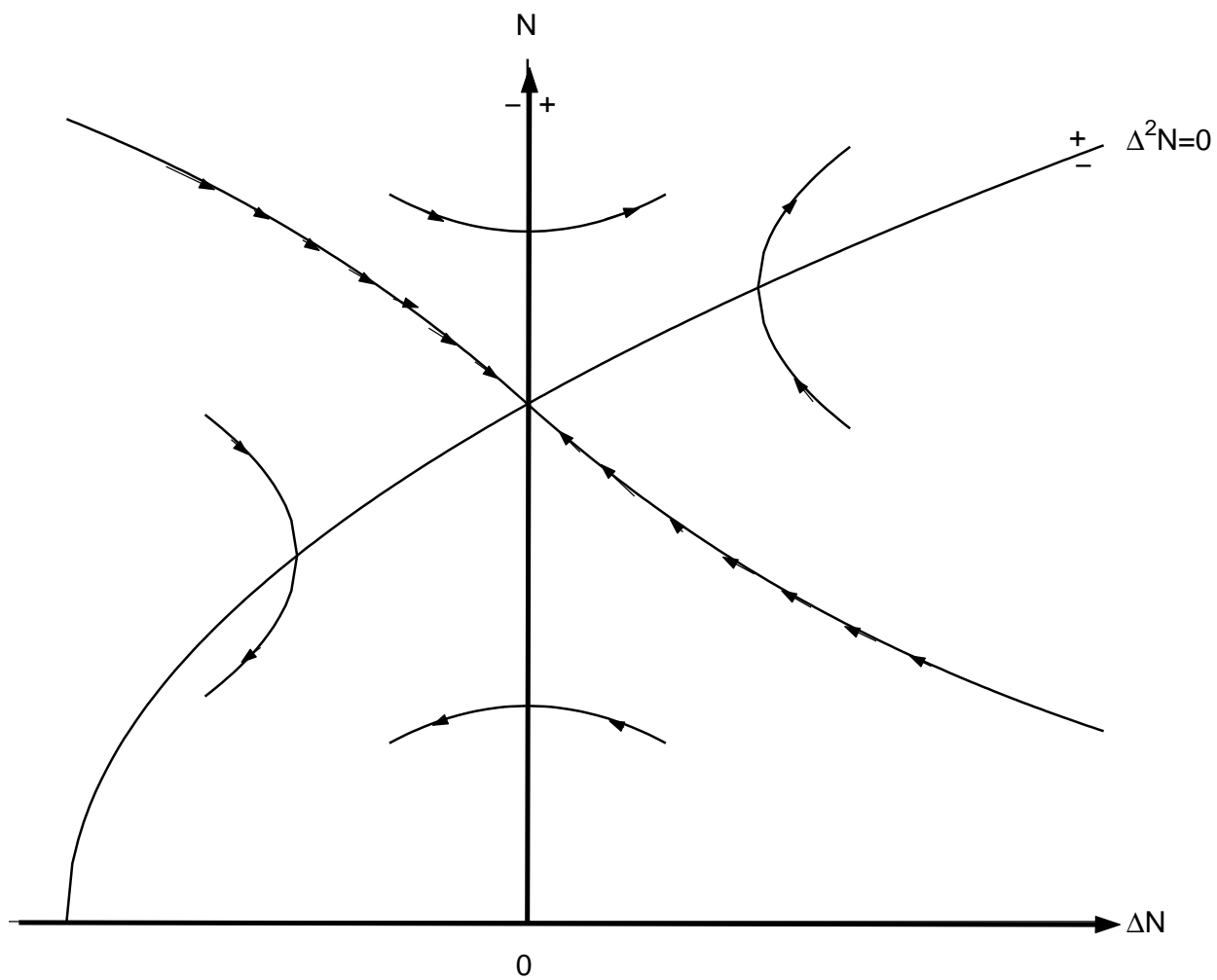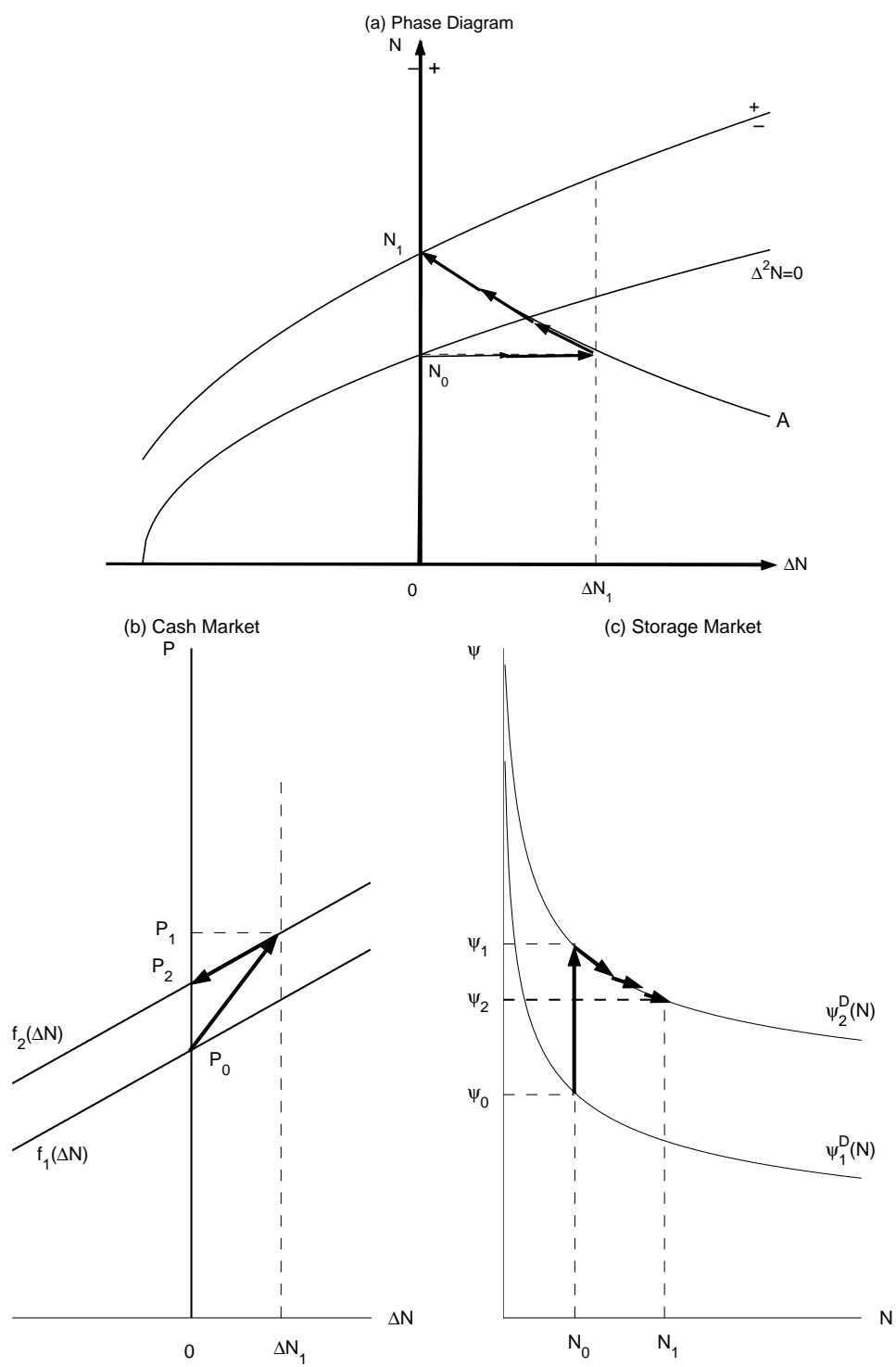
Figure 1: Phase Diagram.

31

(a) Phase Diagram

(b) Cash Market
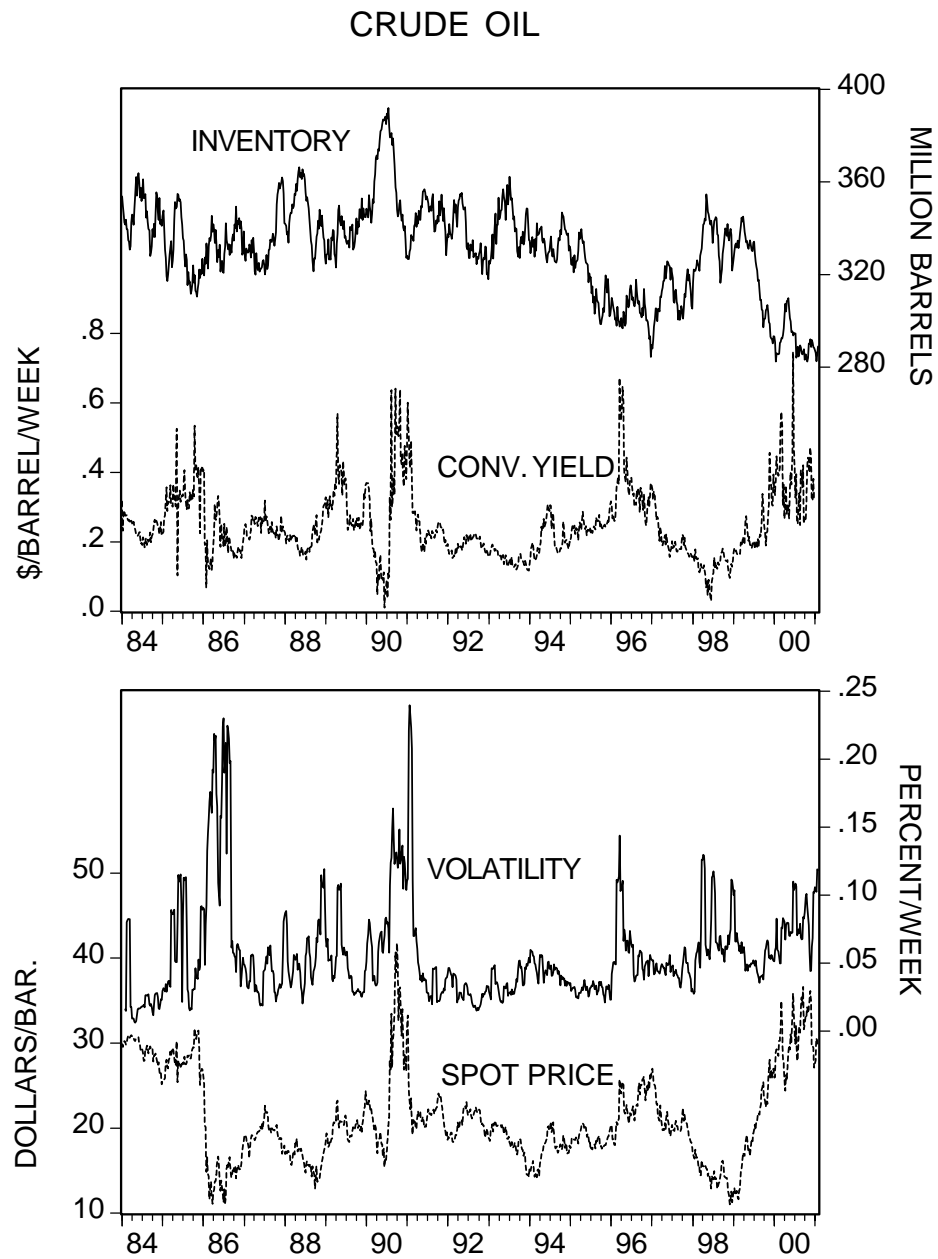
(c) Storage Market

Figure 2: Increase in Volatility.

Figure 3: CRUDE OIL.

33
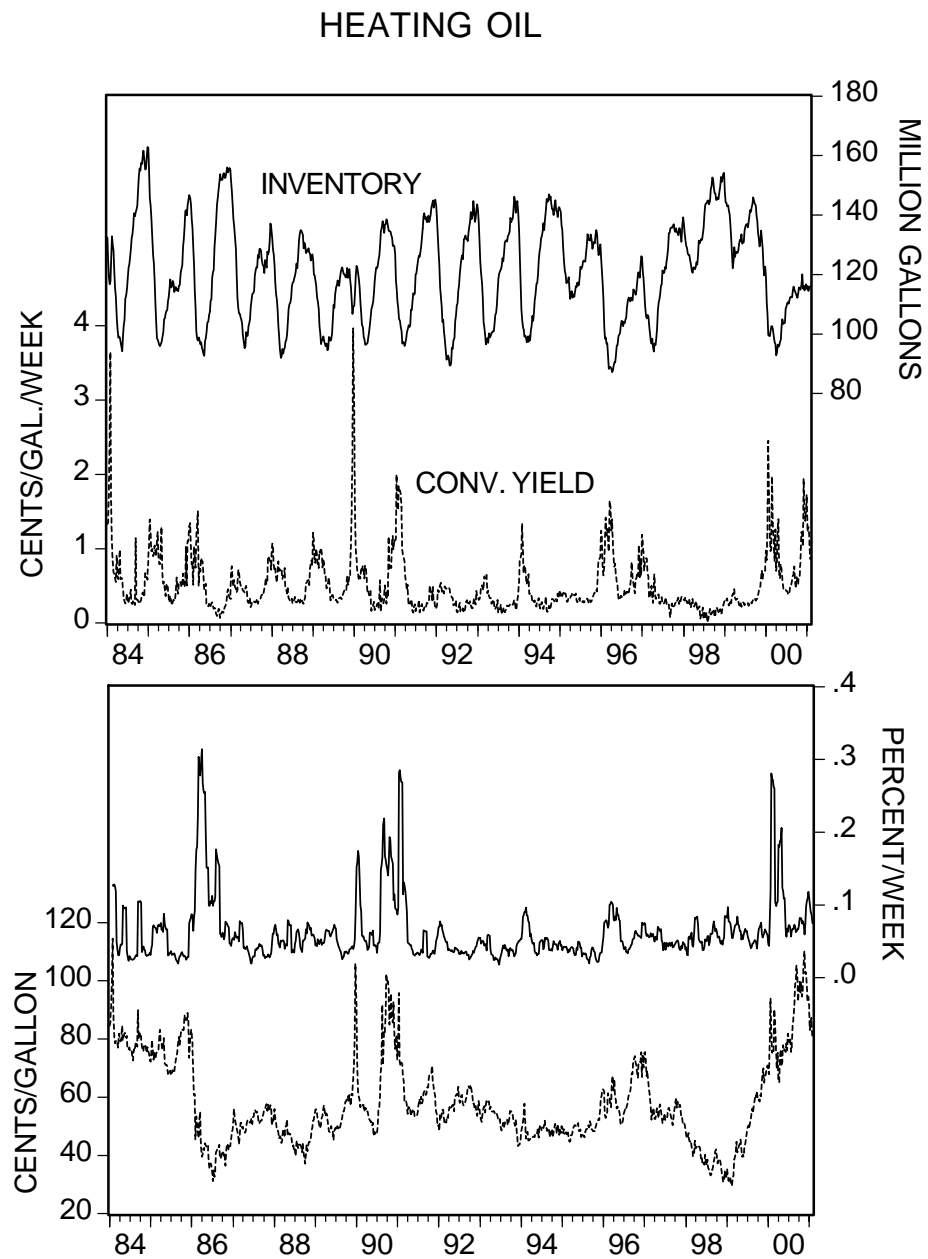
HEATING OIL

Figure 4: HEATING OIL.

GASOLINE

Figure 5: GASOLINE.

35
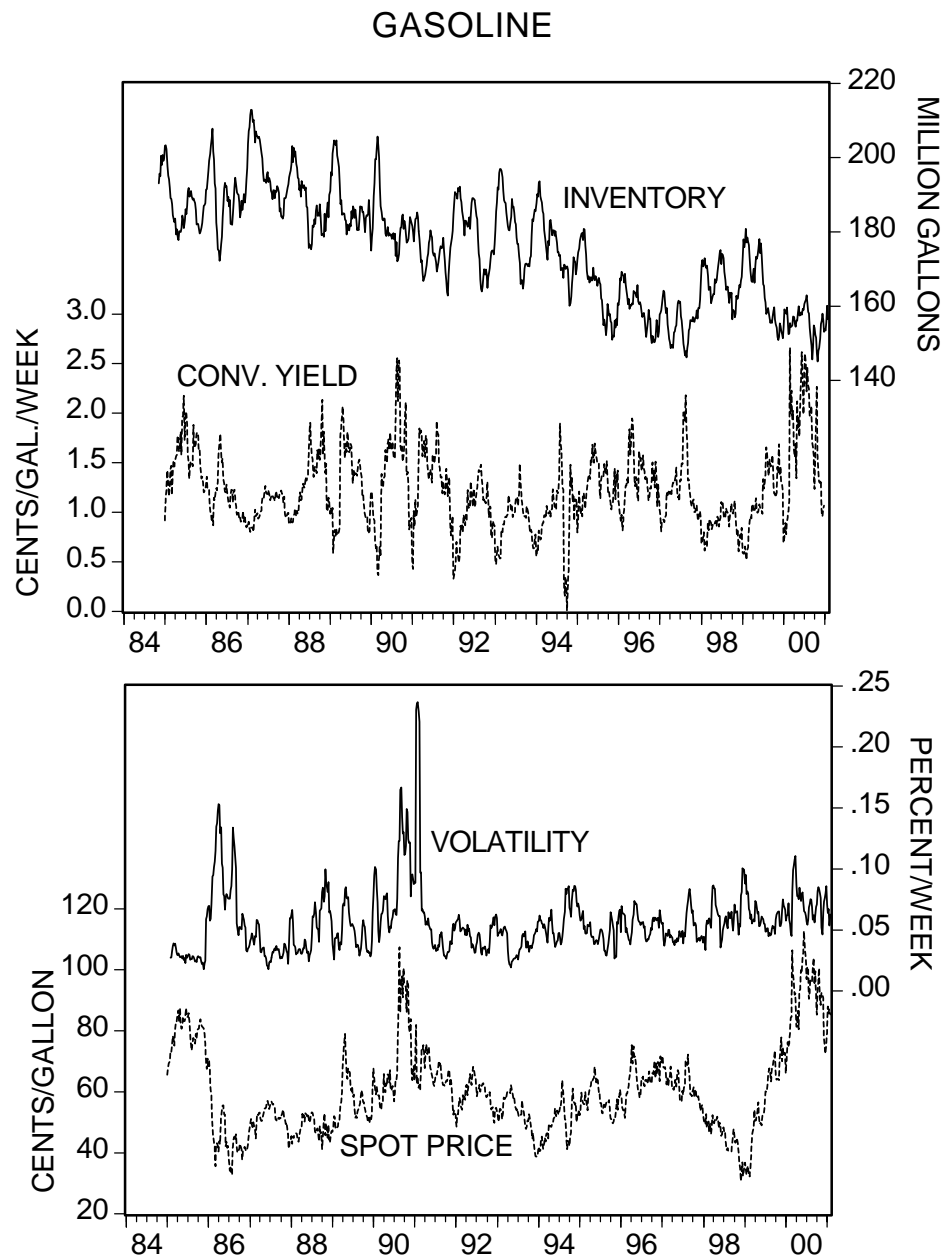
FIGURE 6: HEATING OIL SPOT PRICE

CENTS/GALLON

SIMULATED

ACTUAL

SHOCK

SIM.

ACTUAL

8/1/1990 9/5/1990 12/6/2000 1/3/2001
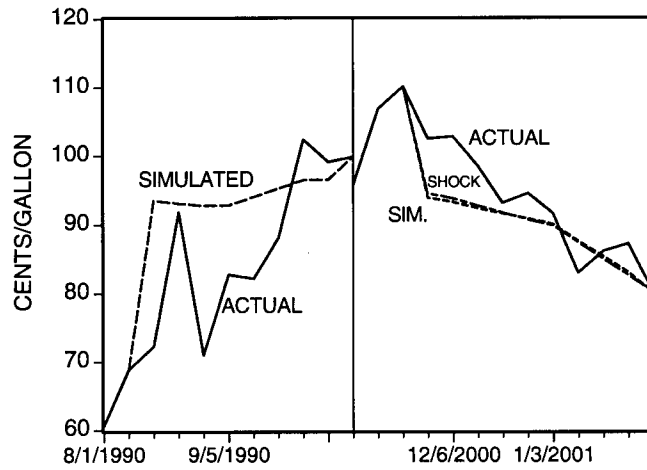
FIGURE 7: HEATING OIL INVENTORIES

MILLION GALLONS

SIMULATED
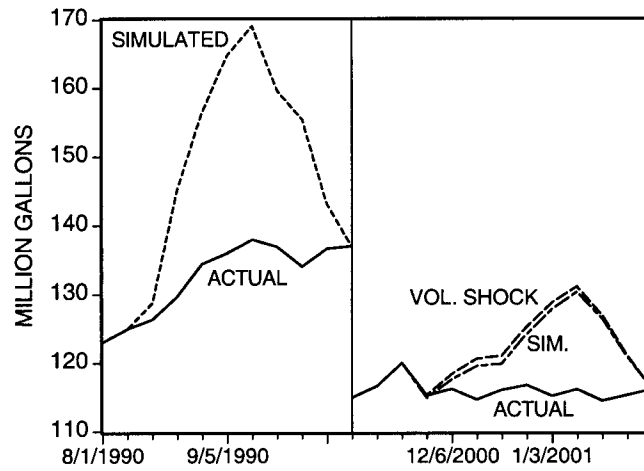
ACTUAL

VOL. SHOCK

SIM.

ACTUAL

8/1/1990 9/5/1990 12/6/2000 1/3/2001

FIGURE 8: HEATING OIL CONVENIENCE YIELD

CENTS/GAL./WEEK

ACTUAL

SHOCK

SIMULATED

SIM. ACTUAL

8/1/1990 9/5/1990 12/6/2000 1/3/2001

36