ABSTRACT

JI, WEN. Essays on Econometric Evaluation of Models of Commodity Futures Prices. (Under the direction of Atsushi Inoue, Paul Fackler, Sastry Pantula.)

This dissertaion is comprised of two essays on econometric evaluation of models of commodity futures prices. The first essay develops a frequency-domain volatility bound approach that can be used to evaluate possibly misspecified models. The proposed method allows us to detect model failures at specific frequencies, for example, the seasonal frequencies, the business cycle frequencies, etc. This is particularly useful when the data exhibit significant cyclical behavior. As an application of the proposed method, the consumption based capital asset pricing models for commodity futures are evaluated using crude oil and corn futures price data. The equilibrium price conditions are derived. Empirical results overwhelmingly reject the consumption based capital asset pricing models across all frequencies. The second essay proposes an econometric method that can be used to determine the appropriate number of factors in affine term structure models. The proposed method estimates the affine model by solving a nonlinear least squares problem, and the number of factors is determined by minimization of a proposed model

selection criteria. Simulation study shows that the proposed method can identify the right number of factors under general conditions. Finally, the empirical issue of how many number of factors are needed for crude oil and corn futures prices is examined using the proposed method.

ESSAYS ON ECONOMETRIC EVALUATION OF MODELS OF COMMODITY FUTURES PRICES

by

WEN JI

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

ECONOMICS
and
STATISTICS

Raleigh
2003

APPROVED BY:

_____	_____
Co-chair of Advisory Committee	**Co-chair of Advisory Committee**
_____	_____
Co-chair of Advisory Committee	

To my parents

with love

BIOGRAPHY

Wen Ji was born on December 16, 1970 in Tianjin, China. She received her Bachelor's degree and Master's degree in Economics at Dongbei University of Finance and Economics in Dalian, China in 1992 and 1995 respectively. She was a lecturer at Dongbei University of Finance and Economics from 1995 to 1996. She received a Master's degree in Economics at the University of Connecticut in August of 1997. She then began her graduate studies at North Carolina State University, received a Master's degree in Statistics in 2000, and passed her final oral exam for Ph.D. in Economics and Statistics in May of 2003.

ACKNOWLEDGEMENTS

I would like to thank the members of my advisory committee for their helpful advice. I thank Atsushi Inoue, for his advice and guidance, for his patience, and for being accessible any time I needed help. I thank Paul Fackler for his helpful advice and for the time he spends with me on my research problems. I thank Sastry Pantula for his advice and comments on the theoretical problems, and for the time he spent in helping me. I also thank David Dickey for his help on this study. Without the helpful advice from my committee members, completion of this dissertation would be impossible.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

This dissertation focuses on econometric evaluation of models of commodity futures prices. The dissertation consists of two essays.

In the first essay, we propose a spectral volatility bound approach and evaluate consumption-based capital asset pricing models (CAPMs) for conam-modity futures.

Evaluation of dynamic structural models is a subject that receives much attention in economics literature. One approach that has been widely used to evaluate dynamic equilibrium models is the Hansen-Jagannathan variance bound approach. Hansen and Jagannathan (1991) derived a lower bound on the volatility of a representative consumer's intertemporal marginal rate of

substitution (IMRS). A model is said to be admissible if the volatility of the IMRS implied by the model is greater than the Hansen-Jagannathan volatility bound.

In this paper, we propose a frequency-domain variance bound that allows us to determine if models fail to capture important aspects of the data at some specific frequencies, for example, the long run frequencies, the seasonal frequencies, the business cycle frequencies, etc. This is particularly useful in modeling prices that exhibit significant cyclical behavior. Hence, the frequency-domain volatility bound provides additional explanations of failures of asset pricing models, and give further guidance on improving the models. The Hansen-Jagannathan variance bound test has been applied to a wide class of asset pricing models.

As an application, we evaluate the consumption-based capital asset pricing models for futures market data using the spectral volatility bound. We derive testable restrictions of the model from representative consumer's dynamic optimization problem. These restrictions implied by the consumption CAPMs are evaluated.

In the second essay, we deal with another class of models, namely, factor models. The factor models have enjoyed great popularity since their

appearance in finance literature. This class of models is parsimonious and computationally tractable. In this thesis, we concentrate on factor models with affine term structure of commodity futures, and examine the method of determining the number of factors in such models.

Undoubtedly, the issue of determining the number of factors is very important to the correct model specification in factor models. Though unfortunately, many studies of factor models have ignored this empirical issue and assumed the number of factors in their application. Statistical methods to estimate and determine the number of factors can be classified into three categories: the maximum likelihood estimation method, the principal component analysis, and the minimum distance method. In the second essay, we propose a method to estimate affine term structure models of futures prices, and develop testing methodologies to determine the appropriate number of factors in the data. We propose to estimate the affine model by solving a nonlinear least squares problem, and determine the number of factors by minimization of a model selection criteria.

We apply the proposed method to crude oil and corn futures prices, and examine the issue of how many number of factors are needed for crude oil and corn futures prices.

The rest of the thesis is organized as follows: chapter two develops the frequency-domain volatility bound approach and evaluates consumption-based CAPMs of futures prices. Chapter three proposes method to determine the number of factors in an affine term structure model of futures prices, and examines the empirical issue for crude oil and corn futures prices data.

Chapter 2

Can the Consumption-Based Capital Asset Pricing Model Explain Prices of Commodity Futures?

2.1 Introduction

A well-known result of asset pricing models with frictionless markets is that asset prices can be represented by a stochastic discount factor or pricing ker-

nel. A stochastic discount factor (SDF) is a random variable that can be used to compute market prices today by discounting, state-by-state, the corresponding payoffs at a future date. Since different asset pricing models imply different SDFs, they imply different asset prices. Consequently, misspecification of the SDF can be used to evaluate asset pricing models. One example of such evaluation methods is the Hansen-Jagannathan volatility bound approach (Hansen and Jagannathan, 1991). The Hansen-Jagannathan bound (HJ bound) is a lower bound on the standard deviation of the SDF as a function of its mean. Any admissible model SDF should lie inside the feasible region in the mean-standard deviation plane defined by the HJ bound. The HJ bound approach is nonparametric and allows the model SDF to be misspecified. It applies to a wide class of economic models, and hence has been commonly used in finance as a model diagnostic test.

In this paper, we develop the volatility bounds in the frequency domain. Frequency-domain analysis has become frequently used in recent macro econometrics literature, e.g., Watson (1993), Diebold, Ohanian and Berkowitz (1998), while its application in finance has been limited. The graphic nature of frequency analysis is particularly useful in dealing with commodities that exhibit significant cyclical behavior. For example, when

modeling commodity futures prices, we may want to know if a specific model fails because it fails to capture seasonal frequency behavior, or because it overlooked the business cycle (long run) frequency, etc. Such information cannot be obtained by the time-domain approach. The frequency-domain volatility bounds approach developed in this paper complements the time-domain approach and provides additional explanations of failures of some asset pricing models and give further guidance on improving the models.

As an application of the proposed frequency-domain variance bounds approach, we evaluate consumption-based Capital Asset Pricing Models(CAPM) using futures prices. The consumption-based CAPM of Merton (1973) and Breeden (1979), have been extensively studied in the asset pricing literature. The empirical evidence in the equity markets has overwhelmingly rejected the consumption-based CAPM with time and state separable preferences, e.g., Hansen and Singleton (1983), Mehra and Prescott (1986). The consumption-based CAPM with time separable preferences implies SDF that lies far outside the admissible region defined by the HJ bound. Such evidence has motivated some researchers to develop models with alternative preference structure which aim at implying more variable SDF.

For example, the habit formation preferences of Constantinides (1990)

and Sundaresan (1989), the non-expected utility or recursive preferences of Epstein and Zinn (1989, 1991) and Weil (1989), and the preferences with seasonal habit formation of Ferson and Harvey (1992). These models with alternative preference structure have been shown to perform reasonably better than the time-separable preferences using equity market data.

However, there has been no empirical evidence suggesting their relative performance in the commodity futures market. Extant empirical studies of the consumption-based CAPM in the commodity futures markets have been quite limited. Examples of such work include Dusak (1973) who estimates market betas and Breeden (1980) who estimates consumption betas. No empirical work exist regarding the newly developed alternative preferences using futures prices. In this paper, we fill in this gap in literature by evaluating the empirical validity of these different versions of consumption-based CAPM in the commodity futures markets. Since there are significant institutional differences between equity markets and commodity futures markets, the pricing relations implied by the consumption-based CAPM can not be used directly in the futures markets. We derive the corresponding pricing relations in the futures markets from a model of representative agent's dynamic optimization problem. The resulting pricing restrictions are shown to have an SDF

representation. Hence the time-domain variance bound and the proposed frequency-domain variance bounds approach can be applied.

The rest of the paper is organized as follows. Section 2 considers the determination of futures prices in a simple pure exchange economy. Testable pricing restrictions are derived from representative agent dynamic optimization problem. These are illustrated in the context of four different specifications of utility functions, which include constant relative risk aversion (CRRA) utility, recursive utility, habit persistence utility, and seasonal habit persistence utility function. Section 3 develops the frequency-domain version of Hansen-Jagannathan (1991) volatility bounds approach. Section 4 applies the proposed methodology to crude oil and corn futures prices. We estimate and evaluate the four versions of CAPM models and a simple linear stochastic discount factor (SDF) model. Section 5 concludes.

2.2 Models

In this section, we consider the determination of equilibrium futures prices in a model that is a slight modification of Lucas' (1978) one-good pure exchange economy. In this economy, there is a tree that produces random amounts of

fruits and there is a competitive stock market of ownership in this tree. An asset is a claim to the fruits of the tree. The fruits are perishable and follow a stationary Markov process. There are a large number of consumers with identical preferences and identical endowments. The representative agent model described here itself is not needed to obtain the general no arbitrage equilibrium conditions, however, it is used to derive specific forms for the stochastic discount factor for the models considered.

The representative consumer maximizes his expected lifetime utility expressed as

$$E\left\{\sum_{t=1}^{\infty}\beta^t\, u(\cdot)\right\} \tag{2.1}$$

where $u(\cdot)$ is a current period utility function, β is a discount factor, and $E\{\cdot\}$ is an expectation operator.

The only asset we consider in this economy is the futures contract for the fruits. Adding additional asset to the model will imply additional pricing relations. In this analysis, we limit to the one asset model. There is a futures market where futures contracts for the fruits are made each period. The futures contract made at date t is a commitment to pay an amount f_t at date $\tau_t = \min_{n\in I_+}\{T+12n\ :\ t \le T+12n\}$ in return for a fruit at date τ_t, where $0 \le T < 12$. In the definition of τ_t, we have assumed that time to

contract expiration is always less than 12 dates for notational simplicity. Let f_{t+1} denote the futures settlement price at date t for delivery of a fruit at date τ_t and q_t denote the corresponding futures position at the end of date t. The futures settlement price f_{t+1} is unknown to the representative agent at date t.

Two institutional features of the futures market will affect the representative agent's budget constraint. First, the representative agent (possibly partially) offsets her futures position of amount Δq_t at price f_t during date t. Thus, her position evolves as

$$q_t = q_{t-1} + \Delta q_t. \tag{2.2}$$

Second, the representative agent pays variation margins at every date. Let c_t denote consumption, p_t denote the asset price, x_t denote the asset holding, and z_t denote the amount of fruits on date t. To simplify the notation, we assume that variation margin is 100α percent and that the spot price is normalized to unity. Then the representative consumer's budget constraint at date t is given by

$$c_t + p_t x_{t+1} \leq (p_t + z_t)x_t + (f_t - f_{t-1})q_{t-1} \tag{2.3}$$

for $t = 0, 1, 2, ...$ with $q_{-1} = 0$.

Given the initial asset holding x_0, the representative agent maximizes the expected utility at date 0 (2.1) subject to a sequence of her budget constraints (2.3). Because all the consumers and endowments are identical, no trade occurs in equilibrium if it exists.

Then, in the economy described as above, the general no arbitrage condition implies that

$$E_t[(\frac{p_{t+1} + z_{t+1}}{p_t} m_{t+1})] = 1 \tag{2.4}$$

and

$$E_t[(f_{t+1} - f_t)m_{t+1}] = 0 \tag{2.5}$$

where m_{t+1} is the stochastic discount factor (SDF), and $E_t[\cdot] = E[\cdot|\Omega]$ is the conditional expectation operator. Equations (2.4) and (2.5) are the no arbitrage restrictions that allow us to determine futures prices in this economy. If either of them is violated, an arbitrage opportunity exists that allows everyone in the economy to increase her utility (a free lunch). For example, if $Pr(m_{t+1} > 0) = 1$ and $E_t[(f_{t+1} - f_t)m_{t+1}] > 0$, where $Pr(A)$ represents the probability that event A occurs, then a consumer can buy a futures contract and increase her expected utility at date t. If $Pr(m_{t+1} > 0) = 1$ and $E_t[(f_{t+1} - f_t)m_{t+1}] < 0$, then a consumer will sell a futures contract and increase her expected utility at date t. In fact, each one will do so in unlimited

quantities! Because there is no arbitrage in equilibrium, (2.5) must hold in the economy. Note that when m_{t+1} is adapted to the information set at date t (e.g., constant), then (2.5) becomes "the martingale hypothesis" which has been extensively studied in the literature (see Raynauld and Tissier, 1984, for example). Many existing literature examine the pricing restriction (2.4). However, the restriction (2.5) has not been explored except for the simplest case of martingale hypothesis. In the equilibrium model, restriction (2.5) can be interpreted that the change in futures prices is conditionally orthogonal to the gradient of the utility function since in equilibrium models, the SDF m_{t+1} is the intertemporal rate of substitution.

In order to derive a specific functional form of m_{t+1}, we need to specify the representative agent's preferences. In the rest of this section, we derive the pricing restrictions (2.4) and (2.5) under four different preference structures. First, we consider the time- and state-separable preferences with constant coefficient of relative risk aversion (CRRA). It is well known that the intertemporal elasticity of substitution is equal to the reciprocal of the coefficient of relative risk aversion. Next, we consider three alternative preferences developed in the asset pricing literature, which include the non-expected utility preferences of Epstein and Zinn (1989, 1991) and Weil (1989), the preferences

with habit formation of Constantinides (1990) and Sundaresan (1989), and the preferences with seasonal habit formation of Ferson and Harvey (1992). A common feature of these alternative preferences is that they try to disentangle the two preference parameters, namely, the intertemporal elasticity of substitution and the coefficient of relative risk aversion.

2.2.1 Time- and State-Separable Utility

Consider the time- and state-separable utility functions, i.e.,

$$U_0 = \sum_{t=0}^{\infty} \beta^t u(c_t) \tag{2.6}$$

where $0 < \beta < 1$. The representative agent maximize her lifetime utility (2.6) subject to her budget constraint (2.3). Solving the dynamic optimization problem, we obtain the pricing restrictions expressed as

$$E_t \left[(\frac{p_{t+1} + z_{t+1}}{p_t}) \beta \left(\frac{u'(c_{t+1})}{u'(c_t)} \right) \right] = 1 \tag{2.7}$$

and

$$E_t \left[(f_{t+1} - f_t) \beta \left(\frac{u'(c_{t+1})}{u'(c_t)} \right) \right] = 0. \tag{2.8}$$

These correspond to (2.4) and (2.5) with $m_{t+1} = \beta u'(c_{t+1})/u'(c_t)$.

If we set $u(c) = c^{1-\gamma}/(1-\gamma)$, where γ is the coefficient of relative risk aversion, we obtain constant relative risk aversion (CRRA) utility function. In fact, this is the most commonly used utility function in empirical studies. The pricing restrictions equation then further simplifies to

$$E_t\left[(\frac{p_{t+1}+z_{t+1}}{p_t})\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\right]=1 \tag{2.9}$$

and

$$E_t\left[(f_{t+1}-f_t)\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\right]=0. \tag{2.10}$$

Details of the derivation for the general time- and state-separable utility functions are provided in Appendix 2A.

2.2.2 Non-Expected Utility

Epstein and Zinn (1989, 1991) propose a recursive, but not necessarily expected utility, preferences model. In such models, systematic risk of an asset is determined by covariance with both the return to the market portfolio and consumption growth.

The recursive preferences can be specified as in Weil (1991):

$$U_t = W(c_t, E_t U_{t+1}), \tag{2.11}$$

where

$$W(c,U) = \frac{\left\{(1-\beta)c^{1-\rho} + \beta[1+(1-\beta)(1-\gamma)U]^{\frac{1-\rho}{1-\gamma}}\right\}^{\frac{1-\gamma}{1-\rho}} - 1}{(1-\beta)(1-\gamma)} \tag{2.12}$$

for $\gamma \neq 1$ and $\rho \neq 1$. ρ is the inverse of the elasticity of intertemporal substitution, and γ is the coefficient of relative risk aversion. Thus it is easily seen that in this kind of models, the elasticity of substitution and risk aversion parameter are separated. The first-order condition and the envelope condition imply

$$E_t[(f_{t+1} - f_t)W_1(c_{t+1}, E_{t+1}V_{t+1})] = 0, \tag{2.13}$$

where W_i is the partial derivative of W with respect to its i-th argument. Following Weil(1989), the marginal rate of substitution for this model can be expressed as

$$m_{t+1} = \frac{W_2(c_t, E_tU_t)W_1(c_{t+1}, E_{t+1}U_{t+1})}{W_1(c_t, E_tU_t)}. \tag{2.14}$$

From equation (A9) of Weil (1989), this is equivalent to

$$m_{t+1} = \left[\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\rho}\right]^{\frac{1-\gamma}{1-\rho}} R_{t+1}^{\frac{1-\gamma}{1-\rho}-1} \tag{2.15}$$

where R_t is the market return.

Substituting the expression for the marginal rate of substitution into (2.4), (2.5), we obtain corresponding conditions for this type of utility functions.

2.2.3 Habit Persistence Utility

The preferences with habit persistence are introduced by Constantinides (1990) and Sundaresan (1989). Habit persistence is modeled by assuming that consumer's subsistence level is a weighted sum of past and current consumption expenditures. Such preferences imply that the utility for consumption expenditure is not time-separable.

The utility function can be written as the following (see Ferson and Constantinides (1991)),

$$U_0 = \sum_{t=0}^{\infty} \beta^t \frac{(c_t - \delta c_{t-1})^{1-\gamma} - 1}{1-\gamma}. \tag{2.16}$$

where γ is the coefficient of relative risk aversion, δ is the habit persistence/durability parameter, $\delta > 0$ indicates habit persistence, and $\delta < 0$ indicates durability.

Solving the representative agent's utility maximization problem, we ob-

tain the pricing restrictions (2.4) and (2.5) with

$$m_{t+1} = \beta \frac{(c_{t+1} - \delta c_t)^{-\gamma} - \beta\delta E_{t+1}(c_{t+2} - \delta c_{t+1})^{-\gamma}}{(c_t - \delta c_{t-1})^{-\gamma} - \beta\delta E_t(c_{t+1} - \delta c_t)^{-\gamma}}, \tag{2.17}$$

Details of the derivation is provided in Appendix 2B.

2.2.4 Seasonal Habit Persistence Utility

The preferences with seasonal habit persistence are proposed by Ferson and Harvey (1992). The utility specification is based on the habit persistence utility, but augmented with seasonal considerations. In this model, the consumer's subsistence level depends on a previous consumption level only in the same season. Such specification implies seasonal time-nonseparability.

The seasonal habit persistence utility function can be written as:

$$U_0 = \sum_{t=0}^{\infty} \beta^t \frac{(c_t - \delta c_{t-12})^{1-\gamma} - 1}{1-\gamma} \tag{2.18}$$

Similar to the habit persistence preferences, $\delta > 0$ indicates seasonal habit persistence, and $\delta < 0$ indicates seasonal durability.

Solving the dynamic optimization problem yields the pricing restrictions (2.4) and (2.5) with corresponding SDF given by

$$m_{t+1} = \beta \frac{(c_{t+1} - \delta c_{t-11})^{-\gamma} - \beta^{12}\delta E_{t+1}(c_{t+13} - \delta c_{t+1})^{-\gamma}}{(c_t - \delta c_{t-12})^{-\gamma} - \beta^{12}\delta E_t(c_{t+12} - \delta c_t)^{-\gamma}}. \tag{2.19}$$

The four models described above will be referred to as Models 1, 2, 3, and 4 respectively.

2.3 Econometric Methodology

In this section, we extend the Hansen and Jagannathan (1991) volatility bounds to the frequency domain. To make the paper self contained, we first describe the time-domain Hansen-Jagannathan variance bounds. We then present the frequency domain version of the volatility bounds.

2.3.1 The Time-domain Volatility Bound

Consider a general pricing relation:

$$q_t = E_t(x_{t+1}m_{t+1}) \tag{2.20}$$

where x_{t+1} is an n-dimensional vector of asset payoffs at date $t+1$, q_t is the random vector of prices of these assets, m_{t+1} is the stochastic discount factor which captures corrections for risk.

Taking expectations on both sides of equation (2.20), we obtain the unconditional version of the pricing relation:

$$E(q_t) = E(x_{t+1}m_{t+1}) \tag{2.21}$$

Equation (2.21) says that the market value today of an uncertain payoff tomorrow is represented by multiplying the payoff by the discount factor and adding across states of nature using the underlying probabilities as argued by Hansen and Jagannathan (1997).

First, assume market complete, we consider derivation of the bounds in two cases, with or without risk free asset. Define

$$P = \{c \cdot \ x : c \in \Re^n\} \tag{2.22}$$

First consider the case when P contains a payoff that is equal to 1 with probability 1. Note that the existence of a riskless asset is not needed in the derivation of the H-J bound. It is only for ease of derivation. The case of no risk free asset will be considered next. Then finding a random variable m^* in P that satisfies equation (2.21) amounts to finding a vector α_0 in $\Re^n$ such that

$$E(xx^T\alpha_0) = Eq \tag{2.23}$$

where $m^* = x \cdot \alpha_0$. Solving (2.23) yields $\alpha_0 = (Exx^T)^{-1}Eq$. Note here that nonsingularity of the second moment of x is assumed. Since P contains a unit payoff, we have $Em = Em^*$. Because $E[x(m - m^*)] = 0$, the difference between m and m^* is orthogonal to the random vector x. Since m^* is in P, m^* is the least squares projection of m onto P, and hence, $Cov(m^*, m - m^*) = 0$, thus

$$\sigma^2(m) = \sigma^2(m^*) + \sigma^2(m - m^*) \tag{2.24}$$

This in turn implies that

$$\begin{aligned} \sigma(m) &\geq \sigma(m^*) \\ E(m^*) &= E(m) \end{aligned} \tag{2.25}$$

So $\sigma^2(m^*)$ is the lower bound on the variance of any m that satisfies the pricing relation (2.21).

Now consider the case when there is no riskless payoff. Then in order to derive a similar bound as above we can assign a number μ to a hypothetical

value of the expected price of a unit payoff. Thus for any μ, we construct a random variable m_μ in P such that,

$$E(xm_\mu) = Eq,\ Em_\mu = \mu \tag{2.26}$$

and obtain volatility bound

$$\sigma(m) \geq \sigma(m_\mu) \tag{2.27}$$

Replicate the construction of m_μ for all real number μ, we get an indexed collection $\{m_\mu : \mu \in \Re\}$ of random variables, each of which satisfies equation (2.26). So for any m satisfying the pricing relation (2.26), $[Em, \sigma(m)]$ belongs to the region $S = \{(\mu, w) \in \Re^2 : w \geq \sigma(m_\mu)\}$. This region is referred to as the mean-standard deviation frontier. Thus, it is easily seen that one criteria to assess pricing models is to compare the model implied SDF variance with the variance bound at the corresponding mean. If the model implied SDF has variance smaller than the variance bound, then the model is not admissible.

The above argument also implies that the minimum variance stochastic discount factor m^* at any specified mean μ can be obtained by solving

$$\begin{aligned} \min_m \quad & Em^2 \\ s.t. \quad & Exm = Eq \\ & Em = \mu \end{aligned} \tag{2.28}$$

The solution to this problem can be shown to be equal to

$$m^* = \mu + [E(xx^T)^{-1}(Eq - \mu\, Ex)]^T(x - Ex) \tag{2.29}$$

and the standard deviation of m^* is equal to

$$\sigma(m^*) = [(Eq - \mu\, Ex)^T \Sigma^{-1}(Eq - \mu\, Ex)]^{\frac{1}{2}} \tag{2.30}$$

For the futures markets pricing relations (2.5), the corresponding unconditional implication can be expressed as

$$E\left[(f_{t+1} - f_t)m_{t+1}w_t\right] = 0, \tag{2.31}$$

where w_t is a k by 1 vector of instrumental variables that are measurable with respect to the information set at date t. Then the minimum variance stochastic discount factor can be obtained by solving the following problem analogous to (2.28) with Eq being set to zero vector, and x at date t being equal to $(f_{t+1} - f_t)w_t$:

$$\begin{aligned} \min_{m \in L^2} \quad & E(m_{t+1}^2) \\ \text{subject to} \quad & E\left[(f_{t+1} - f_t) m_{t+1} w_t\right] = 0 \\ & E(m_{t+1}) = \mu, \end{aligned} \tag{2.32}$$

Solving the above minimization problem, we obtain the minimum variance SDF as

$$\begin{aligned} m_{t+1}^* &= \mu - \mu E\left[(f_{t+1} - f_t) w_t'\right] \left\{\text{Var}\left[(f_{t+1} - f_t) w_t\right]\right\}^{-1} \\ &\quad \times \left\{(f_{t+1} - f_t) w_t - E\left[(f_{t+1} - f_t) w_t\right]\right\}. \end{aligned} \tag{2.33}$$

and the variance of the bound is

$$\text{Var}(m_{t+1}^*) = \mu^2 E\left[(f_{t+1} - f_t) w_t'\right] \left\{\text{Var}\left[(f_{t+1} - f_t) w_t\right]\right\}^{-1} E\left[(f_{t+1} - f_t) w_t\right]. \tag{2.34}$$

Thus to evaluate a pricing model, we compare the variance of the model implied SDF with that of the variance bound. If the variance of the model implied SDF is greater than or equal to the bounds, then the model is admissible. It should be noted here that an admissible model implied SDF does not necessarily suggest that the model is 'correct'. It only says that the model SDF lies within the admissible set.

2.3.2 The Frequency-domain Volatility Bound

Using the volatility bound described above, we can evaluate asset pricing models by looking at the model implied SDF variance. This approach can be applied to a wide class of asset pricing models that have a stochastic discount factor representation. In fact, it has become a standard method in financial economics to detect for model inadmissibility. However, there is still some important information that cannot be obtained by the above method. It does not provide information about the frequency at which the model fails. Such information is useful to us since some particular frequency may be extremely important to the commodity we are studying. For example, seasonal frequency plays an important role in explaining agricultural commodity prices. A model may be inadmissible simply because it fails to capture seasonal behavior of the prices. In such a case, to be able to identify the frequency on which the model fails will be important for the econometrician to detect the source of model misspecification and modify the model correspondingly. In addition, we may be only interested in price behavior on some specific frequencies. For example, the business cycle frequency is of particular interest to macroeconomists. Such situations motivate the frequency-domain volatility bounds developed in this paper.

Let γ_k and γ_k^* be the autocovariance functions for m_{t+1} and m_{t+1}^* respectively. Assume that the sequences $\{\gamma_k\}$ and $\{\gamma_k^*\}$ are absolutely summable, that is,

$$\sum_{k=-\infty}^{\infty} \mid \gamma_k \mid < \infty \tag{2.35}$$

and

$$\sum_{k=-\infty}^{\infty} \mid \gamma_k^* \mid < \infty \tag{2.36}$$

Then the spectral representation of γ_k and γ_k^* can be expressed as

$$\gamma_k = \int_{-\pi}^{\pi} s(\omega) e^{i\omega\, k} d\omega \tag{2.37}$$

$$\gamma_k^* = \int_{-\pi}^{\pi} s^*(\omega) e^{i\omega\, k} d\omega \tag{2.38}$$

where

$$s(\omega) = \frac{1}{2\pi}\left[\gamma_0 + 2\sum_{k=1}^{\infty} \gamma_k \cos\omega\, k\right], \tag{2.39}$$

$$s^*(\omega) = \frac{1}{2\pi}\left[\gamma_0^* + 2\sum_{k=1}^{\infty} \gamma_k^* \cos\omega\, k\right], \tag{2.40}$$

are the Fourier transform, or the spectrum of γ_k and γ_k^* respectively.

Take $k = 0$ in (2.37) and (2.38), we obtain a decomposition of the variance of the processes m_{t+1} and m_{t+1}^* as the following:

$$\begin{aligned} \mathrm{Var}(m_{t+1}) &= \int_{-\pi}^{\pi} s(\omega) d\omega, \qquad (2.41) \\ \mathrm{Var}(m_{t+1}^*) &= \int_{-\pi}^{\pi} s^*(\omega) d\omega, \qquad (2.42) \end{aligned}$$

Such decomposition implies that the term $s(\omega)d\omega$ and $s^*(\omega)d\omega$ can be interpreted as the contribution to the variance attributable to the component of the process with frequencies in the interval $(\omega, \omega + d\omega)$. A peak in the spectrum indicates an important contribution to the variance from the components at frequencies around the peak. This naturally leads to a frequency-domain version of the volatility bounds considered previously. Since the variance of m_{t+1}^* is the lower bound on the variance of admissible SDFs, the spectral density $s^*(\omega)$ can be interpreted as the lower bound on the spectral density at frequency ω of admissible SDFs. To evaluate a particular model, we can compare the spectrum of the model implied SDF and the spectral volatility bounds. If the spectrum of the model implied SDF lies below the spectral bound at a particular frequency, then we say that the model is not admissible in that frequency. In particular, if a model implied

SDF has variance smaller than the lower bound, and the spectrum of the model SDF exceeds the spectral bound at all frequencies except around, say, seasonal frequency, then we can conclude that failure to capture seasonality in the data is the source of model misspecification. To make the model admissible, the model needs to be augmented with seasonal considerations. Thus, the spectral bound described above also provides a way to identify the source of model misspecification.

2.4 Empirical Applications

In this section, we apply the econometric method proposed in the previous section to evaluate the four versions of consumption-based CAPMs of commodity futures prices discussed in Section 2. For comparison purposes, we also consider a linear stochastic discount factor model which specifies the SDF as a linear function of macroeconomic variables.

We use monthly data on crude oil and corn futures prices. The data covers from January 1984 through December 1998. Details of data source are listed in Table 2.1.

Real per capita consumption series are obtained by $c_t = (C_{1t} + C_{2t})/N_t$,

and real return series are obtained by $r_t = (1 + R_t)P_{t-1}/P_t$.

Our procedure will be as follows:

First, we estimate the four consumption-based CAPMs by generalized method of moments (GMM) of Hansen (1982).

Consider the pricing relations (2.4) and (2.5). Each equation implies an error term u_{t+1} that has conditional mean of zero, given information at time t. For example, (2.4) implies that

$$E\left\{[(\frac{p_{t+1} + z_{t+1}}{p_t} m_{t+1} - 1] \mid w_t\right\} = 0 \tag{2.43}$$

and (2.5) implies

$$E\left\{[(f_{t+1} - f_t)m_{t+1}] \mid w_t\right\} = 0 \tag{2.44}$$

where w_t is a set of instruments known at time t.

Thus from $E(u_{t+1} \mid w_t) = 0$, we obtain orthogonality conditions between u_{t+1} and w_t written as

$$E(u_{t+1}w_t) = 0 \tag{2.45}$$

Therefore, the parameters of the models can be estimated by exploiting these orthogonality conditions using GMM, which is based on minimizing a

quadratic form ($g \top W\, g$), where $g = vec(\frac{\sum_t^T u_{t+1} w_t}{T})$ is a vector of sample orthogonality conditions. The weighting matrix is the inverse of the covariance matrix of the orthogonality conditions. The four models differ only in their expressions for u_{t+1}. For example, for model 1, we estimate γ in

$$E\left[w_t(f_{t+1} - f_t)\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\right] = 0, \tag{2.46}$$

$$E\left[w_t r_{t+1}\beta\left(\frac{c_{t+1}}{c_t}\right)^{-\gamma}\right] = 1 \tag{2.47}$$

where β is set to 0.99 following the majority of literature. [1] Similarly the parameters in the other three models are estimated. [2]

After obtaining parameter estimates, we compute the sample mean of the model implied SDF, $\hat{\mu}$, by

[1] Most of the literature set β at specific values rather than empirically estimating it from the data. This is because estimating such parameter is likelly to yield negative discount rate, i.e., β greater than 1. For example, in Weil (1989), β is set to be equal to 0.95 and 0.98. An example of estimate of such parameter is in Ferson and Consitantinides (1991) which estimated β to be equal to 0.993 for monthly data.

[2] For Model 4, we used a slight modification of the specification used by Ferson and Harvey (1992).

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} \beta \left(\frac{c_{t+1}}{c_t} \right)^{-\hat{\gamma}}. \tag{2.48}$$

Second, at each of the CAPM model implied SDF mean $\hat{\mu}$, we estimate the linear SDF model by minimizing the squared distance from the Hansen-Jagannathan variance bound. That is, we estimate β_0, β_1 and β_2 by minimizing

$$\min_{m_{t+1} \in \mathcal{M}} E[(\beta_0 + \beta_1 i_{t+1} + \beta_2 ex_{t+1} - m_{t+1})^2], \tag{2.49}$$

where $\mathcal{M}$ is the set of nonnegative SDFs that satisfy $E(m_t) = \hat{\mu}$, which is obtained from the estimated CAPM models, and $E[w_t(f_{t+1} - f_t)m_{t+1}] = 0$, i_t is the interest rate and ex_t is the broad index of real exchange rate returns. The factors are chosen from a larger sets of macro economic variables using BIC model selection criteria.

The squared distance above is the measure of model misspecification proposed by Hansen, Heaton and Luttmer (1995) and Hansen and Jagannathan (1997).

Third, for each of the CAPM model implied SDF mean $\hat{\mu}$, we also compute the Hansen and Jagannathan variance bound and its frequency-domain analog.

The time domain variance bound is obtained by equation (2.33) and (2.34). The frequency-domain variance bound is estimated nonparametrically using the Bartlett kernel with bandwidth set to 6.

We conduct the above procedure for each maturity month. Throughout the analysis we use the following instruments as w_t: constant, c_t/c_{t-1}, r_t and the current futures price difference $f_t - f_{t-1}$.

Parameter estimates are reported in Tables 2.2, 2.3, 2.4 and 2.5. Tables 2.2 and 2.4 present GMM estimation results for the four consumption-based CAPMs for crude oil and corn, respectively.

The second, third and fourth columns of Tables 2.2 and 2.4 contain parameter estimates of the four consumption-based CAPMs for each contract. The estimates for γ are all equal to 1 for model 2. The parameter estimates show significant differences for contracts maturing in different months. This is true for both crude oil and corn. This is consistent with significant seasonality observed in crude oil and corn futures prices. One reasonable explanation for seasonality in crude oil prices is the significant difference in demands for electricity, gasoline and heating oil in different seasons. The difference in the liquidity in corn futures market across maturity months may be able to explain seasonality in corn prices. Note that the estimated utility parameters

and coefficients of risk aversion may differ for futures contracts maturing in different months, they are constants that do not change with time. However, even based on 10% significance level, these seasonal differences are not statistically significant in most of the cases. Column five of Tables 2.2 and 2.4 report the test statistics for overidentifying restrictions (the J statistics). Our results show that the four consumption-based CAPM models pass the overidentifying restrictions tests in most cases at 10% significance level, with a few exceptions for crude oil contracts which pass the tests at 5% significance level. The last three columns of Tables 2.2 and 2.4 report the sample mean of the model implied SDF calculated from the parameter estimates, the model SDF variance, and the (time-domain) Hansen-Jagannathan variance bound at the model SDF sample mean. We see that the model implied SDF variances of all four CAPMs are much smaller than the corresponding Hansen-Jagannathan variance bounds, and hence their SDFs are inadmissible.

Tables 2.3 and 2.5 provide estimation results for the linear SDF models for crude oil and corn. The second, third, and fourth columns of these two tables present the parameter estimates for the linear SDF model. The parameter estimates are all negative for interest rate except for one case (June crude

oil contract). This means that the linear SDF decreases with interest rate. Since the linear SDF is estimated by setting its mean to the inverse of the return from the risk-free asset, thus the higher the interest rate, the smaller the estimated linear SDF. For convenience, column five presents the mean of the linear SDF, which is in fact equal to the mean of the estimated CAPM SDF reported in Tables 2.2 and 2.4 previously. Column six reports the variance of the linear SDF. The variance of the Hansen-Jagannathan bound at each of the model means is also listed in the last column for comparison convenience. The results show that all the linear SDF variances exceed the Hansen-Jagannathan bounds, and thus the linear SDF model is admissible.

The frequency-domain SDF evaluation is illustrated through figures 1, 2, 3 and 4. Since the variances of the consumption-based CAPM SDF are much smaller than the corresponding linear SDF variance, we present their results in separate figures.

Figure 1 depicts frequency-domain variance decomposition for the linear SDF and the Hansen-Jagannathan bounds for crude oil futures contracts maturing in even months. The figures for odd months are similar and thus are omitted to save space. Figure 3 depicts the frequency-domain variance decomposition for corn. The linear SDF exhibits high variability in low

frequencies for both crude oil and corn. This may be best explained by the persistence in many macroeconomic variables, in this case, the interest rate and the exchange rate return. Figure 3 also shows that for corn contracts, the variances of the linear SDF in March and May are not as high as the variances in the other maturity months. This may be due to the seasonality in the liquidity in the corn futures market.

Figures 2 and 4 illustrate the frequency-domain SDFs of the CAPMs and the Hansen-Jagannathan variance bounds for crude oil and corn, respectively. As expected from the results in tables 2.2 and 2.4, the consumption-based CAPMs perform poorly.

Although none of these models are admissible, Figures 2 and 4 show that Model 2 performs better than the other three models at all frequencies for both crude oil and corn. This is consistent with the findings in the time domain. Models 1, 3, and 4 tend to show higher volatility in higher frequencies. The seasonal habit persistence model (Model 4) outperforms models 1 and 3 in all frequencies for crude oil. This evidence is stronger in higher frequencies than in lower frequencies. Even though corn contracts also show significant seasonal variability, the model that is designed to model seasonality better (Model 4) does not show any advantage in explaining futures

price behavior for corn. Figures 2 and 4 also plot the spectral density of the Hansen-Jagannathan variance bounds. Figure 2 shows that the spectral density bound for crude oil reaches its peak in the lowest frequency. This implies that the longer run frequencies, for example, the business cycle frequencies, play an important role in explaining crude oil futures price dynamics. This is not surprising considering the fact that oil consumption is highly correlated with macroeconomic variables. In addition, the spectral density bound also exhibits a second peak around frequency 1.75, which corresponds to a period of approximately 3-4 months, which is a quarterly seasonality. Note that this is not seasonality in the usually sense, considering the fact that we are using monthly data. This may suggest the importance of seasonal frequency behavior for crude oil. Figure 4 exhibits a V-shape spectral density bound for corn. This may be related to the seasonality in corn production. A comparison of the CAPMs and the Hansen-Jagannathan bound in the frequency-domain reveals that all four consumption-based CAPMs are inconsistent with the features shown in the spectral bounds.

2.5 Concluding Remarks

In this paper, we proposed a frequency-domain variance bounds approach for assessing misspecified models. The proposed method is very useful for detecting possible sources of model misspecification. For example, it could tell us whether the model fails because it ignores the seasonal frequency, or, the business cycle frequency. Using crude oil and corn futures prices, we evaluated four versions of the consumption-based CAPMs for commodity futures.

Our results show that all four consumption-based CAPMs fail to explain commodity futures price variability. The failure of CAPMs are particularly pronounced in low frequencies and seasonal frequencies. In contrast, a simple linear stochastic discount factor model performs better than the CAPMs, particularly in low frequencies. Though it also fails in seasonal frequencies.

Appendix 2A

Derivation of Representative Agent's Utility Optimization Problem (Time Separable Utility)

$$\begin{aligned} &\max \quad E[\textstyle\sum_{t=0}^{\infty} \beta^t u(c_t)] \\ &\text{s.t.} \;\; c_t + p_t x_{t+1} \leq (p_t + z_t)x_t + (f_t - f_{t-1})q_{t-1} \end{aligned} \tag{2.50}$$

Solving the dynamic optimization problem

$$V(q_{t-1}, x_t) = max\left\{u(c_t) + \beta\, E_t V(q_t, x_{t+1})\right\} \tag{2.51}$$

The first order conditions imply

$$\frac{\partial\, V_t}{\partial\, c_t} = u'(c_t) - \lambda_t = 0 \tag{2.52}$$

$$\frac{\partial\, V_t}{\partial\, x_{t+1}} = \beta\, E_t V'_{x_{t+1}}(q_t, x_{t+1}) - \lambda_t\, p_t = 0 \tag{2.53}$$

$$\frac{\partial\, V_t}{\partial\, q_t} = \beta\, E_t V'_{q_t}(q_t, x_{t+1}) = 0 \tag{2.54}$$

Envelope conditions:

$$\begin{aligned} \frac{\partial V_t}{\partial x_t} &= \lambda_t(p_t + z_t) \\ \frac{\partial V_t}{\partial q_{t-1}} &= \lambda_t(f_t - f_{t-1}) \end{aligned} \tag{2.55}$$

Shifting the above three equations by one period ahead, we have

$$\begin{aligned} \frac{\partial V_{t+1}}{\partial x_{t+1}} &= \lambda_{t+1}(p_{t+1} + z_{t+1}) \\ \frac{\partial V_{t+1}}{\partial q_t} &= \lambda_{t+1}(f_{t+1} - f_t) \end{aligned} \tag{2.56}$$

Substituting into the first order conditions, we obtain

$$\begin{aligned} \beta\, E_t[\lambda_{t+1}(p_{t+1} + z_{t+1})] - \lambda_t p_t = 0 \\ \beta\, E_t[\lambda_{t+1}(f_{t+1} - f_t)] = 0 \end{aligned} \tag{2.57}$$

hence

$$\beta\, E_t[u'(c_{t+1})(p_{t+1} + z_{t+1})] = u'(c_t)p_t \tag{2.58}$$

and

$$\beta\, E_t[u'(c_{t+1})(f_{t+1} - f_t)] = 0 \tag{2.59}$$

further simplifications easily reduce to

$$\beta\, E_t \left\{ \frac{u'(c_{t+1})}{u'(c_t)} \frac{p_{t+1} + z_{t+1}}{p_t} \right\} = 1 \tag{2.60}$$

and

$$\beta\, E_t \left\{ u'(c_{t+1})(f_{t+1} - f_t) \right\} = 0 \tag{2.61}$$

Appendix 2B

Derivation of Representative Agent's Utility Optimization Problem (Habit Persistence Utility)

$$\begin{aligned} &\max \quad E[\textstyle\sum_{t=0}^{\infty} \beta^t u(c_t - \delta\, c_{t-1})] \\ &\text{s.t.} \;\; c_t + p_t x_{t+1} \leq (p_t + z_t)x_t + (f_t - f_{t-1})q_{t-1} \end{aligned} \tag{2.62}$$

Solving the dynamic optimization problem

$$V(q_{t-1}, c_{t-1}, x_t) = max\,\{u(c_t - \delta\, c_{t-1}) + \beta\, E_t V(q_t, c_t, x_{t+1})\} \tag{2.63}$$

The first order conditions imply

$$\frac{\partial\, V_t}{\partial\, c_t} = u'(c_t - \delta\, c_{t-1}) + \beta\, E_t V'_{c_t}(q_t, c_t, x_{t+1}) - \lambda_t = 0 \tag{2.64}$$

$$\frac{\partial\, V_t}{\partial\, x_{t+1}} = \beta\, E_t V'_{x_{t+1}}(q_t, c_t, x_{t+1}) - \lambda_t\, p_t = 0 \tag{2.65}$$

$$\frac{\partial\, V_t}{\partial\, q_t} = \beta\, E_t V'_{q_t}(q_t, c_t, x_{t+1}) = 0 \tag{2.66}$$

Envelope conditions:

$$\begin{aligned}
\frac{\partial V_t}{\partial x_t} &= \lambda_t(p_t + z_t) \\
\frac{\partial V_t}{\partial q_{t-1}} &= \lambda_t(f_t - f_{t-1}) \\
\frac{\partial V_t}{\partial c_{t-1}} &= -\delta\, u'(c_t - \delta\, c_{t-1}) \qquad (2.67)
\end{aligned}$$

Shifting the above three equations by one period ahead, we have

$$\begin{aligned}
\frac{\partial V_{t+1}}{\partial x_{t+1}} &= \lambda_{t+1}(p_{t+1} + z_{t+1}) \\
\frac{\partial V_{t+1}}{\partial q_t} &= \lambda_{t+1}(f_{t+1} - f_t) \\
\frac{\partial V_{t+1}}{\partial c_t} &= -\delta\, u'(c_{t+1} - \delta\, c_t) \qquad (2.68)
\end{aligned}$$

Substituting into the first order conditions, we obtain

$$\begin{aligned}
u'(c_t - \delta\, c_{t-1}) + \beta\, E_t[-\delta\, u'(c_{t+1} - \delta\, c_t)] - \lambda_t = 0 \\
\beta\, E_t[\lambda_{t+1}(p_{t+1} + z_{t+1})] - \lambda_t p_t = 0 \\
\beta\, E_t[\lambda_{t+1}(f_{t+1} - f_t)] = 0 \qquad (2.69)
\end{aligned}$$

hence

$$\beta\, E_t\,\{[u'(c_{t+1}-\delta\, c_t)-\beta\, E_{t+1}(\delta\, u'(c_{t+2}-\delta\, c_{t+1})](p_{t+1}+z_{t+1})\}$$

$$=[u'(c_t-\delta\, c_{t-1})-\beta\, E_t(\delta\, u'(c_{t+1}-\delta\, c_t))]p_t \qquad (2.70)$$

and

$$\beta\, E_t\,\{[u'(c_{t+1}-\delta\, c_t)-\beta\, E_{t+1}(\delta\, u'(c_{t+2}-\delta\, c_{t+1}))](f_{t+1}-f_t)\}=0 \qquad (2.71)$$

further simplifications easily reduce to

$$\beta\, E_t\left\{\frac{u'(c_{t+1}-\delta\, c_t)-\beta\, E_{t+1}(\delta\, u'(c_{t+2}-\delta\, c_{t+1}))}{u'(c_t-\delta\, c_{t-1})-\beta\, E_t(\delta\, u'(c_{t+1}-\delta\, c_t))}\frac{p_{t+1}+z_{t+1}}{p_t}\right\}=1 \qquad (2.72)$$

and

$$\beta\, E_t\,\{[u'(c_{t+1}-\delta\, c_t)-\beta\, E_{t+1}(\delta\, u'(c_{t+2}-\delta\, c_{t+1}))](f_{t+1}-f_t)\}=0 \qquad (2.73)$$

Table 2.1. Monthly Data: January 1984 – December 1998

	Variables	Sources
C_{1t}	Real Personal Consumption Expenditures: Nondurable Goods Billions of Chained 1996 Dollars, Seasonally Adjusted Annual Rate	U.S. Department of Commerce Bureau of Economic Analysis
C_{2t}	Real Personal Consumption Expenditures: Service Goods Billions of Chained 1996 Dollars, Seasonally Adjusted Annual Rate	U.S. Department of Commerce Bureau of Economic Analysis
ex_t	Trade-Weighted Real Exchange Rate Index (Broad)	G.5 Release, Federal Reserve Board of Governors
f_t	Futures Prices on First Wednesday	Crude Oil: New York Board of Trade Corn: Chicago Board of Trade
i_t	3-Month Treasury Bill Rate, Secondary Market Averages of Business Days, Discount Basis Percent	H.15 Release, Federal Reserve Board of Governors
N_t	Total Population: All Ages Including Armed Forces Overseas Thousands	U.S. Department of Commerce Census Bureau
P_t	Consumer Price Index for All Urban Consumers: All Items 1982-84=100, Not Seasonally Adjusted	U.S. Department of Labor Bureau of Labor Statistics
R_t	Value-Weighted Returns	Center for Research in Security Prices

Notes: The Models 1, 2, 3 and 4 refer to (2.4), (2.5) with alternative specifications for m_{t+1}. β is set to 0.99. $*$ denotes the 10% level of statistical significance.

Table 2.2. Consumption-based CAPM Models: Crude Oil

Model	Month	γ	ρ	δ	J statistic	mean	variance	bound
1	JAN	0.745			9.532	0.989	0.000	0.038
	FEB	0.441			9.851	0.989	0.000	0.048
	MAR	0.414			8.460	0.989	0.000	0.045
	APR	0.681			9.120	0.989	0.000	0.042
	MAY	0.582			9.548	0.989	0.000	0.044
	JUN	0.612			9.585	0.989	0.000	0.054
	JUL	0.413			9.761	0.989	0.000	0.054
	AUG	1.177			11.514	0.988	0.000	0.059
	SEP	1.480			11.832	0.987	0.000	0.071
	OCT	0.001			7.481	0.990	0.000	0.049
	NOV	0.001			12.731*	0.990	0.000	0.063
	DEC	0.001			8.361	0.990	0.000	0.051
2	JAN	1.000	7.862		6.871	0.990	0.002	0.038
	FEB	1.000	7.867		8.264	0.991	0.002	0.049
	MAR	1.000	7.867		7.639	0.990	0.002	0.045
	APR	1.000	8.506		7.039	0.990	0.002	0.042
	MAY	1.000	7.878		7.354	0.991	0.002	0.044
	JUN	1.000	7.897		8.780	0.991	0.002	0.054
	JUL	1.000	8.332		8.018	0.992	0.002	0.054
	AUG	1.000	8.508		8.241	0.992	0.002	0.060
	SEP	1.000	7.930		9.913	0.992	0.002	0.072
	OCT	1.000	7.871		6.303	0.993	0.002	0.049
	NOV	1.000	7.229		7.468	0.994	0.002	0.064
	DEC	1.000	8.068		5.768	0.993	0.002	0.051
3	JAN	0.419		0.312	8.734	0.989	0.000	0.035
	FEB	0.012		0.810	9.945	0.990	0.000	0.050
	MAR	0.201		0.350	8.309	0.990	0.000	0.045
	APR	0.050		0.712	8.962	0.990	0.000	0.043
	MAY	0.016		0.809	9.145	0.990	0.000	0.044
	JUN	0.016		0.812	9.215	0.990	0.000	0.054
	JUL	0.004		0.904	8.447	0.990	0.000	0.054
	AUG	0.598		0.306	10.660*	0.989	0.000	0.060
	SEP	0.657		0.392	9.973	0.989	0.000	0.070
	OCT	0.001		0.957	7.020	0.990	0.000	0.050
	NOV	0.000		0.399	12.383*	0.990	0.000	0.064
	DEC	0.001		0.964	7.304	0.990	0.000	0.052
4	JAN	0.117		0.692	8.538	0.990	0.000	0.037
	FEB	0.064		0.796	10.208	0.990	0.000	0.059
	MAR	0.214		0.000	8.654	0.990	0.000	0.053
	APR	0.052		0.831	8.956	0.990	0.033	0.051
	MAY	0.243		0.739	8.817	0.990	0.000	0.052
	JUN	0.065		0.796	8.518	0.990	0.000	0.057
	JUL	0.008		0.959	8.403	0.990	0.000	0.065
	AUG	0.944		0.485	10.144	0.989	0.000	0.069
	SEP	0.928		0.623	9.829	0.989	0.001	0.080
	OCT	0.049		0.762	6.762	0.990	0.000	0.058
	NOV	0.021		0.000	12.663*	0.990	0.000	0.079
	DEC	0.121		0.616	6.948	0.990	0.000	0.059

Notes: The linear stochastic discount factor model takes the form of $m_{t+1} = \beta_0 + \beta_1 i_{1t+1} + \beta_2 ex_{2t+1}$, where i_{1t} is the 3 month treasury bill rate and ex_{2t} is the broad index of real exchange rate returns.

Table 2.3. Linear SDF Models: Crude Oil

Model	Month	β_0	β_1	β_2	mean	variance	bounds
1	JAN	2.328	-0.200	-0.002	0.989	0.127	0.038
	FEB	0.883	-0.303	0.020	0.989	0.192	0.048
	MAR	0.909	-0.312	0.020	0.989	0.214	0.045
	APR	1.822	-0.246	0.006	0.989	0.163	0.042
	MAY	1.365	-0.277	0.013	0.989	0.179	0.044
	JUN	0.914	-0.314	0.020	0.989	0.217	0.054
	JUL	1.419	-0.315	0.015	0.989	0.233	0.054
	AUG	1.612	-0.292	0.012	0.988	0.218	0.059
	SEP	2.760	-0.185	-0.007	0.987	0.162	0.071
	OCT	2.256	-0.229	0.001	0.990	0.199	0.049
	NOV	0.932	-0.320	0.021	0.990	0.288	0.063
	DEC	1.620	-0.283	0.011	0.990	0.258	0.051
2	JAN	2.349	-0.199	-0.002	0.990	0.127	0.038
	FEB	0.884	-0.303	0.020	0.991	0.193	0.049
	MAR	0.910	-0.312	0.020	0.990	0.215	0.045
	APR	1.587	-0.264	0.010	0.990	0.175	0.042
	MAY	1.367	-0.278	0.013	0.991	0.180	0.044
	JUN	0.916	0.314	0.020	0.991	0.218	0.054
	JUL	1.423	-0.316	0.015	0.992	0.234	0.054
	AUG	1.620	-0.293	0.012	0.992	0.220	0.060
	SEP	2.782	-0.185	-0.007	0.992	0.163	0.072
	OCT	2.333	-0.225	0.000	0.993	0.198	0.049
	NOV	0.936	-0.321	0.021	0.994	0.290	0.064
	DEC	1.663	-0.281	0.011	0.993	0.257	0.051
3	JAN	2.313	-0.201	-0.002	0.989	0.128	0.035
	FEB	0.884	-0.303	0.020	0.990	0.193	0.050
	MAR	0.910	-0.312	0.020	0.990	0.215	0.045
	APR	1.614	-0.262	0.010	0.990	0.173	0.043
	MAY	1.365	-0.278	0.013	0.990	0.179	0.044
	JUN	0.915	-0.314	0.020	0.990	0.217	0.054
	JUL	1.420	-0.316	0.015	0.990	0.233	0.054
	AUG	1.612	-0.293	0.012	0.989	0.219	0.060
	SEP	2.702	-0.190	-0.006	0.989	0.164	0.070
	OCT	2.306	-0.226	0.000	0.990	0.197	0.050
	NOV	0.932	-0.320	0.021	0.990	0.288	0.064
	DEC	1.684	-0.278	0.011	0.990	0.253	0.052
4	JAN	2.330	-0.200	-0.002	0.990	0.127	0.037
	FEB	0.884	-0.303	0.020	0.990	0.193	0.059
	MAR	0.910	-0.312	0.020	0.990	0.215	0.053
	APR	1.546	-0.267	0.011	0.990	0.177	0.051
	MAY	1.365	-0.278	0.013	0.990	0.179	0.052
	JUN	0.915	-0.314	0.020	0.990	0.217	0.057
	JUL	1.420	-0.316	0.015	0.990	0.233	0.065
	AUG	1.609	-0.293	0.012	0.989	0.219	0.069
	SEP	2.691	-0.191	-0.006	0.989	0.164	0.080
	OCT	2.370	-0.221	-0.000	0.990	0.196	0.058
	NOV	0.932	-0.320	0.021	0.990	0.288	0.079
	DEC	1.610	-0.283	0.012	0.990	0.258	0.059

Table 2.4. Consumption-based CAPM Models: Corn

Model	Month	γ	ρ	δ	J statistic	mean	variance	bound
1	MAR	0.726			2.581	0.989	0.000	0.010
	MAY	0.717			2.050	0.989	0.000	0.005
	JUL	1.017			2.832	0.988	0.000	0.014
	SEP	0.803			3.614	0.989	0.000	0.017
	DEC	0.802			3.526	0.989	0.000	0.015
2	MAR	1.000	0.531		1.846	0.990	0.002	0.010
	MAY	1.000	0.583		0.979	0.990	0.002	0.005
	JUL	1.000	0.887		8.348	0.990	0.002	0.014
	SEP	1.000	0.496		3.123	0.990	0.002	0.017
	DEC	1.000	0.573		2.826	0.990	0.002	0.015
3	MAR	0.338		0.292	3.090	0.989	0.000	0.013
	MAY	0.415		0.273	2.231	0.989	0.000	0.007
	JUL	0.675		0.154	3.003	0.989	0.000	0.015
	SEP	0.471		0.168	3.843	0.989	0.000	0.019
	DEC	0.502		0.204	3.606	0.989	0.000	0.016
4	MAR	0.295		0.249	2.010	0.990	0.000	0.011
	MAY	0.157		0.458	1.686	0.990	0.000	0.007
	JUL	0.564		0.025	3.845	0.989	0.000	0.023
	SEP	0.425		0.015	3.159	0.989	0.000	0.019
	DEC	0.294		0.021	2.241	0.990	0.000	0.012

Notes: Models 1, 2, 3 and 4 refer to (2.4), (2.5) with alternative specifications for m_{t+1}. β set to 0.99.

Table 2.5. Linear SDF Models: Corn

Model	Month	β_0	β_1	β_2	mean	variance	bounds
1	MAR	1.220	-0.183	0.009	0.989	0.078	0.010
	MAY	0.515	-0.078	0.010	0.989	0.013	0.004
	JUL	0.059	-0.296	0.028	0.988	0.182	0.014
	SEP	0.965	-0.221	0.014	0.989	0.107	0.017
	DEC	1.521	-0.260	0.010	0.989	0.164	0.015
2	MAR	1.288	-0.230	0.011	0.990	0.123	0.010
	MAY	0.514	-0.078	0.010	0.991	0.013	0.005
	JUL	0.148	-0.298	0.027	0.990	0.184	0.014
	SEP	1.447	-0.266	0.011	0.991	0.169	0.017
	DEC	1.519	-0.261	0.010	0.990	0.165	0.015
3	MAR	0.895	-0.162	0.011	0.989	0.057	0.013
	MAY	0.514	-0.078	0.010	0.989	0.013	0.007
	JUL	0.068	-0.296	0.028	0.989	0.182	0.015
	SEP	1.866	-0.236	0.005	0.989	0.149	0.019
	DEC	1.558	-0.258	0.010	0.989	0.162	0.016
4	MAR	1.426	-0.191	0.007	0.990	0.089	0.011
	MAY	0.515	-0.078	0.010	0.990	0.013	0.007
	JUL	-0.181	-0.294	0.030	0.989	0.180	0.023
	SEP	1.666	-0.250	0.008	0.990	0.157	0.019
	DEC	1.359	-0.272	0.013	0.990	0.173	0.012

Notes: The linear stochastic discount factor model takes the form of $m_t = \beta_0 + \beta_1 i_{1t+1} + \beta_2 ex_{2t+1}$, where i_{1t} is the 3 month treasury bill rate and ex_{2t} is the broad index of real exchange rate returns.

Figure 1. Frequency−Domain SDF Variance Decomposition: Crude Oil

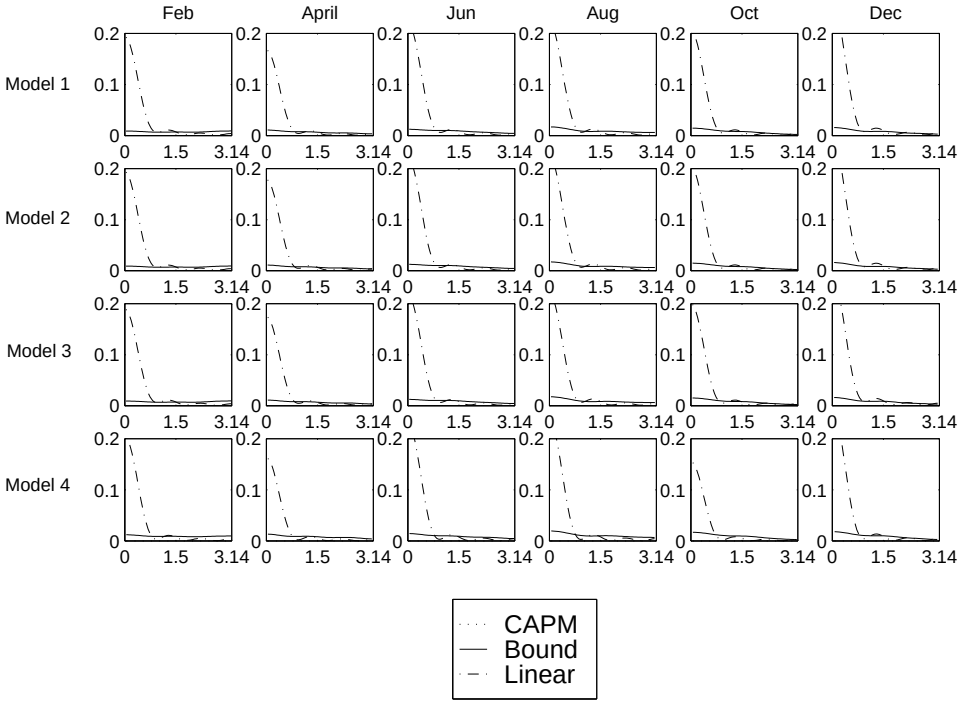


Figure 2. Fequency−Domain SDF Variance Decomposition: Crude Oil
Models 1, 2, 3, 4 and Variance Bound (September Contract)

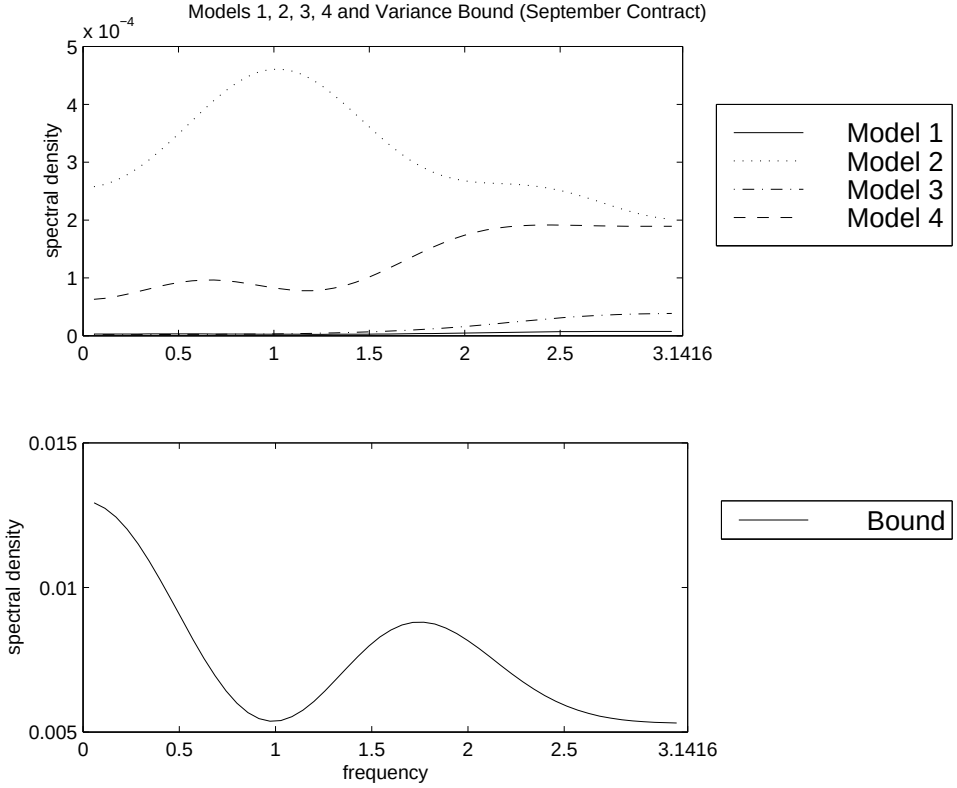


Figure 3. Frequency−Domain SDF Variance Decomposition: Corn

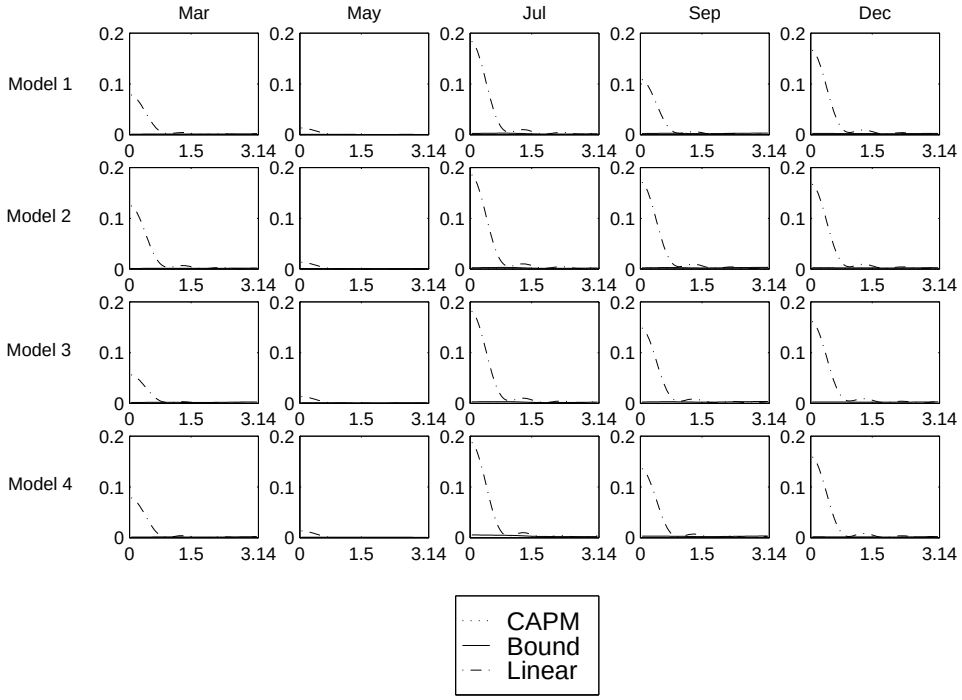
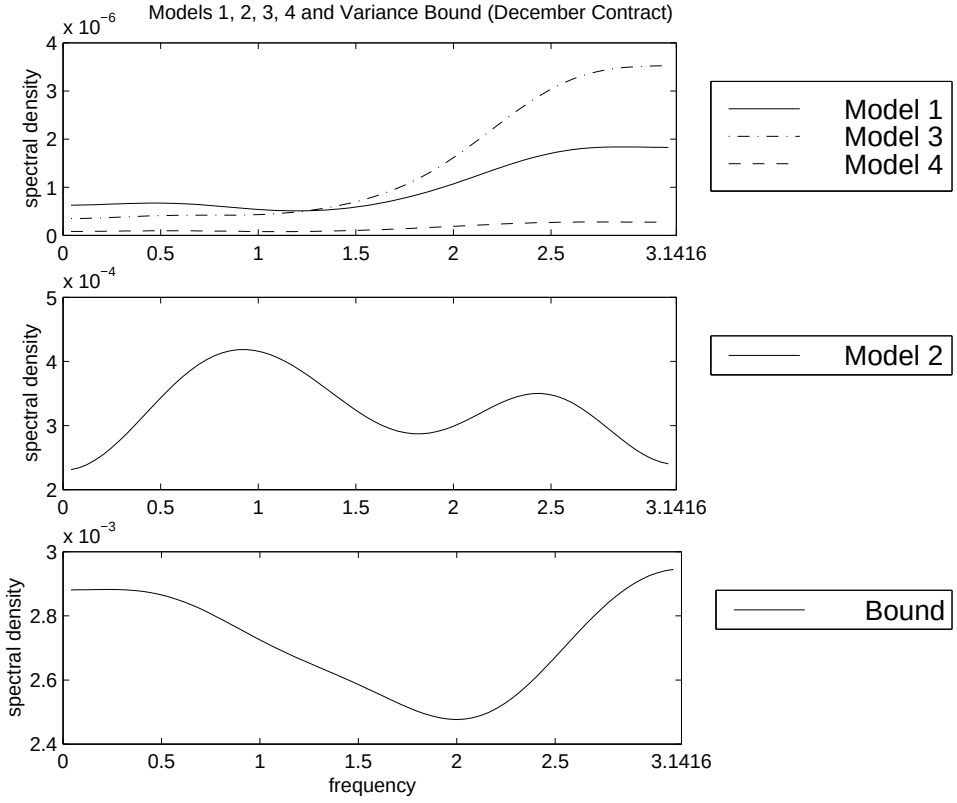


Figure 4. Frequency−Domain SDF Variance Decomposition: Corn
Models 1, 2, 3, 4 and Variance Bound (December Contract)



Chapter 3

Determining the Number of Factors in an Affine Term Structure Model of Futures Prices

3.1 Introduction

Factor models with an affine term structure have been widely used in the finance literature. The affine term structure models imply that both the drift

and variance functions of the state variables or factors are affine functions of the state variables themselves. Thus this class of models are quite tractable computationally. Such models were discussed in detail in Duffie and Kan (1996). In fact many of the well-known models in the literature are nested as special cases of the affine class, e.g., Vasicek (1977); Cox, Intersoll, and Ross (1985, hereafter CIR); Longstaff and Schwartz (1992) ; Dai and Singleton (1998).

One issue that is important to both theoretical and empirical validity of the affine factor models is the correct specification of the number of factors. Many of previous studies of affine factor models have assumed the number of factors rather than empirically testing it from the data. For example, Ghysels and Ng (1998) assumed two factors and tested the affine structure; Longstaff and Schwartz (1992) studied a two-factor model of interest rate.

Some statistical methodologies have previously been examined concerning the estimation of factor models and the determination of the number of factors. Basically these methods can be classified into three categories: the maximum likelihood estimation method, the principal component analysis, and the minimum distance method.

3.1.1 The Maximum Likelihood Estimation of Factor Models

The maximum likelihood estimation of factor models derived by Lawley (1940) is quite widely used in factor analysis. For a standard factor model

$$y = \lambda_0 + \Lambda\, x + e \tag{3.1}$$

where y is a p by 1 vector of random variables, λ_0 is a p by 1 vector of constants, x is a $k < p$ by 1 vector of factors, e is a p by 1 vector of random errors, and Λ a p by k matrix of factor loadings. It is assumed that x and e are independent with $E(x) = 0$, $E(e) = 0$, $E(xx^\top) = I$, and $E(ee') = \Psi$, a diagonal matrix. The covariance matrix for y is then

$$\Sigma = \Lambda\Lambda' + \Psi \tag{3.2}$$

Assume that the random vector y has a multivariate normal distribution, and we draw a sample of size N from the population. Let S be the sample covariance matrix of y, which is written as

$$S = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})(y_i - \bar{y})^\top \tag{3.3}$$

Then S has the Wishart distribution. The log likelihood function is expressed as

$$\ln L = -\frac{1}{2}(N-1)[\ln \mid \Sigma \mid + tr(S\Sigma^{-1})] \tag{3.4}$$

The values of Λ and Ψ that maximize $\ln L$ are the maximum likelihood estimates of the true Λ and Ψ, which are denoted as $\hat{\Lambda}$ and $\hat{\Psi}$.

Then a likelihood ratio test can be derived to test the hypothesis of the appropriate number of factors for given k. The null hypothesis is that $\Sigma = \Lambda\Lambda' + \Psi$ holds for a given value of k, and the alternative hypothesis is that Σ is any positive definite matrix of order p by p.

The likelihood ratio statistic

$$U_k = (N-1)[\ln \mid \hat{\Sigma} \mid - \ln \mid S \mid + tr(S\hat{\Sigma}^{-1}) - p] \tag{3.5}$$

is χ^2 with degrees of freedom equal to

$$d_k = \frac{1}{2}[(p-k)^2 - (p+k)] \tag{3.6}$$

where $\hat{\Sigma} = \hat{\Lambda}\hat{\Lambda}^\top + \hat{\Psi}$, and $\hat{\Lambda}$ and $\hat{\Psi}$ are the maximum likelihood estimates of the true Λ and Ψ.

If the null hypothesis is rejected, the conclusion is that at least $k + 1$ common factors are required.

In exploratory analysis where the problem is to determine the appropriate number of factors, a step by step procedure is usually needed. One can take a small value of k and test the hypothesis. If U_k is significant then increase k by 1 until a value of k is reached for which U_k is nonsignificant. This is a specific to general approach.

Though there has been some debate about whether the maximum likelihood test indicates the correct number of factors, and some studies have shown different evidence of its performance, the maximum likelihood test is still more widely used than other statistical tests in factor analysis. It would appear to give an upper bound to the number of factors that could legitimately be extracted.

3.1.2 The Principal Component Analysis

The principal component analysis is another widely used multivariate technique. Similar to the factor analysis, the principal component analysis is also based on the analysis of the covariance matrix.

Consider the random vector y described as in equation (3.1), with co-

variance matrix $cov(y) = \Sigma$. Different from the factor analysis, the principal component analysis does not assume specific structures on Σ, the covariance matrix of y, as expressed in equation (3.2) for the factor analysis model.

The principal components of y are linear combinations $\tilde{\epsilon_i} = \tilde{\beta_i} y$, $i = 1, \cdots, p$, of the elements of y with $\tilde{\beta}'\tilde{\beta} = 1$ such that each $\tilde{\epsilon_i}$ has maximal variance λ_i subject to the condition of being uncorrelated with $\tilde{\epsilon_1}, \cdots, \tilde{\epsilon_{i-1}}$. The weight vectors $\tilde{\beta_i}$ and the variances λ_i are found as solutions of the eigenvalue problem

$$\Sigma\tilde{\beta_i} = \lambda_i\tilde{\beta_i} \tag{3.7}$$

The total variance of y is defined as

$$\sigma^2_{total} = trace(\Sigma) = \sum_{i=1}^{p} \lambda_i \tag{3.8}$$

If we order the eigenvalues in decreasing order, and choose the first k largest eigenvalues, then the first k principal components account for

$$r = \frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{p} \lambda_i} \tag{3.9}$$

of the variance. Hence the appropriate number of principal components in the random vector y can be determined by the percentage of variance accounted

for by the first k principal components.

The principal component analysis is closely related to the factor analysis despite their fundamental difference in their assumptions on the covariance matrix at the outset.

3.1.3 The Minimum Distance Method

The minimum distance method can also be used to estimate factor models and determine the number of factors. As illustrated in Craigg and Donald (1997), the problem can be considered as estimating a matrix rank, and the number of factors is determined by choosing the rank that minimizes a criteria function. The minimum distance method covers standard model selection criteria in time series analysis, e.g., Akaike Information Criteria (AIC), Schwarz Bayesian Information Criteria (BIC), etc.

Akaike's AIC is based on the use of the Kullback-Leibler information. The Kullback-Leibler information between the true model f and an approximating model g is defined as

$$I(f,g) = \int f(x) \ln \{\frac{f(x)}{g(x)}\} \, \mathrm{d}x \tag{3.10}$$

The quantity $I(f,g)$ measures the information loss when g is used to

approximate f. If $f = g$, then $I(f, g) = 0$. So we choose the model that loses as little information as possible or the one that is closest to the true model.

Akaike showed that a relative K-L information, which is $I(f, g)$ minus a constant that depends only on the unknown truth f, can be estimated from the empirical data and the maximized log-likelihood function. A suitable estimator of the relative K-L information is

$$maximized \;\; loglikelihood - K \tag{3.11}$$

where K is the number of parameters in the model. Akaike defines an information criterion (AIC) by multiplying both terms by 2, i.e.,

$$AIC = -2 * maximized \;\; loglikelihood + 2K \tag{3.12}$$

This is known as Akaike's information criterion (AIC). According to AIC, one should select the model that yields the smallest value of AIC, which is the closest to the unknown true model, among the candidate models. As more parameters are added to the model, the K-L information can be made smaller. However, at some point, the addition of still more estimated parameters will have the opposite effect and the estimate of the relative K-L information will increase. Hence the appropriate model will be chosen based on the Principle

of Parsimony.

Following Akaike's AIC, alternative criterion has been provided for model selection. One important criterion is the criterion of Schwartz (1978) developed in the Bayesian context, which is termed BIC for Bayesian Information Criterion in literature. The BIC is expressed as

$$BIC = -2 * maximized \;\; loglikelihood + K * log(T) \tag{3.13}$$

where T is sample size.

As pointed out by Schwartz (1978), since the BIC differs from AIC only in that the number of parameters is multiplied by $\frac{1}{2} \ln T$, it tends more than AIC to choose lower-dimensional models.

Some other suggested criterion functions are also covered by the minimum distance method, though they may differ in the details of the functional form.

3.1.4 Other Related Literature

Application of Factor Analysis

The application of factor analysis is mainly in cross-sectional studies in psychological and sociological sciences. Some economic studies also apply the factor analysis to determine the number of factors, for example, Litterman

and Sheinkman (1991) studied bond returns using factor analysis. They extracted three factors and found that the three factors can account for as much as 97% of the variance, and hence they conclude three factors are sufficient to explain bond returns.

Some other application of factor analysis also exists in the literature. In fact, ever since the introduction of the Arbitrage Pricing Theory (APT) of Ross (1976), factor analysis has become commonly used in finance to determine the number of factors generating asset returns. In his original development of the APT, Ross (1976) assumes a strict factor structure in which the idiosyncratic components have zero correlations across assets. Such assumption makes possible direct application of standard factor analysis.

Chamberlain and Rothschild (1983) relax the strict factor structure and assume an approximate factor structure in which residual returns may be weakly correlated across securities. The implication is that, for an approximate k-factor structure, exactly k of the eigenvalues of the asset return covariance matrix grow without limit as the number of assets increase, and all remaining eigenvalues remain constant. Hence, Chamberlain and Rothschild (1983)'s work leads naturally to some empirical efforts to determine the number of factors based on the behavior of the eigenvalues of the sample

covariance matrix, examples of this type of analysis include Trzcinka (1986) and Luedecke (1984). Using return data on 865 securities for 1069 eeks, Trzcinka (1986) computes sample covariance matrix of successive higher dimension, and examine the behavior of the eigenvalues. He finds that the first eigenvalue contributes a large fraction of the variance of returns, and its relative importance increases dramatically with the number of securities. The conclusion is that, while there may be many factors responsible for security returns, one factor is more important than the others. Luedecke (1984) also obtains similar result. The problems with such analysis are discussed in Brown (1989), who shows that, evidence that one market factor explains the major part of security returns is consistent with an economy where there are actually k "equally important" priced factors, and that the eigenvalue analysis in the context of such an economy will lead to a false inference that the one important factor is the return on an equally weighted market index.

Other Large Sample Tests to Determine the Number of Factors

Alternative to the factor analysis and principal component analysis, a number of more recent studies have also considered tests to determine the number of factors in different contexts. For example, Connor and Korajcyk (1993),

Stock and Watson (1998), Bai and Ng (1999). Connor and Korajcyk (1993) develop a test to determine the number of factors in an approximate factor model of asset returns. Their test is an asymptotic test that depends on large number of cross-sectional assets, N, and fixed number of observations in time, T. Stock and Watson (1998) develop a test for the number of factors in an approximate dynamic factor model. The test is developed for the purpose of forecasting a single time series when there are many predictors (N), and time series observations (T), and that the predictors can be summarized by a small number of indexes (factors). Their test is developed for the case when N, $T \to \infty$, and $\frac{\sqrt{N}}{T} \to \infty$. Bai and Ng (1999) set up the determination of number of factors as a model selection problem. Their test is valid for large cross-sections (N) and large time dimensions (T). None of these tests are relevant for the futures price context since the futures market has only a small number of contracts on any trading day.

Most of these extant studies regarding the number of factors apply to markets with large number of cross-sectional assets. The study in the futures market is quite limited. One example is in Schwartz (1997), who studies one-factor, two-factor, and three-factor affine models of commodity futures prices and evaluates their relative performance using out-of-sample cross-section

tests and time-series tests. Another example is in Cortazar and Schwartz (1994), who conducted principal component analysis on copper futures returns. The difficulty with the futures markets is that the number of futures contracts is usually quite small, about ten to fifteen contracts on each day. Moreover, these contracts only exist for short period of time before they mature. Thus more stringent assumptions are needed to obtain appropriate series of futures data for the analysis. In Cortazar and Schwartz (1994), they roll over nearby contracts to obtain the corresponding data matrix in the principal component analysis.

In this paper, we propose new methods to estimate affine term structure models of futures prices, and develop testing methodologies to determine the appropriate number of factors in the data. Before we describe the proposed approach in detail, we give a brief discussion of the method and the relations between the method and other related literature in the next subsection.

3.1.5 Some Discussions

While there exists several well established methods to estimate factor models and determine the appropriate number of factors, there are important differences.

First, in the standard factor models, the p by 1 vector y can be written as

$$y = \lambda_0 + \Lambda\, x + e$$

where Λ, the factor loadings, are constants to be estimated. While in the futures term structure models, for a specific futures contract with time to maturity τ, the futures price on such contract $y(\tau)$ can be written as

$$y(\tau) = \lambda_0(\tau) + \Lambda(\tau)\, x + e$$

where $y(\tau)$ is a scalar function of the time to maturity, and the factor loadings $\Lambda(\tau)$ are functions of time to maturity. The variables y are observed at different values of time to maturity. This makes the model much more complicated than the standard model setting. We propose to use polynomials to approximate the factor loading functions, then estimate the polynomial coefficients and hence obtain factor loadings.

Second, as already pointed out in the previous subsection, the number of futures contracts per day is not large, usually only about ten to fifteen contracts per day, hence the large sample tests discussed in the previous subsection are not applicable. Also, the futures contracts mature within a short period of time, hence, in order to apply the factor analysis or principal

component analysis to futures data, previous researchers roll over to the nearby contracts to obtain the corresponding data matrix for factor analysis. In our proposed approach, we avoid rolling over nearby contracts and instead estimate factor loadings on a day by day basis.

Third, we estimate the model by solving a nonlinear least squares problem, and determine the number of factors by minimization of a model selection criteria. Unfortunately, the penalty term in the AIC and BIC as a candidate penalty in our model does not penalize enough on the number of factors. Simulation results show that if we use MSE plus the penalty term in either AIC or BIC as model selection criterion, the criterion will choose the maximum number of factors regardless of what the true data generating processes are. In order to find the appropriate penalty term, we perform a simulation experiment.

Assume a criteria function in the form:

$$CR_k = \ln \hat{\sigma^2} + \nu * k * log(T)/T \tag{3.14}$$

where $\hat{\sigma^2}$ is estimate of the variance of measurement error, ν is a constant to be determined, k is the number of factors, T is the sample size.

We choose the value of ν that could enable us to choose the right model

with the highest probability based on simulations in the experiment. The number of factors will be determined such that the model achieves the smallest value of the criterion function. This belongs to the minimum distance method discussed in the previous section.

The rest of the paper is organized as follows: Section 2 presents the affine term structure model of futures prices and describes estimation issues of the model. Section 3 develops methods to determine the number of factors. Section 4 applies the approach to crude oil and corn futures price data. Section 5 concludes.

3.2 The Affine Term Structure Model of Futures Prices

In this section, the affine class of term structure models is presented in the context of futures prices.

3.2.1 Assumptions

Consider a futures contract traded in period t with time to maturity $\tau_{t,l}$, where l is the index of individual contract. Let $f_{t,\tau_{t,l}}$, a scalar, denote the log

of futures price on such a contract. Then, $f_{t,\tau_{t,l}}$ can be expressed as:

$$f_{t,\tau_{t,k}} = \beta_0(\tau_{t,k}) + \beta_1(\tau_{t,k})X_t + e_{t,\tau_{t,k}} \tag{3.15}$$

where $\beta_0(\cdot)$ is a function of time to maturity, $\beta_1(\cdot)$ is a $1 \times k$ row vector, each element of $\beta_1(\cdot)$ is a scalar function of time to maturity, X_t is an $k \times 1$ column vector of state variables, or factors, and $e_{t,\tau_{t,l}}$ is the disturbance of the l^{th} contract.

In period t, there are p_t futures contracts traded in the market, each with different time to maturity, $\tau_{t,l}$, where $l = 1, 2, \cdots, p_t$. Let f_t be the p_t by 1 vector of log futures prices on these contracts. Then we can express the model in a matrix form as follows:

$$f_t = B_0(\tau_t) + B_1(\tau_t)X_t + e_t \tag{3.16}$$

where $\tau_t = (\tau_{t,1}, \tau_{t,2}, \cdots, \tau_{t,p_t})$ is a p_t by 1 vector,

$$f_t = \left\{ \begin{array}{c} f_{t,\tau_{t,1}} \\ f_{t,\tau_{t,2}} \\ . \\ f_{t,\tau_{t,p_t}} \end{array} \right\}_{p_t \times 1}, \tag{3.17}$$

$$B_0(\tau_t) = \begin{Bmatrix} \beta_0(\tau_{t,1}) \\ \beta_0(\tau_{t,2}) \\ . \\ \beta_0(\tau_{t,p_t}) \end{Bmatrix}_{p_t \times 1}, \tag{3.18}$$

$$B_1(\tau_t) = \begin{Bmatrix} \beta_1(\tau_{t,1}) \\ \beta_1(\tau_{t,2}) \\ . \\ \beta_1(\tau_{t,p_t}) \end{Bmatrix}_{p_t \times k}. \tag{3.19}$$

and

$$e_t = \begin{Bmatrix} e_{t,\tau_{t,1}} \\ e_{t,\tau_{t,2}} \\ . \\ e_{t,\tau_{t,p_t}} \end{Bmatrix}_{p_t \times 1}, \tag{3.20}$$

where e_t is a p_t by 1 vector of measurement errors that are assumed to follow p_t-variate normal distribution with $E(e_t) = 0$, $Var(e_t) = H_t = \sigma^2 \cdot I_{p_t}$, where I_{p_t} is an identity matrix of dimension p_t.

The assumptions of the model can be summarized as follows:

(a) The futures prices f_t are affine functions of the state variables x_t.

(b) The state variables x_t and the random errors e_t are uncorrelated with each other, i.e., $E(xe) = 0$.

(c) The state variables x_t are stationary.

3.2.2 Estimation

To estimate the model, we consider polynomial approximations of the unknown functions $\beta_0(\tau_{t,l})$ and $\beta_1(\tau_{t,l})$. Let $\beta_0(\tau_{t,l})$ and $\beta_1(\tau_{t,l})$ be approximated by $\phi(\tau_{t,l})c$ and $\phi(\tau_{t,l})C$ respectively, where $\phi(\cdot)$ is a 1 by $m+1$ vector of polynomials in $\tau_{t,l}$ from order 0 to order m, c is an $m+1$ by 1 vector of polynomial coefficients for $\beta_0(\tau_{t,l})$, C is $m+1$ by k matrix of polynomial coefficients for $\beta_1(\tau_{t,l})$.

For notational simplicity, define

$$\Phi_t = \Phi(\tau_t) = \left\{ \begin{array}{c} \phi(\tau_{t,1}) \\ \phi(\tau_{t,2}) \\ . \\ \phi(\tau_{t,p_t}) \end{array} \right\}_{p_t \times k} . \tag{3.21}$$

In order to obtain estimate of the state variable X_t, we consider estimate X_t as least squares solutions to the linear regression model:

$$f_t - \Phi_t\, c = (\Phi_t\, C)X_t + e_t \tag{3.22}$$

where $E(e_t) = 0$, $Var(e_t) = H_t = \sigma^2 \cdot I$, hence the unobservable state variable X_t is estimated by:

$$\hat{X}_t = [(\Phi_t\, C)^\top H_t^{-1}(\Phi_t\, C)]^{-1}(\Phi_t\, C)^\top H_t^{-1}(f_t - \Phi_t\, c) \tag{3.23}$$

Thus the log of futures prices can be expressed in the following equation:

$$f_t = \Phi_t\, c + \Phi_t\, C(C^\top \Phi_t^\top H_t^{-1}\Phi_t\, C)^{-1}C^\top \Phi_t^\top H_t^{-1}(f_t - \Phi_t\, c) + \hat{e_t} \tag{3.24}$$

Rearranging terms and after some algebra, we obtain the expression for $\hat{e_t}$ as:

$$\hat{e_t} = (I_{p_t} - J_t H_t^{-1})e_t \tag{3.25}$$

where $J_t = \Phi_t\, C(C^\top \Phi_t^\top H_t^{-1}\Phi_t\, C)^{-1}C^\top \Phi_t^\top$.

It then follows that $\hat{e_t}$ follows multivariate normal distribution with zero mean and covariance matrix Σ_t, which can be shown to be equal to $H_t - J_t$. Details are contained in Appendix $3A$.

The expression for Σ_t suggests that $\hat{e_t}$ is an p_t-variate normal random variables with rank $p_t - k < p_t$. In order to write the log likelihood function, we perform the following transformation to obtain a full rank $(p_t - k)$-variate normal vector.

Consider a spectral decomposition of the symmetric matrix Σ_t. Let $\Sigma_t = Z_t D_t Z_t^\top$, where Z_t is a p_t by $p_t - k$ matrix such that $Z_t^\top Z_t = I_{p_t - k}$, D is a diagonal matrix of dimension $p_t - k$ with diagonal elements equal to the eigenvalues of Σ. Let $R_t = D_t^{-\frac{1}{2}} Z_t^\top$. Then $\tilde{e_t} = R_t \hat{e_t}$ is a $(p_t - k)$-variate normal random vector with zero mean, and covariance matrix $I_{p_t - k}$. Proof is given in Appendix $3B$.

Assuming $\tilde{e_t}$ is independent across time, which means that random shocks to the futures contracts at different time periods are uncorrelated with each other, the joint likelihood function for $(\tilde{e_1}, \tilde{e_2}, \cdots, \tilde{e_T})$ can be written as

$$f(\tilde{e_t}) = \prod_{t=1}^{T} (2\pi)^{-\frac{p_t - k}{2}} exp\{-\frac{1}{2} \tilde{e_t}^T \tilde{e_t}\} \tag{3.26}$$

The log likelihood function is

$$\log\ L = -\frac{1}{2} \sum_{t=1}^{T} (p_t - k) \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} \tilde{e_t}^\top \tilde{e_t} \tag{3.27}$$

So maximizing the log likelihood function with respect to the polynomial coefficients c and C is equivalent to finding the c and C that minimize the

transformed sum of squared errors $\sum_{t=1}^{T} \tilde{e_t}^{\top} \tilde{e_t}$, which in fact, after some algebra, can be shown to be equal to $\sum_{t=1}^{T} \hat{e_t}^{\top} H_t^{-1} \hat{e_t}$, where H_t is the covariance matrix for the original error term e_t. The proof is provided in Appendix $3C$ in order to save space. So we can obtain estimates of the polynomial coefficients by solving the following minimization problem:

$$\min_{c,C} \sum_{t=1}^{T} \hat{e_t}^{\top} H_t^{-1} \hat{e_t} \tag{3.28}$$

3.2.3 Some Empirical Issues

In this section, we discuss several empirical issues in estimation.

Identification

The true parameters in the $m+1$ by $k+1$ matrix $\bar{C}$ where $\bar{C} = (c\ C)$ are not identified in our model, that is, there are multiple solutions that obtain the maximum value for the likelihood function. In order to obtain consistent estimates, we impose the identification restrictions on the parameters.

The identification condition is stated in the following proposition:

Proposition 3.3.1*:* For the model as described in equation (3.16) through (3.24), the likelihood function (**??**) achieves unique maximum if C is re-

stricted to be

$$C = \left\{ \begin{array}{c} I_k \\ \\ C^* \end{array} \right\} \tag{3.29}$$

where C^* is of dimension $(m + 1 - k)$ by n, and c is restricted to be

$$c = \left\{ \begin{array}{c} 0_k \\ \\ c^* \end{array} \right\} \tag{3.30}$$

where c^* is an $m + 1 - k$ by 1 vector, 0_k is an k by 1 vector of zeros.

Proof: See Appendix $3D$.

Heteroscedasticity

In empirical analysis, how to deal with heteroscedasticity is an important issue to consider. It is generally expected that the futures price behavior across different maturities are quite different, since at different time to maturities, the liquidity of the contracts are different.

We model liquidity implicitly as functions of time to maturity. As a simplification, we assume a diagonal structure of the covariance matrix for e_t, which implies that shocks to the cross-sectional futures contracts are mutually independent.

Thus the covariance matrix for e_t can be written as:

$$H_t = diag(\phi(\tau_t^1)\gamma, \phi(\tau_t^2)\gamma, \cdots, \phi(\tau_t^{p_t})\gamma) \tag{3.31}$$

where γ is $m+1$ by 1 vector of polynomial coefficients.

We can thus estimate parameters in c, C, and γ together by minimizing the least squares objective (3.28).

Seasonality

Commodity futures market generally exhibit significant seasonality. In order to account for seasonality, we augment the proposed method in the previous section with seasonal considerations.

Assume that the seasonal effects are deterministic, and that they do not have influence on how the state variables affect the futures prices. This assumption implies that seasonality only enters through c. Thus, we can add dummy variables that represent contracts maturing in different months. In this way, we will obtain different estimates for c, and hence for β_0. The fitted value at time t will therefore fluctuate around the fitted value of $\beta_1 X_t$ due to different estimate for β_0. The whole term structure shifts up and down due to seasonality.

3.3 Determine the Number of Factors

We first estimate the polynomial coefficients by solving the nonlinear least squares problem (3.28) and obtain the sum of squared residuals. To determine the appropriate number of factors, we consider a penalized MSE comparison method and a sequential SSE test approach.

3.3.1 Penalized MSE Approach

We first consider determining the number of factors through penalized MSE comparison across alternative models.

In order to determine the appropriate penalty term, we perform a simulation study and choose the criteria that could achieve the highest probability of choosing the right number of factors.

Consider a criteria of the form:

$$CR_k = \log \hat{\sigma^2} + \nu * k * \log T/T \tag{3.32}$$

where $\hat{\sigma^2} = \frac{SSE}{\sum_{t=1}^{T}(p_t - k)}$, where p_t is the number of contracts on date t, T is the number of trading dates, k is the number of factors, SSE is the total sum of squared errors obtained from estimation, ν is a constant to be determined.

Note first that this estimate for the mean squared error corrects for the degree of freedom by subtracting the number of factors in the denominator, since you lose one degree of freedom by adding one more factor.

A more complicated penalty function that involves both the number of factors and the number of contracts could have been used, but simulation shows that including the number of contracts in the penalty term does not have any effect on the choice probabilities, hence they are eliminated in the criteria function.

This criteria satisfies the two conditions for weak consistency in selecting models stated in Sin and White (1996):

(a)

$$\nu \cdot \ k \cdot \ln T = o_p(T) \tag{3.33}$$

and

(b)

$$P[\lim_{T \to \infty} \nu \cdot \ k \cdot \ln T = \infty] = 1 \tag{3.34}$$

where P is the probability operator.

These two conditions will be sufficient for the criteria to choose model 1

if it has lower K-L information criteria, and choose more parsimonious model 2 if the two models have the same average K-L information criteria.

In order to determine the value of the constant ν in the criteria, we perform a simulation study described as follows.

We first simulate futures data under different scenarios. We simulate futures data using $0, 1, 2, 3, 4, 5$ factors, and for each true number of factors, we examine the probability that each number of factors is chosen for given value of ν, and hence determine the value of ν that could achieve the highest overall probabilities of choosing the right number of factors. The criteria used for the choice of ν will be discussed later in the text.

We simulate the data according to the following scheme.

$$f_t = \Phi_t \; c + \Phi_t \; C \; X_t + e_t \tag{3.35}$$

where f_t is p_t by 1 vector of futures prices on date t, Φ_t is basis functions for date t, c is vector of polynomial coefficients restricted to be zeros, C is polynomial coefficient matrix,which is set to be:

$$C = \left\{ \begin{array}{c} I_k \\ \\ 0 \end{array} \right\} \tag{3.36}$$

X_t is state variable on date t, which follows a vector autoregressive process of order 1. e_t is random noise which follows standard normal distribution.

The state variable process is simulated according to the following:

$$X_{t+1} = \rho\ I_k \cdot X_t + \sigma \cdot u_{t+1} \tag{3.37}$$

where we consider two choices for ρ, namely, $\rho = 0.5$, and $\rho = 0.9$. The choice for σ is such that the model maintains a constant signal-noise ratio in the sense that $\theta = \sum_i^{p_t} var(u_{i,t}) / \sum_i^{p_t} var(f_{i,t})$ is a constant. This is achieved by choosing the σ that satisfies

$$\sum_i^{p_t} var(f_{i,t}) = \frac{1}{T}\ trace(\frac{\sigma^2}{1-\rho^2} \times \sum_{t=1}^{T}(\Phi_t\ C\ C^\top \Phi_t^\top)) + \sum_i^{p_t} var(u_{i,t}) \tag{3.38}$$

and

$$\sum_i^{p_t} var(f_{i,t}) = \sum_i^{p_t} var(u_{i,t})/\theta \tag{3.39}$$

We set two choices for θ, $\theta = 0.005^2$, and $\theta = 0.025^2$ in simulation, and change σ accordingly for different choice of d in order to maintain constant signal-noise ratio.

For the case of zero factors, we simulate data according to:

$$f_t = \Phi_t \cdot c + e_t \tag{3.40}$$

where c is a vector of zeros in simulation.

Simulation size is set to 1000. Sample size of 200 and 1000 are considered. In simulation, smaller sample size data is obtained as a subset of larger sample size data directly. The futures price data is simulated such that there are equal number of contracts on all trading days. To examine possible effects of the number of contracts on the performance of the proposed approach, we simulate three sets of data, one with 6 contracts per day, with every adjacent contracts have one month difference in maturity, another with 12 contracts per day with one month difference in maturity, and the last case with 6 contracts per day but with two months difference in maturity. The purpose of such comparison is to see if the proposed approach may work better for some futures term structure than for others. The shortest time to maturity in the simulated data will be set to be one month. That is, for 6 contracts per day with one month difference in maturity, we start from first day's contracts with $2, 3, 4, 5, 6, 7$ months maturity, on the next week's trading day, each contract will decrease its maturity by one week, continue until the shortest contract on a trading day becomes shorter than one month, then we

will add an additional contract and exclude the shortest contract. For the 12 contracts per day case, we will start from $2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13$ months maturity, and do the same as 6 contracts case. For the 6 contracts per day but with two months difference in maturity case, we start from $2, 4, 6, 8, 10, 12$ months maturity, and do the same as the other two sets of data.

The simulation is implemented in Matlab. The polynomial coefficients are estimated using a quasi-newton method. For the first replication, we use the true parameters as starting values if the estimated number of factors is the same as the true model, and we use the same structure as starting values but with different dimension if the estimated number of factors differs from the true model. In later replications, the resulting estimates from the first replication is used as starting values.

For a given value of ν in the criteria, and for each of the 24 sets of data, each with alternative parameter combinations, we do the following:

First, we estimate the polynomial coefficients $\hat{C}$ by minimizing sum of squared errors over all contracts across all the trading days.

Second, we calculate SSE at the estimated $\hat{C}$, and calculate the selection criteria correspondingly.

Third, we choose among the alternative models the one that gives the minimum of the criteria.

We repeat the process 1000 times, and calculate the six by six probability matrix that each number of factors is chosen for each of the true number of factors. The probability matrix is constructed such that the column represents the true number of factors, the row represents the probability that a specific number of factors is chosen.

We choose ν in the range of 0 to 400 such that the overall distance of the probability matrices from identity matrices is minimized, i.e.,

$$\min_{\nu} \sum_{i \in M} trace(P_i(\nu) - I)^{\top}(P_i(\nu) - I) \tag{3.41}$$

where P_i is the ith probability matrix obtained through simulation, I is the identity matrix, $M = \{1, 2, 3, \cdots, 24\}$, which includes all the cases in the simulation. This includes both sample size $T = 1000$ and $T = 200$.

The results are that for $\nu = 48, 49, 50$, the objective function achieves minimum value of 0.00004. Most of the probability matrices are identity, that is, at these values of ν, the criteria can select the right number of factors with 100% probability.

Based on findings in the simulation, we recommend choosing $\nu = 50$.

Some Further Discussions

The criteria chosen in the previous subsection performs quite well in the simulation study, and satisfies the conditions for weak consistency. Here, we explore the criteria a little bit further.

We explore alternative data generating processes, and examine if the criteria works well in situations other than the one considered in the previous subsection.

We first consider the cases when the factor loadings are smaller, and hence, the contribution of the factors are not significant.

We changed the specification of the C matrix in the simulation. Instead of specifying C to be

$$C = \left\{ \begin{array}{c} I_k \\ \\ 0 \end{array} \right\} \tag{3.42}$$

We tried the several cases when the C matrices are

$$C = \left\{ \begin{array}{c} \delta \cdot I_k \\ \\ 0 \end{array} \right\} \tag{3.43}$$

where $\delta = 0.5, 0.08, 0.05, 0.01$. The results show that the criteria works fine for $\delta >= 0.08$, but when $\delta = 0.05$, the criteria can identify up to 4 factors,

i.e., if the true number of factors is less than or equal to 4, the criteria can choose the right number, but if the true number of factors is 5, the criteria chooses 4 factors with 100% probability.

When $\delta = 0.01$, the criteria can pick up to 2 factors, that is, if the true number of factors is less than or equal to 2, the criteria can choose the right number of factors, but if the true number of factors is greater than 2, the criteria chooses 1 less the correct number of factors.

This pattern actually shows that the performance of the criteria depends on the relative contribution of factors versus the noise. When the factor loading is too small, the contribution of the factors to the variance is small, and it is harder for the method to pick the right number of factors.

We also consider the several cases when the factor loadings for the factors are not equal, hence, the contribution of each individual factor is different.

We tried the following cases, in which $UpperC$ denotes the upper part of the C matrix.

(a)

$$diag(UpperC) = \{1, 0.5, 0.2, 0.1, 0.05\} \tag{3.44}$$

In this case, the criteria can identify the true number of factors with 100%

probability.

(b)

$$diag(UpperC) = \{1, 0.5, 0.2, 0.1, 0.01\} \tag{3.45}$$

The criteria works fine for up to 4 factors, but for 5 factors, it does not work as well.

(c)

$$diag(UpperC) = \{1, 0.5, 0.02, 0.01, 0.01\} \tag{3.46}$$

The criteria can identify the true number of factors with 100% probability for up to 3 factors, but for 4 and 5 factors, it picks 1 less than the true number of factors with 100% probability. This means that when the factor loadings on a factor is small, our criteria will not be able to identify such a factor, but will attribute to the random error.

There may also be a situation when a factor contributes quite significantly during some specific period of time, but has no impact during other times (the Peso problem). Such a situation may also cause problems using our criteria, since the effect of such factor only takes place in short period of time, the method may not be able to identify such a factor. However, if the

duration of the effect of a factor is very short, we might be able to consider such factors unimportant over the long term, and hence ignoring such a factor may not lead to serious problems.

Thus, based on our findings of the performance of the criteria in various data generating processes, the penalized in-sample MSE criteria can identify the right number of factors in general situations. However, when a factor contributes a small amount of variance, the criteria may under estimate the number of factors. It can be argued that a factor that has very low factor loadings is in fact ignorable itself, and hence it may not cause serious problems in choosing the appropriate model.

3.3.2 Sequential SSE Test Approach

In this section we consider a sequential SSE test approach. Along the same lines as the regression model,

$$\frac{SSE_k}{\hat{\sigma^2}} \sim \chi^2 \tag{3.47}$$

with $T(p-k) - Q_k$ degrees of freedom, where SSE_k is the sum of squared errors in the k factor model, $\hat{\sigma^2}$ is the estimated variance of the measurement error, p is the number of contracts per day, k is the number of factors, Q_k is

the number of parameters in the k factor model. Note that in situations when the number of contracts per day is unequal across trading days, $\sum_{t=1}^{T} p_t - T \cdot k$, where p_t represents the number of contracts on date t, should be used in place of $T(p-k)$.

Similarly for $k+1$ factor model,

$$\frac{SSE_{k+1}}{\hat{\sigma^2}} \sim \chi^2_{T(p-k-1)-Q_{k+1}} \tag{3.48}$$

Assume that additional factors are orthogonal to the other factors, then

$$\frac{SSE_k - SSE_{k+1}}{SSE_{k+1}/(T(p-k-1))} \sim \chi^2_{T+r_k} \tag{3.49}$$

where $SSE_{k+1}/(T(p-k-1)$ is an estimate of the MSE in the $k+1$ factor model, r_k is the difference in the number of parameters in the k factor model and k+1 factor model. In our model, r_k is obtained as the order of polynomial coefficients as they represent the number of elements in an additional column of the polynomial coefficient matrix C.

So, for small $\frac{SSE_k - SSE_{k+1}}{SSE_{k+1}}$,

$$T(p-k-1) \cdot \ln\left(1 + \frac{SSE_k - SSE_{k+1}}{SSE_{k+1}}\right) \sim \chi^2_{T+r_k} \tag{3.50}$$

hence

$$T(p-k-1)\cdot(\ln SSE_k - \ln SSE_{k+1}) \sim \chi^2_{T+r_k} \tag{3.51}$$

which can be viewed as sum of $T + r_k$ independent $\chi^2(1)$ random variables. Using the CLT,

$$\sqrt{\frac{T+r_k}{2}}(\frac{1}{T+r_k}\cdot T(p-k-1)\cdot(\ln SSE_k - \ln SSE_{k+1})-1) \sim N(0,1) \tag{3.52}$$

hence,

$$\frac{1}{\sqrt{2(T+r_k)}}\cdot T(p-k-1)\cdot(\ln SSE_k - \ln SSE_{k+1}) - \frac{\sqrt{T+r_k}}{\sqrt{2}} \sim N(0,1) \tag{3.53}$$

So, we reject H_0 of k factors in favor of H_a of $k+1$ factors if

$$\ln SSE_k - \ln SSE_{k+1} > \frac{(\sqrt{2}z_\alpha + \sqrt{T+r_k})\sqrt{T+r_k}}{T(p-k-1)} \tag{3.54}$$

where z_α is the critical value for standard normal distribution at significance level α.

We choose the number of factors using two alternative approaches.

A.The General-to-Specific Approach

The general-to-specific approach starts from a general model and eliminates higher order models if accept the null hypothesis.

In our case, we start from 4 vs. 5 factor models comparison test, reject H_0 of 4 factors and accept H_a of 5 factors if equation (3.54) is true, if so, then we stop further testing of 3 vs. 4 and choose 5 factors. If the equation does not hold, then we continue testing 3 vs. 4 factors, and so on, until the null hypothesis is rejected. We use $\alpha = 0.01$ critical value which is equal to 2.33. This will make it harder to reject the null hypothesis than if we use size $\alpha = 0.05$ critical value, and hence, less likely to choose too many factors.

The probability matrices obtained using this general-to-specific sequential approach are shown in table 3.7. The results show that for sample size $T = 1000$, this method can identify the right number of factors with high probability. However, for sample size $T = 200$, the approach does not do as well. In addition, the method works better when the number of contracts per day is 6 than when the number of contracts per day is 12. Overall, there is a slight tendency to choose more than the true number of factors, particularly when the true number of factors is small.

B. The Specific-to-General Approach

The specific-to-general approach works opposite to the general-to-specific approach.

In our model, we start from 0 vs. 1 factor models comparison test. If equation (3.54) does not hold, then we accept the null hypothesis of zero factor, and stop further testing. Otherwise, if equation (3.54) holds, then reject H_0 of 0 factor and accept H_a of 1 factor model. Then we continue testing 1 factor vs. 2 factor models, and so on, until the null hypothesis is accepted.

We also use $\alpha = 0.01$ critical value which is equal to 2.33.

The probability matrices obtained using this specific-to-general sequential approach are shown in table 3.8. The results show that the specific-to-general approach works better than the general-to-specific approach, particularly for the 12 contract case. The diagonal elements of the probability matrices are close to 1 for both sample size $T = 1000$ and $T = 200$.

Overall, the specific-to-general approach can choose the right number of factors with very high probability.

3.4 Applications to Commodity Futures Data

We apply the proposed approaches to weekly futures price data for crude oil and corn. Crude oil futures price data covers the first week of 1985 through the week of February 23rd, 2000. Corn futures price data covers the first week of 1980 through the week of September 8th, 1999. We use Wednesday data for each week. If in a week no Wednesday data is available, we use Thursday data for that week.

For the crude oil data, there are 793 weeks of futures data. The maximum number of contracts on a day is 15, and the minimum number of contracts on a day is 8. The crude oil contracts used in the study have contract maturities as short as less than one month, and as long as 15 months. The average maturity for crude oil futures is 0.63 years, corresponding to approximately seven to eight months. The contract maturity months cover all twelve months. For corn futures data, there are 767 weeks of data. The maximum number of contracts on a day is 7, and the minimum number of contracts on a day is 5. The corn contracts used in the study also have maturities as short as less than one month and as long as 15 months. The average contract maturity for corn futures is 0.65 years, about seven to eight months. Corn contracts include March, May, July, September and December contracts. A description

of the crude oil and corn futures data is included in table 3.0.

3.4.1 In-sample MSE Comparison

We first apply the in-sample MSE comparison method to the data.

First, we estimate the polynomial coefficients using the sample data for each number of factors models, ranging from zero to five factors.

Second, we then calculate the sum of squared residuals for each model, and calculate mean squared error according to

$$MSE = \frac{SSE}{\sum_{t=1}^{T} p_t - T \cdot k} \tag{3.55}$$

where p_t is the number of contract on date t, T is the total number of trading days, k is the number of factors in the model.

Third, we then calculate the model selection criteria CR_k for each model at $\nu = 50$. The model with the smallest value of the criteria is chosen.

The estimated polynomial coefficients for the non seasonal models are reported in table 3.5 for crude oil and table 3.6 for corn.

We first follow the above procedure using polynomial order $n = 6$, we then repeat the same procedure with polynomial order $n = 12$. For crude oil, the minimum number of contracts per day is 8, hence we consider zero

to seven factor models.

The results for polynomial order of 6 are reported in part a of tables 3.1 through 3.6

Table 3.1 presents the in-sample root mean squared error(RMSE), and the penalized MSE (CR_k) for crude oil futures across 6 candidate models(0 factor models through 5 factor models) for both nonseasonal and seasonal case.

The first part of table 3.1 presents the nonseasonal model RMSE results. The in-sample RMSE is reduced significantly from 0.1695 to 0.0286 by adding one factor to a non-factor model, which is about 83% reduction. When adding a second factor, the in-sample RMSE reduced to 0.0069, about 76% reduction. Adding more factors will reduce in-sample RMSE even further. The 5 factor model can achieve in-sample RMSE as small as 0.0010. The second column reports the in-sample model selection criteria for the models at $\nu = 50$. This column shows that for crude oil, five factor model achieves the smallest of the criteria, which is -11.7929. Hence the data suggests five factor model for crude oil using non-seasonal model.

The second part of table 3.1 presents the seasonal model results for crude oil. The model selection criteria reported in the second column shows that for

seasonal models, five-factor model still achieves the minimum of the criteria, which is -11.5299.

Results for corn futures are presented in tables 3.2.

Table 3.2 presents in-sample RMSE, and in-sample CR_k criteria for corn futures. For non-seasonal models, in-sample RMSE is reduced significantly from 0.1599 to 0.0339 by adding one factor to the zero factor model, which is 79% reduction. The in-sample RMSE is reduced from 0.0339 to 0.0237 by adding a second factor, which is about 30% reduction. By adding the fifth factor, the in-sample RMSE is reduced from 0.0131 to 0.0091, which is 31% reduction. The in-sample criteria listed in the second column shows that five factor model achieves the minimum of the criteria, which is -7.2351.

The seasonal model for corn is listed in the second part of table 3.5. A comparison between the non-seasonal model and seasonal model RMSE suggests that incorporating seasonality into the model decreases RMSE. All values in the seasonal model are apparently smaller than their counterparts in the non-seasonal model. This suggests that seasonality may play a relatively more important role in explaining corn futures than crude oil futures. The RMSE is reduced from 0.1581 to 0.0299, which is about 81% reduction. By adding a second factor, the RMSE is reduced to 0.0165, which is 45%

reduction. The RMSE is reduced by 27% from 0.0165 to 0.0140 by adding a third factor. Comparison of the in-sample CR_k for the seasonal models suggests that four factors are sufficient for corn futures. The four factor model achieves the minimum criteria among all candidate models, which is equal to -7.8182. This is different from the result using non-seasonal models which requires five factors to explain corn futures.

The results for polynomial order of 12 are reported in part b of table 3.1 through 3.4. The results show that for crude oil, the non-seasonal model chooses the six factor model, while the seasonal model chooses the seven factor model. For corn futures data, both the non-seasonal model and seasonal model choose the five factor model.

3.4.2 Sequential SSE Test Results

We apply the two types of sequential SSE test approach described previously to the crude oil and corn futures data, again for polynomial order of 6 and 12.

First, we obtain sum of squared errors for zero to five factor models from estimation (zero to seven factor models for crude oil for polynomial order of 12).

Second, we calculate $\ln SSE_k - \ln SSE_{k+1}$, for $k = 0, 1, 2, 3, 4$, and compare the values with $\frac{(\sqrt{2}z_\alpha+\sqrt{T+r_k})\sqrt{T+r_k}}{\sum_{t=1}^{T} p_t - T(k+1)}$, where α is set to 0.01, and perform the sequential test according to the general-to-specific approach and the specific-to-general approach described in the previous section.

The SSEs for each number of factor models, the difference in logs of SSEs, and the corresponding critical values for comparison with the log differences, are reported in table 3.3 for crude oil, and table 3.4 for corn. Parts a are for the polynomial order of 6, parts b are for polynomial order of 12.

For polynomial order of 6, both the general-to-specific and the specific-to-general approaches choose five factors for crude oil for both seasonal and non-seasonal model, and five factors for corn for non-seasonal model, but four factors for the seasonal model. These results are the same as those obtained using the penalized MSE approach.

For polynomial order of 12, both the general-to-specific and the specific-to-general approaches choose seven factors for crude oil and five factors for corn for both seasonal and non-seasonal models.

3.5 Concluding Remarks

In this paper, we propose a new methodology to estimate an affine term structure model and new approaches to determining the number of factors in such models. The proposed method is well suited for affine term structure models. where the cross-sectional observations are not large enough for asymptotic tests developed previously. By solving a nonlinear least squares problem, we obtain polynomial approximations to the unknown functions of times to maturity. We then choose the number of factors based on penalized in-sample mean squared error comparison across alternative models. Simulation study shows that the proposed penalized in-sample MSE method can choose the right number of factors for most cases. Though in some cases, for example, when a factor's factor loadings are small, hence the factor does not contribute much to the variance, the approach may fail to identify such a factor. However, such cases are rarely of much concern in practice since the factors that are left out are those that do not contribute much to the variance. In addition to the penalized in-sample mean squared error method, we also considered a sequential SSE test approach.

Finally, we apply the proposed methods to crude oil and corn futures price data. Using the penalized in-sample MSE comparison approach, the

data suggests five factors for the crude oil futures. For corn futures, the non-seasonal model chooses five factors, while the seasonal model chooses four factors.

Table of Notations

x	state variable
e	measurement error
f	futures price
t	time index of trade day
τ	time to maturity of a futures contract
T	sample size
N	number of cross sectional observations
p_t	number of contracts on date t
k	number of factors
m	polynomial order
N_s	simulation size

Table 3.0. Crude Oil and Corn Futures Data

Crude Oil Data from 1/2/1985 to 2/23/2000: 793 weeks of observations		
	Price	Maturity
Mean	$19.1820	0.6332
Standard Error	3.3452	0.3321
Maximum	38.69	1.2493
Minimum	10.61	0.0603

Corn Data from 1/2/1985 to 9/8/1999: 767 weeks of observations		
	Price	Maturity
Mean	254.3161(cents)	0.6451
Standard Error	40.9127	0.3501
Maximum	487.50	1.2493
Minimum	144.25	0.0411

Table 3.1(a) In-Sample RMSE and CR_k Comparison: Crude Oil 1/2/85 to 2/23/00

Non-seasonal Models, polynomial order = 6		
	In-Sample RMSE	In-Sample CR_k
0 Factor	0.1695	-3.5493
One Factor	0.0286	-6.6871
Two Factor	0.0069	-9.1041
Three Factor	0.0024	-10.8217
Four Factor	0.0013	-11.5399
Five Factor	0.0010	-11.7929
Seasonal Models, polynomial order = 6		
	In-Sample MSE	In-Sample CR_k
0 Factor	0.1683	-3.5643
One Factor	0.0283	-6.7058
Two Factor	0.0066	-9.1859
Three Factor	0.0023	-10.8748
Four Factor	0.0019	-10.8980
Five Factor	0.0011	-11.5299

Table 3.1(b) In-Sample RMSE and CR_k Comparison: Crude Oil 1/2/85 to 2/23/00

Non-seasonal Models, polynomial order=12		
	In-Sample RMSE	In-Sample CR_k
0 Factor	0.1695	-3.5494
One Factor	0.0286	-6.6873
Two Factor	0.0068	-9.1269
Three Factor	0.0024	-10.8376
Four Factor	0.0013	-11.6085
Five Factor	0.0009	-11.9535
Six Factor	0.0007	-12.1061
Seven Factor	0.0005	-12.0975
Seasonal Models, polynomial order = 12		
	In-Sample RMSE	In-Sample CR_k
0 Factor	0.1679	-3.5684
One Factor	0.0282	-6.7142
Two Factor	0.0066	-9.2043
Three Factor	0.0023	-10.8697
Four Factor	0.0013	-11.6338
Five Factor	0.0011	-11.4790
Six Factor	0.0007	-11.9996
Seven Factor	0.0005	-12.1033

Table 3.2(a) In-Sample RMSE and CR_k Comparison: Corn 1/2/85 to 9/8/99

Non-seasonal Models, polynomial order = 6		
	In-Sample RMSE	In-Sample CR_k
0 Factor	0.1599	-3.6666
One Factor	0.0339	-6.3345
Two Factor	0.0237	-6.6174
Three Factor	0.0177	-6.7700
Four Factor	0.0131	-6.9458
Five Factor	0.0091	-7.2351
Seasonal Models, polynomial order = 6		
	In-Sample RMSE	In-Sample CR_k
0 Factor	0.1581	-3.6895
One Factor	0.0299	-6.5880
Two Factor	0.0165	-7.3395
Three Factor	0.0140	-7.2446
Four Factor	0.0084	-7.8182
Five Factor	0.0082	-7.4506

Table 3.2(b) In-Sample RMSE and CR_k Comparison: Corn 1/2/85 to 9/8/99

Non-seasonal Models, polynomial order = 12		
	In-Sample RMSE	In-Sample CR_k
0 Factor	0.1599	-3.6666
One Factor	0.0339	-6.3358
Two Factor	0.0238	-6.6114
Three Factor	0.0174	-6.8060
Four Factor	0.0096	-7.5582
Five Factor	0.0071	-7.7208
Seasonal Models, polynomial order = 12		
	In-Sample RMSE	In-Sample CR_k
0 Factor	0.1579	-3.6919
One Factor	0.0271	-6.6063
Two Factor	0.0135	-7.3749
Three Factor	0.0096	-7.3180
Four Factor	0.0053	-7.7718
Five Factor	0.0016	-9.0135

Table 3.3(a) Sequential SSE Test Results: Crude Oil 1/2/85 to 2/23/00

Non-seasonal Models, polynomial order = 6

	SSE	$\ln SSE_k - \ln SSE_{k+1}$	Critical Value
0 Factor	301.6607	3.6374	0.0920
One Factor	7.9401	2.9231	0.1002
Two Factor	0.4269	2.2318	0.1099
Three Factor	0.0458	1.2420	0.1218
Four Factor	0.0132	0.7876	0.1366
Five Factor	0.0060	N/A	N/A

Seasonal Models, polynomial order = 6

	SSE	$\ln SSE_k - \ln SSE_{k+1}$	Critical Value
0 Factor	297.1649	3.6410	0.0920
One Factor	7.7936	2.9864	0.1002
Two Factor	0.3933	2.2030	0.1099
Three Factor	0.0435	0.5469	0.1218
Four Factor	0.0251	1.1675	0.1366
Five Factor	0.0078	N/A	N/A

Table 3.3(b) Sequential SSE Test Results: Crude Oil 1/2/85 to 2/23/00

Non-seasonal Models, polynomial order = 12

	SSE	$\ln SSE_k - \ln SSE_{k+1}$	Critical Value
0 Factor	301.6284	3.6374	0.0926
One Factor	7.9389	2.9458	0.1009
Two Factor	0.4173	2.2249	0.1107
Three Factor	0.0451	1.2946	0.1227
Four Factor	0.0124	0.8806	0.1376
Five Factor	0.0051	0.7030	0.1566
Six Factor	0.0025	0.5611	0.1818
Seven Factor	0.0014	N/A	N/A

Seasonal Models, polynomial order = 12

	SSE	$\ln SSE_k - \ln SSE_{k+1}$	Critical Value
0 Factor	295.9531	3.6453	0.0926
One Factor	7.7280	2.9963	0.1009
Two Factor	0.3862	2.1795	0.1107
Three Factor	0.0437	1.2879	0.1227
Four Factor	0.0120	0.3807	0.1376
Five Factor	0.0082	1.0711	0.1566
Six Factor	0.0028	0.6733	0.1818
Seven Factor	0.0014	N/A	N/A

Table 3.4(a) Sequential SSE Test Results: Corn 1/2/85 to 9/8/99

Non-seasonal Models, polynomial order = 6			
	SSE	$\ln SSE_k - \ln SSE_{k+1}$	Critical Value
0 Factor	118.7204	3.2815	0.2230
One Factor	4.4607	0.9364	0.2780
Two Factor	1.7488	0.8688	0.3690
Three Factor	0.7336	1.0054	0.5486
Four Factor	0.2684	1.3892	1.0687
Five Factor	0.0691	N/A	N/A

Seasonal Models, polynomial order = 6			
	SSE	$\ln SSE_k - \ln SSE_{k+1}$	Critical Value
0 Factor	116.0253	3.5121	0.2230
One Factor	3.4617	1.4049	0.2780
Two Factor	0.8495	0.6214	0.3690
Three Factor	0.4563	1.4031	0.5486
Four Factor	0.1122	0.7323	1.0687
Five Factor	0.0539	N/A	N/A

Table 3.4(b) Sequential SSE Test Results: Corn 1/2/85 to 9/8/99

Non-seasonal Models, polynomial order = 12			
	SSE	$\ln SSE_k - \ln SSE_{k+1}$	Critical Value
0 Factor	118.7139	3.2827	0.2247
One Factor	4.4550	0.9291	0.2801
Two Factor	1.7594	0.9109	0.3717
Three Factor	0.7076	1.5817	0.5526
Four Factor	0.1455	1.2625	1.0766
Five Factor	0.0412	N/A	N/A

Seasonal Models, polynomial order = 12			
	SSE	$\ln SSE_k - \ln SSE_{k+1}$	Critical Value
0 Factor	115.7522	3.5280	0.2247
One Factor	3.3990	1.4220	0.2801
Two Factor	0.8199	0.6593	0.3717
Three Factor	0.4241	1.2834	0.5526
Four Factor	0.1175	2.3416	1.0766
Five Factor	0.0113	N/A	N/A

Table 3.5. Estimates of Polynomial Coefficients: Crude Oil 1/2/85 to 2/23/00

Zero Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
2.9462					
-0.0086					
0.0152					
-0.0013					
-0.0005					
-0.0036					

One Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1				
0.7368	-0.2578				
-0.0129	0.0075				
0.0276	-0.0097				
0.0269	-0.0089				
-0.0051	0.0016				

Two Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1	0			
0	0	1			
0.0931	-0.0314	-0.3007			
-0.0353	0.0119	0.0816			
0.0242	-0.0082	-0.0264			
-0.0063	0.0021	0.0064			

Three Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1	0	0		
0	0	1	0		
0	0	0	1		
0.0048	-0.0014	-0.1084	-0.6378		
0.0052	-0.0018	0.0524	0.2509		
0.0019	-0.0006	-0.0221	-0.0817		

Four Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1	0	0	0	
0	0	1	0	0	
0	0	0	1	0	
0	0	0	0	1	
0.0065	-0.0021	-0.0225	-0.1980	-0.6859	
-0.0027	0.0008	0.0168	0.1314	0.3073	

Five Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
0	0	0	0	1	0
0	0	0	0	0	1
0.0006	-0.0002	-0.0104	-0.0928	-0.3878	-0.9029

Table 3.6 Estimates of Polynomial Coefficients: Corn 1/2/85 to 9/8/99

Zero Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
5.5244					
0.0204					
-0.0018					
-0.0008					
-0.0010					
-0.00003					
One Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1				
1.4533	-0.2592				
-0.0897	0.0152				
-0.0005	0.0087				
0.0900	-0.0162				
-0.0024	0.0005				
Two Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1	0			
0	0	1			
0.3830	-0.0687	-0.3431			
-0.1292	0.0228	0.0682			
0.1200	-0.0213	-0.0354			
-0.1186	0.0211	0.0892			
Three Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1	0	0		
0	0	1	0		
0	0	0	1		
0.0579	-0.0103	-0.0944	-0.2842		
0.1789	-0.0327	-0.0757	-0.3500		
-0.0861	0.0156	0.0608	0.0459		
Four Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1	0	0	0	
0	0	1	0	0	
0	0	0	1	0	
0	0	0	0	1	
-26499.7	4699.3351	21011.02	53010.416	-9759.8599	
5749.2605	-1019.5475	-4558.4514	-11500.93	2117.4091	
Five Factor Model					
C_0	C_1	C_2	C_3	C_4	C_5
0	1	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0
0	0	0	0	1	0
0	0	0	0	0	1
-0.0646	0.0114	0.0547	0.2219	-0.0778	0.0521

Notes: Contracts 1, 2, 3 refer to 6 contracts per day with one month difference between two adjacent contracts, 12 contracts per day with one month difference between two adjacent contracts, and 6 contracts per day with two months difference between two adjacent contracts respectively.

Table 3.7 Sequential SSE Test Simulation Results (General-to-Specific)

T	ρ	contract	θ	d	0	1	2	3	4	5
1000	0.5	1	0.005^2	0	.991	.000	.000	.000	.000	.000
				1	.009	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.5	1	0.025^2	0	.954	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.046	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.5	2	0.005^2	0	.625	.000	.000	.000	.000	.000
				1	.000	.641	.000	.000	.000	.000
				2	.347	.000	.974	.000	.000	.000
				3	.000	.358	.024	.997	.000	.000
				4	.028	.001	.001	.003	1.000	.000
				5	.000	.000	.001	.000	.000	1.000
1000	0.5	2	0.025^2	0	.959	.000	.000	.000	.000	.000
				1	.000	.959	.000	.000	.000	.000
				2	.011	.029	1.000	.000	.000	.000
				3	.002	.005	.000	1.000	.000	.000
				4	.012	.007	.000	.000	1.000	.000
				5	.016	.000	.000	.000	.000	1.000
1000	0.5	3	0.005^2	0	.994	.000	.000	.000	.000	.000
				1	.005	1.000	.000	.000	.000	.000
				2	.001	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.5	3	0.025^2	0	.979	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.010	.000	1.000	.000	.000	.000
				3	.011	.000	.000	1.00	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000

Table 3.7 Sequential SSE Test Simulation Results (General-to-Specific, Continued)

T	ρ	contract	θ	k	0	1	2	3	4	5
1000	0.9	1	0.005^2	0	.991	.000	.000	.000	.000	.000
				1	.009	.999	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.001	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.9	1	0.025^2	0	.954	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.046	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.9	2	0.005^2	0	.625	.000	.000	.000	.000	.000
				1	.000	.496	.000	.000	.000	.000
				2	.347	.000	.973	.000	.000	.000
				3	.000	.503	.011	.997	.000	.000
				4	.028	.000	.000	.003	1.000	.000
				5	.000	.001	.016	.000	.000	1.000
1000	0.9	2	0.025^2	0	.959	.000	.000	.000	.000	.000
				1	.000	.958	.000	.000	.000	.000
				2	.011	.039	1.000	.000	.000	.000
				3	.002	.000	.000	1.000	.000	.000
				4	.012	.003	.000	.000	1.000	.000
				5	.016	.000	.000	.000	.000	1.000
1000	0.9	3	0.005^2	0	.994	.000	.000	.000	.000	.000
				1	.005	1.000	.000	.000	.000	.000
				2	.001	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.9	3	0.025^2	0	.979	.000	.000	.000	.012	1.000
				1	.000	1.000	.000	.000	.000	.000
				2	.010	.000	1.000	.000	.000	.000
				3	.011	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000

Notes: Data generating process:

$$f_t = \Phi_t \; c + \Phi_t \; C \; X_t + e_t$$

$$X_t = \rho \cdot I_d \cdot X_{t-1} + \sigma \cdot u_t$$

Table 3.7 Sequential SSE Test Simulation Results (General-to-Specific, Continued)

T	ρ	contract	θ	d	0	1	2	3	4	5
200	0.5	1	0.005^2	0	.860	.000	.000	.000	.000	.000
				1	.114	.939	.000	.000	.000	.000
				2	.022	.022	1.000	.000	.000	.000
				3	.004	.000	.000	1.000	.000	.000
				4	.000	.039	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.5	1	0.025^2	0	.791	.000	.000	.000	.000	.000
				1	.000	.995	.000	.000	.000	.000
				2	.208	.000	.998	.000	.000	.000
				3	.001	.000	.002	1.000	.000	.000
				4	.000	.005	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.5	2	0.005^2	0	.500	.000	.000	.000	.000	.000
				1	.002	.633	.000	.000	.000	.000
				2	.409	.000	.942	.000	.000	.000
				3	.001	.366	.052	.944	.000	.000
				4	.085	.001	.003	.001	.998	.000
				5	.003	.000	.003	.055	.002	1.000
200	0.5	2	0.025^2	0	.895	.000	.000	.000	.000	.000
				1	.000	.919	.000	.000	.000	.000
				2	.041	.066	.999	.000	.000	.000
				3	.005	.005	.000	.987	.000	.000
				4	.035	.010	.000	.000	.998	.000
				5	.024	.000	.001	.013	.002	1.000
200	0.5	3	0.005^2	0	.837	.000	.000	.000	.000	.000
				1	.097	.969	.000	.000	.000	.000
				2	.066	.031	.998	.000	.000	.000
				3	.000	.000	.002	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.5	3	0.025^2	0	.735	.000	.000	.000	.012	1.000
				1	.000	.989	.000	.000	.000	.000
				2	.073	.000	.997	.000	.000	.000
				3	.192	.011	.003	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000

Notes: Contract 1,2,3 refer to 6 contracts per day with one month difference between two adjacent contracts, 12 contracts per day with one month difference between two adjacent contracts, and 6 contracts per day with two months difference between two adjacent contracts respectively.

Table 3.7 Sequential SSE Test Simulation Results (General-to-Specific, Continued)

T	ρ	contract	θ	d	0	1	2	3	4	5
200	0.9	1	0.005^2	0	.860	.000	.000	.000	.000	.000
				1	.114	.954	.000	.000	.000	.000
				2	.022	.026	1.000	.000	.000	.000
				3	.004	.000	.000	1.000	.000	.000
				4	.000	.020	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.9	1	0.025^2	0	.791	.000	.000	.000	.000	.000
				1	.000	.998	.000	.000	.000	.000
				2	.208	.000	.998	.000	.000	.000
				3	.001	.000	.002	1.000	.000	.000
				4	.000	.002	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.9	2	0.005^2	0	.500	.000	.000	.000	.000	.000
				1	.002	.539	.000	.000	.000	.000
				2	.409	.000	.935	.000	.000	.000
				3	.001	.460	.052	.986	.000	.000
				4	.085	.001	.000	.014	.998	.000
				5	.003	.000	.013	.000	.002	1.000
200	0.9	2	0.025^2	0	.895	.000	.000	.000	.000	.000
				1	.000	.901	.000	.000	.000	.000
				2	.041	.088	.995	.000	.000	.000
				3	.005	.000	.001	.987	.000	.000
				4	.035	.011	.000	.013	.998	.000
				5	.024	.000	.004	.000	.002	1.000
200	0.9	3	0.005^2	0	.837	.000	.000	.000	.000	.000
				1	.097	.932	.000	.000	.000	.000
				2	.066	.028	.997	.000	.000	.000
				3	.000	.000	.003	1.000	.000	.000
				4	.000	.040	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.9	3	0.025^2	0	.735	.000	.000	.000	.000	.000
				1	.000	.967	.000	.000	.000	.000
				2	.073	.033	.996	.000	.000	.000
				3	.192	.000	.004	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000

Notes: Data generating process:

$$f_t = \Phi_t \; c + \Phi_t \; C \; X_t + e_t$$

$$X_t = \rho \cdot I_d \cdot X_{t-1} + \sigma \cdot u_t$$

Notes: Contracts 1, 2, 3 refer to 6 contracts per day with one month difference between two adjacent contracts, 12 contracts per day with one month difference between two adjacent contracts, and 6 contracts per day with two months difference between two adjacent contracts respectively.

Table 3.8 Sequential SSE Test Simulation Results (Specific-to-General)

T	ρ	contract	θ	d	0	1	2	3	4	5
1000	0.5	1	0.005^2	0	.991	.000	.000	.000	.000	.000
				1	.009	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.5	1	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.5	2	0.005^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.000	.000	.976	.000	.000	.000
				3	.000	.000	.024	.997	.000	.000
				4	.000	.000	.000	.003	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.5	2	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	.971	.000	.000	.000	.000
				2	.000	.029	1.000	.000	.000	.000
				3	.000	.000	.000	.999	.000	.000
				4	.000	.000	.000	.001	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.5	3	0.005^2	0	.995	.000	.000	.000	.000	.000
				1	.005	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.5	3	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.00	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000

Table 3.8 Sequential SSE Test Simulation Results (Specific-to-General, Continued)

T	ρ	contract	θ	k	0	1	2	3	4	5
1000	0.9	1	0.005^2	0	.991	.000	.000	.000	.000	.000
				1	.009	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.9	1	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.9	2	0.005^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.000	.000	.988	.000	.000	.000
				3	.000	.000	.012	.997	.000	.000
				4	.000	.000	.000	.003	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.9	2	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	.959	.000	.000	.000	.000
				2	.000	.041	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.9	3	0.005^2	0	.995	.000	.000	.000	.000	.000
				1	.005	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
1000	0.9	3	0.025^2	0	1.000	.000	.000	.000	.012	1.000
				1	.000	1.000	.000	.000	.000	.000
				2	.000	.000	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000

Notes: Data generating process:

$$f_t = \Phi_t \; c + \Phi_t \; C \; X_t + e_t$$

$$X_t = \rho \cdot I_d \cdot X_{t-1} + \sigma \cdot u_t$$

Table 3.8 Sequential SSE Test Simulation Results (Specific-to-General, Continued)

T	ρ	contract	θ	d	0	1	2	3	4	5
200	0.5	1	0.005^2	0	.881	.000	.000	.000	.000	.000
				1	.116	.974	.000	.000	.000	.000
				2	.003	.026	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.5	1	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	1.000	.000	.000	.000	.000
				2	.000	.000	.998	.000	.000	.000
				3	.001	.000	.002	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.5	2	0.005^2	0	.998	.000	.000	.000	.000	.000
				1	.002	1.000	.000	.000	.000	.000
				2	.000	.000	.948	.000	.000	.000
				3	.000	.000	.052	.999	.000	.000
				4	.000	.000	.000	.001	.998	.000
				5	.000	.000	.000	.000	.002	1.000
200	0.5	2	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	.933	.000	.000	.000	.000
				2	.000	.067	1.000	.000	.000	.000
				3	.000	.000	.000	.987	.000	.000
				4	.000	.000	.000	.013	.998	.000
				5	.000	.000	.000	.000	.002	1.000
200	0.5	3	0.005^2	0	.902	.000	.000	.000	.000	.000
				1	.097	.969	.000	.000	.000	.000
				2	.001	.031	.998	.000	.000	.000
				3	.000	.000	.002	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.5	3	0.025^2	0	1.000	.000	.000	.000	.012	1.000
				1	.000	1.000	.000	.000	.000	.000
				2	.000	.000	.997	.000	.000	.000
				3	.000	.000	.003	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000

Notes: Contract 1,2,3 refer to 6 contracts per day with one month difference between two adjacent contracts, 12 contracts per day with one month difference between two adjacent contracts, and 6 contracts per day with two months difference between two adjacent contracts respectively.

Table 3.8 Sequential SSE Test Simulation Results (Specific-to-General, Continued)

T	ρ	contract	θ	d	0	1	2	3	4	5
200	0.9	1	0.005^2	0	.881	.000	.000	.000	.000	.000
				1	.116	.973	.000	.000	.000	.000
				2	.003	.027	1.000	.000	.000	.000
				3	.000	.000	.000	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.9	1	0.025^2	0	1.000	.000	.000	.000	.000	.001
				1	.000	1.000	.000	.000	.000	.017
				2	.000	.000	.998	.000	.000	.000
				3	.000	.000	.002	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	.982
200	0.9	2	0.005^2	0	.998	.000	.000	.000	.000	.000
				1	.002	1.000	.000	.000	.000	.040
				2	.000	.000	.945	.000	.000	.000
				3	.000	.000	.055	.986	.000	.000
				4	.000	.000	.000	.014	.998	.000
				5	.000	.000	.000	.000	.002	.960
200	0.9	2	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	.907	.000	.000	.000	.000
				2	.000	.093	.999	.000	.000	.000
				3	.000	.000	.001	.987	.000	.000
				4	.000	.000	.000	.013	.998	.000
				5	.000	.000	.000	.000	.002	1.000
200	0.9	3	0.005^2	0	.902	.000	.000	.000	.000	.000
				1	.097	.969	.000	.000	.000	.000
				2	.001	.031	.997	.000	.000	.000
				3	.000	.000	.003	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000
200	0.9	3	0.025^2	0	1.000	.000	.000	.000	.000	.000
				1	.000	.967	.000	.000	.000	.000
				2	.000	.033	.996	.000	.000	.000
				3	.000	.000	.004	1.000	.000	.000
				4	.000	.000	.000	.000	1.000	.000
				5	.000	.000	.000	.000	.000	1.000

Notes: Data generating process:

$$f_t = \Phi_t \; c + \Phi_t \; C \; X_t + e_t$$

$$X_t = \rho \cdot I_d \cdot X_{t-1} + \sigma \cdot u_t$$

Appendix 3A

For notational simplicity, the time t subscripts are eliminated in the Appendix.

Derivation of $\hat{e}$ and $Cov(\hat{e})$.

$$\begin{aligned}
\hat{e} &= f - \Phi\, c - \Phi\, C(C^{\top}\Phi^{\top}H^{-1}\Phi\, C)^{-1}C^{\top}\Phi^{\top}H^{-1}(f - \Phi\, c) \\
&= (I - JH^{-1})(f - \Phi\, c) \\
&= (I - JH^{-1})(\Phi\, C\, X + e) \\
&= (\Phi\, C - \Phi\, C(C^{\top}\Phi^{\top}H^{-1}\Phi\, C)^{-1}(C^{\top}\Phi^{\top}H^{-1}\Phi\, C))X + (I - JH^{-1})e \\
&= (I - JH^{-1})e \qquad (3.56)
\end{aligned}$$

where $J = \Phi\, C(C^{\top}\Phi^{\top}H^{-1}\Phi\, C)^{-1}C^{\top}\Phi^{\top}$.

$$\begin{aligned}
\Sigma &= (I - JH^{-1})H(I - JH^{-1}) \\
&= (I - JH^{-1})H(I - H^{-1}J) \\
&= (H - J)H^{-1}H(I - H^{-1}J) \\
&= (H - J)(I - H^{-1}J) \\
&= H - J - J + JH^{-1}J \\
&= H - J \qquad (3.57)
\end{aligned}$$

Appendix 3B

Derivation of covariance matrix of $\tilde{e}$.

$$\begin{aligned} Cov(\tilde{e}) &= E(\tilde{e}\tilde{e}^{\top}) \\ &= E(R^{\top}\hat{e}\hat{e}^{\top}R) \\ &= E(D^{-\frac{1}{2}}Z^{\top}\hat{e}\hat{e}^{\top}ZD^{-\frac{1}{2}}) \\ &= D^{-\frac{1}{2}}Z^{\top}\Sigma\, ZD^{-\frac{1}{2}} \\ &= D^{-\frac{1}{2}}Z^{\top}ZDZ^{\top}ZD^{-\frac{1}{2}} \\ &= D^{-\frac{1}{2}}DD^{-\frac{1}{2}} \\ &= I \end{aligned} \tag{3.58}$$

Appendix 3C

Proof of equivalence between $\sum_{t=1}^{T} \tilde{e}^T \tilde{e}$ and $\sum_{t=1}^{T} \hat{e}^\top H^{-1} \hat{e}$. By (2.4), $\hat{e} = (I - JH^{-1})e$, hence

$$\begin{aligned} \hat{e}^\top H^{-1} \hat{e} &= e^\top (I - H^{-1}J)H^{-1}(I - JH^{-1})e \\ &= e^\top (H^{-1} - H^{-1}JH^{-1} - H^{-1}JH^{-1} + H^{-1}JH^{-1}JH^{-1})e \\ &= e^\top (H^{-1} - H^{-1}JH^{-1})e \end{aligned} \tag{3.59}$$

While $\tilde{e} = R\hat{e}$, so

$$\begin{aligned} \tilde{e}^\top \tilde{e} &= \hat{e}^\top R^\top R \hat{e} \\ &= e(I - H^{-1}J)R^\top R(I - JH^{-1})e \\ &= e^\top H^{-1}(H - J)R^\top R(H - J)H^{-1}e \\ &= e^\top H^{-1}(ZDZ^\top)ZD^{-\frac{1}{2}}D^{-\frac{1}{2}}Z^\top (ZDZ^\top)H^{-1}e \\ &= e^\top H^{-1}(ZDZ^\top)H^{-1}e \\ &= e^\top H^{-1}(H - J)H^{-1}e \\ &= e^\top (H^{-1} - H^{-1}JH^{-1})e \end{aligned} \tag{3.60}$$

Hence the result.

Appendix 3D

Proof of Proposition 3.3.1.

Consider an arbitrary affine transformation on X and let $\tilde{X} = a + AX$, where a is $n \times 1$ vector, A is $n \times n$ nonsingular matrix.

From

$$\begin{aligned} f &= \Phi\, c + \Phi\, C\, X + e \\ &= \Phi\tilde{c} + \Phi\tilde{C}\tilde{X} + e \\ &= \Phi\tilde{c} + \Phi\tilde{C}(a + AX) + e \\ &= \Phi(\tilde{c} + \tilde{C}a) + \Phi\tilde{C}AX \end{aligned} \tag{3.61}$$

thus

$$\begin{aligned} \tilde{c} + \tilde{C}a &= c \\ \tilde{C}A &= C \end{aligned} \tag{3.62}$$

Since a is an arbitrary $k \times 1$ vector , we can perform arbitrary transformation to k elements of c without changing the fit of the model. The simplest case is to take the upper k elements to be all zeros.

Also, A is an arbitrary nonsingular $n \times n$ matrix, so we can post-multiply C by an arbitrary nonsingular $k \times k$ matrix without changing the fit of the model. From standard linear algebra result, such multiplication is equivalent to a series of fundamental column transformation to C, the simplest case is to find some A such that

$$CA = \left\{ \begin{array}{c} I_k \\ C^* \end{array} \right\} \tag{3.63}$$

Using random values for the c and C restricted as above, we find that the Hessian matrix obtained is in fact full rank. Hence the result.

Appendix 3E

Gradient Results For $H = I$, the residual at time t is equal to

$$\begin{aligned} \hat{e} &= [I - \Phi\, C(C^\top \Phi^\top \Phi\, C)^{-1} C^\top \Phi^\top](f - \Phi\, c) \\ &= [I - \Theta\, C^\top \Phi^\top]\tilde{f} \end{aligned} \tag{3.64}$$

where $\Theta = \Phi\, C(C^\top \Phi^\top \Phi\, C)^{-1}$, $\tilde{f} = f - \Phi\, c$.

$$\begin{aligned} \frac{d\hat{e}}{dc} &= (I - \Theta\, C^\top \Phi^\top)(-\Phi) \\ &= \Psi \end{aligned} \tag{3.65}$$

where Ψ is defined as $\Psi = [\Theta\, C^\top \Phi^\top \Phi - \Phi]$.

$$\frac{d\hat{e}}{dC} = (\tilde{f}^\top \Theta \otimes\ \Psi) + (\Theta \otimes\ \tilde{f}^\top \Psi) \tag{3.66}$$

where $\otimes$ is the Kronecker product operator.

The derivatives of the squared errors with respect to c and C are

$$\frac{d\hat{e}^\top \hat{e}}{dc} = 2\hat{e}^\top \Phi \tag{3.67}$$

$$\frac{d\hat{e}^\top \hat{e}}{dC} = -2(\hat{x} \otimes \Phi^\top \hat{e}) \tag{3.68}$$

where $\hat{x} = (C^\top \Phi^\top \Phi\, C)^{-1} C^\top \Phi^\top \tilde{f}$.

Appendix 3F

Hessian Results

The second derivatives are

$$\frac{d^2\hat{e}^\top\hat{e}}{dcdc^\top} = -2\Phi^\top\Psi \tag{3.69}$$

$$\frac{d^2\hat{e}^\top\hat{e}}{dcdC^\top} = -2[\Phi^\top\Theta \otimes \tilde{f}^\top\Psi + \hat{x} \otimes \Phi^\top\Psi] \tag{3.70}$$

$$\begin{aligned}\frac{d^2\hat{e}^\top\hat{e}}{dCdC^\top} = \; & -2[(\Theta^\top\Phi \otimes \Psi^\top\tilde{f}\tilde{f}^\top\Theta)T_{N,n} + (\Theta^\top\tilde{f}\tilde{f}^\top\Psi \otimes \Phi^\top\Theta)T_{N,n} \\ & +\hat{x}\hat{x}^\top \otimes \Phi^\top\Psi + (C^\top\Phi^\top\Phi\, C)^{-1} \otimes \Phi^\top\hat{e}\hat{e}^\top\Phi]\end{aligned} \tag{3.71}$$

where $T_{N,n}$ is the commutation matrix that transforms $vec(A)$ into $vec(A^\top)$, i.e., $T_{N,n}vec(A) = vec(A^\top)$. The *vec* operator vectorizes a matrix by stacking its columns. The size of $T_{N,n}$ is Nn by Nn.

Details of the Gradient and Hessian derivation is available from the author upon request. The rules for matrix differential calculus can be found in Magnus and Neudecker(1988), and Fackler (1999), where in the former, the commutation matrix $T_{N,n}$ is denoted by $K_{N,n}$.

Bibliography

[1] Anderson, T. W. (1963), "The Use of Factor Analysis in the Statistical Analysis of Multiple Time Series," *Psychometrica,* 28, 1-25

[2] Bai, J. and S. Ng (2001), "Determining the Number of Factors in Approximate Factor Models," *Econometrica,*

[3] Breeden, D. (1979), "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment," *Journal of Financial Economics,* 7, 265–296.

[4] Breeden, D. (1980), "Consumption Risk in Futures Markets," *Journal of Finance,* 35, 503–520.

[5] Brown, S.J. (1989), "The Number of Factors in Security Returns," *Journal of Finance,* 44, 1247-1262.

[6] Constantinides (1989), "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy,* 98, 519–543.

[7] Connor, G. and R.A. Korajczyk (1993), "A Test for the Number of Factors in an Approximate Factor Model," *Journal of Finance,* XLVIII: 4, 1263-1291.

[8] Cortazar,G. and E. Schwartz (summer 1994), "The Valuation of Commodity Contingent Claims," *The Journal of Derivatives,* 27–39.

[9] Cox, J.C., J.E. Intersolll and S.A. Ross (1981) "The Relation Between Forward and Futures Prices," *Journal of Financial Economics,* 9, 321-346.

[10] Cox, J.C., J.E. Intersolll and S.A. Ross (1985), "A Theory of the Term Structure of Interest Rates," *Econometrica,* 53, 385–407.

[11] Cox, J.C.and S.A. Ross (1976), "The Valuation of Options for Alternative Stochastic Processes." *Journal of Financial Economics,* 3, 145–166.

[12] Diebold, F.X., L.E. Ohanian and J. Berkowitz (1998), "Dynamic Equilibrium Economies: A Framework for Comparing Models and Data," *Review of Economic Studies,* 65, 433–451.

[13] Duffie, D. and R. Kan (1996), "A Yield-factor Model of Interest Rates," *Mathematical Finance,* 6, 379–406.

[14] Epstein, L.G. and S.E. Zinn (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica,* 57, 937–968.

[15] Epstein, L.G. and S.E. Zinn (1991), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy,* 99, 263–286.

[16] Fama, E.F. and MacBeth, J.D. (1973) "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy,,* 81, 607–636.

[17] Ferson, W.E. and G.M. Constantinides (1991), "Habit Persistence and Durability in Aggregate Consumption: Empirical Tests," *Journal of Financial Economics,* 29, 199-240.

[18] Ferson, W.E. and C.R. Harvey (1992), "Seasonality and Consumption-Based Asset Pricing," *Journal of Finance,* 57, 511–547.

[19] Frankel, J.A. (1986), "Expectations and Commodity Price Dynamics: The Overshooting Model," *American Journal of Agricultural Economics,* 344–348.

[20] Hansen, L.P., J. Heaton and E.G.J. Luttmer (1995), "Econometric Evaluation of Asset Pricing Models," *Review of Financial Studies,* 8, 237–274.

[21] Hansen, L.P. and R. Jagannathan (1991), "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy,* 99, 225–262.

[22] Hansen, L.P. and R. Jagannathan (1997), "Assessing Specification Errors in Stochastic Discount Factor Models," *Journal of Finance,* LII No.2, 557–590.

[23] Hansen, L.P. and T.J. Sargent (1993), "Seasonality and Approximation Errors in Rational Expectation Models," *Journal of Econometrics,* 55, 21–55.

[24] Hansen, L.P. and K. Singleton (1983), "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy,* 91, 249–265.

[25] Labys, W.C. and C.W.J. Granger (1970), *Speculation, Hedging and Commodity Price Forecasts,* Lexington Books: Lexington, MA.

[26] Litterman, R. and J. Scheinkman (June 1991), "Common Factors Affecting Bond Returns," *Journal of Fixed Income,* , 54–61.

[27] Longstaff, F.A. and E. Schwartz (1992), "Interest-rate Volatility and the Term Structure: A Two-factor General Equilibrium Model," *Journal of Finance,* 47, 1259–1282.

[28] Lucas, R.E., Jr. (1978), "Asset Prices in an Exchange Economy," *Econometrica,* 46, 1429–1445.

[29] Luedecke, B.P. (1984), "An Empirical Investigation into Arbitrage and Approximate K-factor Structure on Large Asset Markets," *Doctoral Dissertation, Department of Economics, University of Wisconsin.*

[30] Magnus, J.R. and H. Neudecker (1988), "Matrix Differential Calculus with Applications in Statistics and Econometrics," John Wiley and Sons Ltd, Chichester, England.

[31] Mehra, R. and E. Prescott (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics,* 15, 145–161.

[32] Merton, R.C. (1973), "An Intertemporal Capital Asset Pricing Model," *Econometrica,* 41, 867–887.

[33] Pindyck, R.S. (1993), "The Present Value Model of Rational Commodity Pricing," *Economic Journal,* 103, 511–530.

[34] Raynauld, J. and J. Tissier (1984), "Risk Premiums in Futures Markets: An Empirical Investigation," *Journal of Futures Markets,* 4, 189–211.

[35] Ross, S.A. (1976), "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory,* 13, 341-360.

[36] Schwartz, E. (1997), "The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging," *Journal of Finance,* 52, 923–973.

[37] Stock, J.H. and M.W. Watson (1998), "Diffusion Indexes," *NBER Working Paper 6702.*

[38] Sundaresan, S.M. (1989), "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Studies,* 2, 73–89.

[39] Trzcinka, C. (1986), "On the Number of Factors in the Arbitrage Pricing Model," *Journal of Finance,* 41, 347-368.

[40] Vasicek, O. (1977), "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics,* 5, 177–188.

[41] Watson, M.W. (1993), "Measures of Fit for Calibrated Models," *Journal of Political Economy,* 101, 1011–1041.

[42] Weil, P. (1989), "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics,* 24, 401–421.