

NCCC-134

APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

Optimal Hedging in the Presence of Estimation Risk

by

Sergio H. Lence and Dermot J. Hayes

Suggested citation format:

Lence, S. H., and D. J. Hayes. 1992. "Optimal Hedging in the Presence of Estimation Risk." Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago, IL. [<http://www.farmdoc.uiuc.edu/nccc134>].

OPTIMAL HEDGING IN THE PRESENCE OF ESTIMATION RISK

Sergio H. Lence and Dermot J. Hayes*

Many of the contributions of economic theory to the study of financial markets involve the use of parameters that are assumed known or estimable to the economic agents. Examples include the risk-premium coefficient (β) and the security market lines that underlie the capital asset pricing model, the volatility measures required for option pricing, and the minimum variance hedge ratio and expected basis used to construct portfolios of futures contracts and cash positions. Generally, it is left to practitioners to provide estimates of these parameters which are then substituted directly for the actual parameter values of the theoretical model. Whenever textbooks demonstrate how these parameters can be estimated there is usually a cookbook formula provided which requires some relatively straightforward analysis of the relevant data (e.g., Marshall, Ch. 6 and 7; Cox and Rubinstein, Ch. 6).

A problem with the procedure described above is that the theory has not been concerned that the practical application of the model involves parameter estimates. Therefore, important information such as that conveyed by the standard errors of the estimated parameters is not used. Often practitioners (or the market participants) will have information about the size of the standard errors, the sensitivity of the estimated parameters to changes in the time period used for the estimation, or some type of nonsample information (such as insider information or subjective priors) which seems important for decision making, and yet is ignored by the theory. For example, a market participant may have available two estimates of the volatility of a particular price, one based on historic data, and another based on the expert opinion of a market analyst. Each estimate in turn will generally imply a different option premium when employed in the Black-Scholes option-pricing formula. It is not immediately clear that using the sample estimate in place of the expert estimate is the correct approach. Nor is it obvious that the theory is correct in ignoring prior (or nonsample) and other sample information. In other words, if we admit that we do not know the true values of the parameters or moments required by the theory, then the behavioral rules predicted by an appropriate theoretical model may be different.

Hedging is an area in which relaxing the assumption of perfect knowledge about the relevant parameters is potentially important. The most popular paradigm in the theory of hedging is the portfolio approach, according to which the amount hedged depends on the tradeoff between return and risk (Johnson, Stein 1961, Anderson and Danthine). When (i) the decision maker is extremely risk averse or (ii) the futures prices are unbiased, this model indicates that the optimum futures position is the well-known *minimum variance* hedge (Ederington; Benninga, Eldor, and Zilcha). The portfolio model has been developed under the assumption of perfect knowledge about the relevant parameters, but in empirical applications the standard procedure is to substitute the parameter estimates based on sample information for the true parameters. There are at least two shortcomings with this approach. One is that the decision rule so obtained is not necessarily optimal under the given circumstances, i.e., the theory itself may be wrong. The second is that such a procedure offers no guidance as to how to proceed when there is relevant nonsample information.

The objectives of this paper are (i) to show how the standard portfolio model of hedging is modified when there is less than perfect information about the relevant parameters, (ii) to advance a practical hedging model to use when both sample and prior (nonsample) information are available, and (iii) assess the losses involved with employing suboptimal decision rules. The first section reviews the method one would use to determine the optimum futures position if one knew the true

*Sergio H. Lence and Dermot J. Hayes are, respectively, Postdoctoral Research Associate and Associate Professor, Department of Economics, Iowa State University, Ames.

relevant parameters. We then discuss the problems that arise when these parameters are unknown and review the theory of optimal decision making under estimation risk. Next we show how the theory can be applied assuming that cash and futures prices follow a multivariate normal-Wishart distribution. In the fourth section, we combine the material presented in the first three sections to derive the optimum futures position in the presence of estimation risk, and discuss and compare this method with the standard approach. Finally, we summarize the main conclusions of the study.

The Optimum Futures Position in the Absence of Estimation Risk

Since its development by Johnson and Stein (1961), the portfolio approach has become the most widely used model of hedging. The portfolio model states that the optimum futures position depends on both the return and the risk of the futures position. In this section, we will sketch its derivation and succinctly discuss it in reference to the objectives of the paper.¹

Consider a decision maker whose terminal wealth (W_{T+1}) equals his initial wealth (W_T) plus the returns to his cash and the futures positions, i.e.,

$$(1) \quad W_{T+1} = W_T + p_{T+1} Q - c(Q) + (f_{T+1} - f_T) F$$

where p_{T+1} is the random cash price at date $T+1$, Q is the amount of product sold at date $T+1$, $c(Q)$ is the cost of producing (or storing) Q , f_{T+1} is the random futures price prevailing at date $T+1$ for delivery at some date $T+h \geq T+1$, f_T is the current futures price for delivery at date $T+h$, and F is the amount bought in the futures market at date T and sold at date $T+1$.² For convenience of exposition, we will assume that the decision maker is characterized by a negative exponential utility function and that the random variables are jointly normally distributed.³ Then, mean-variance analysis applies and choosing the futures position that maximizes expected utility of terminal wealth is equivalent to

$$(2) \quad \max_F [E(W_{T+1}) - \frac{1}{2} \lambda \text{Var}(W_{T+1})]$$

where: $E(W_{T+1}) = W_T + E(p_{T+1}) Q - c(Q) + [E(f_{T+1}) - f_T] F$

λ = coefficient of absolute risk aversion

$$\text{Var}(W_{T+1}) = Q^2 \text{Var}(p_{T+1}) + 2 Q F \text{Cov}(p_{T+1}, f_{T+1}) + F^2 \text{Var}(f_{T+1})$$

$E(\cdot)$ is the expectation operator and $\text{Var}(\cdot)$ denotes variance. The first order condition (FOC) corresponding to (2) is

$$(3) \quad [E(f_{T+1}) - f_T] - \frac{1}{2} \lambda [2 Q \text{Cov}(p_{T+1}, f_{T+1}) + 2 F \text{Var}(f_{T+1})] = 0$$

which yields the solution

$$(4) \quad F = \frac{[E(f_{T+1}) - f_T]}{\lambda \text{Var}(f_{T+1})} - \frac{\text{Cov}(p_{T+1}, f_{T+1})}{\text{Var}(f_{T+1})} Q$$

¹The reader is referred to Anderson and Danthine for a thorough analysis of the portfolio model of hedging.

²A negative F indicates that the futures contracts are sold at date T and bought at $T+1$.

³The analysis would be analogous if either the returns (p_{T+1}/p_T and f_{T+1}/f_T) or the price differences ($p_{T+1} - p_T$ and $f_{T+1} - f_T$) were assumed to be jointly normally distributed. Although the alternatives seem more realistic than assuming the price levels (p_{T+1} and f_{T+1}) to be jointly normally distributed, we preferred the latter characterization because it is the most straightforward and the conclusions of the study remain the same.

Expression (4) states that the optimal futures position is made of two components (Anderson and Danthine). The first is the *speculative* component and is represented by $[E(f_{T+1}) - f_T]/[\lambda \text{Var}(f_{T+1})]$, whereas the second is the *hedge* component and is given by $[-\text{Cov}(p_{T+1}, f_{T+1})/\text{Var}(f_{T+1}) Q]$.

The speculative term receives its name because it is the optimal futures position for a speculator, i.e., for the case in which there is a null cash position ($Q = 0$). This term equals zero if the decision maker perceives the futures price to be unbiased [i.e., $E(f_{T+1}) = f_T$]. When the agent perceives the futures price to be biased upward [$E(f_{T+1}) < f_T$] the speculative term will be negative, so that he will sell futures contracts at date T hoping to make a profit by buying them at a lower price at date $T+1$. The opposite will be true when the futures price is perceived to be biased downward. The absolute magnitude of the speculative term is inversely related to the futures risk [$\text{Var}(f_{T+1})$] and to the decision maker's degree of absolute risk aversion (λ).

The hedge term is a negative proportion of the cash position, and is so called because it is the optimal futures position when there are no incentives to speculate [i.e., $E(f_{T+1}) = f_T$]. The sign of the hedge component is the opposite of the sign of the cash position, so that the hedger will be short in the futures market if he is long in the cash market and vice versa. The ratio $\text{Cov}(p_{T+1}, f_{T+1})/\text{Var}(f_{T+1})$ is known in the literature as the *minimum variance hedge ratio*. It gives the ratio of futures to cash quantities that the hedger must have in order to minimize the variance of his combined position. The minimum variance hedge ratio is also the optimum ratio of futures and cash positions for an extremely risk-averse decision maker because the speculative term vanishes as the coefficient of absolute risk aversion (λ) tends to infinity. Note that when the individual believes that there is a cost to hedging [i.e., $E(f_{T+1}) > f_T$], he will sell all of his output in the futures market and then purchase additional futures contracts via the speculative term. The net effect is to underhedge when hedging costs exist.

The evidence from empirical studies employing historical data on futures prices is inconclusive as to whether futures prices are biased or not (Peck 1977 Section 1, Baillie and Myers, Martin and Garcia, Rausser and Carter). Stein (1986, Ch. 2) provides a theoretical model showing that biased futures prices are consistent with market equilibrium, and that the bias can be of either sign. Stein (1986, Ch. 5) also provides empirical evidence supporting his hypothesis; he shows, for example, that large speculators tend to profit from futures trading at the expense of small speculators because of their special forecasting ability. These results mean that the speculative term is very important when deciding the optimal futures position, an observation that is supported by the large share of speculation relative to total open interest (Stein 1986, Ch. 1).

Most empirical studies estimate the minimum variance hedge ratio by regressing p_{T+1} on f_{T+1} using historical data and obtaining the estimated ratio as the coefficient corresponding to f_{T+1} (Ederington, Hill and Schneeweis, Myers and Thompson). Other studies use conditional forecasts (Peck 1975) and GARCH models (Baillie and Myers, Cecchetti, Cumby, and Figlewski). Some examples of empirical estimates of the minimum variance hedge ratio for commodities and financial instruments are reported in Table 1. The most salient feature of Table 1 is the wide range of estimates for each particular commodity or financial instrument depending on the data set and the methodology used.

Despite its intuitive appeal, estimating the optimum futures position in this way has a clear shortcoming. This arises because expression (4) gives the optimum futures position *provided* the decision maker knows the *true* values of the relevant parameters [$E(f_{T+1})$, $\text{Var}(f_{T+1})$, and $\text{Cov}(p_{T+1}, f_{T+1})$] with certainty. If this is not true we must consider estimation risk, and it need not be true that (4) yields the optimum futures position when the estimated parameters are substituted for the true parameters. This concern is appropriate to *all* empirical studies that attempt to estimate the optimum futures position because if we knew the actual parameters with certainty there would be no need to estimate the optimum futures position. Furthermore, empirical studies regarding the

Table 1. Estimates of the minimum variance hedge ratio

Study	Corn	Soybeans	Wheat	8 % GNMA	90 Day T-Bills	20 Year T-Bonds
Cecchetti, Cumby and Figlewski						0.73-1.33
Ederington	0.76-1.02		0.78-0.92	0.12-1.04	0.12-0.65	
Myers and Thompson	0.85-1.04	0.87-1.12	0.61-1.10			
Baillie and Myers	0.09-1.53	0.22-1.17				

bias of futures prices and the minimum variance hedge ratio indicate that there are substantial differences among the parameter estimates obtained by different authors, thus suggesting that estimation risk is not negligible.

Another problem with employing (4) in the presence of estimation risk is that it is not clear how to proceed when parameter estimates based on nonsample information (e.g., insider information, expert opinion) are available.

Optimal Decisions under Estimation Risk⁴

Consider a decision maker characterized by a utility function $U[r(\underline{d}, \underline{y})]$, whose argument $r(\underline{d}, \underline{y})$ is a function of a vector of decision variables \underline{d} and a $(k \times 1)$ vector of future random variables $\underline{y} \equiv \underline{x}_{T+1}$ related to the decision problem. The joint probability density function (pdf) corresponding to \underline{y} is $p(\underline{y}|\underline{\theta})$, where $\underline{\theta}$ is the vector of parameters (or moments) that characterizes the pdf. Therefore, if $p(\underline{y}|\underline{\theta})$ is known with certainty, the problem of the decision maker is to find the solution to

$$(5) \quad \max_{\underline{d} \in D} \mathbb{E}_{\underline{y}|\underline{\theta}}(U) = \max_{\underline{d} \in D} \int_{\underline{Y}} U[r(\underline{d}, \underline{y})] p(\underline{y}|\underline{\theta}) d\underline{y}$$

where D is the feasible decision set and Y is the domain of \underline{y} . But in empirical applications we are generally faced with incomplete knowledge about $p(\underline{y}|\underline{\theta})$, or estimation risk (Bawa, Brown, and Klein). Estimation risk may be due to lack of knowledge about (i) the exact functional form of $p(\underline{y}|\underline{\theta})$, or (ii) the parameters contained in $\underline{\theta}$ (given that $p(\underline{y}|\underline{\theta})$ is known). In what follows we will assume a situation characterized by case (ii), i.e., we postulate that the decision maker knows the joint pdf, but he is uncertain about the parameters. For example, the agent knows that a random variable is normally distributed, but he does not know exactly the true mean and variance.

If the decision maker does not know the true values of the parameters in $\underline{\theta}$, the problem as represented in (5) cannot be solved because $\mathbb{E}_{\underline{y}|\underline{\theta}}(U)$ is a function of these unknown parameters and therefore it is also unknown. The standard solution to this problem is to substitute the point estimate $\hat{\underline{\theta}}(\underline{X})$ for the unknown parameter vector $\underline{\theta}$, where $\underline{X} \equiv (\underline{x}_1, \dots, \underline{x}_T)'$ is a $(T \times k)$ matrix of T past realizations of \underline{x} . In such a case, instead of (5) we have

$$(6) \quad \max_{\underline{d} \in D} \mathbb{E}_{\underline{y}|\hat{\underline{\theta}}(\underline{X})}(U) = \max_{\underline{d} \in D} \int_{\underline{Y}} U[r(\underline{d}, \underline{y})] p[\underline{y}|\hat{\underline{\theta}}(\underline{X})] d\underline{y}$$

For example, if $p(\underline{y}|\underline{\theta})$ were the k -variate normal pdf, then $p[\underline{y}|\hat{\underline{\theta}}(\underline{X})]$ would be the k -variate normal pdf with the sample means, variances, and covariances instead of the true means, variances, and covariances. This technique is sometimes called the parameter certainty equivalent (PCE) method

⁴For a detailed exposition of decision making under estimation risk, see DeGroot (Ch. 8).

(Bawa, Brown, and Klein), to emphasize that parameters are taken as if they were known. This approach completely ignores the uncertainty regarding the parameters in $\underline{\theta}$.

Uncertainty about $\underline{\theta}$ can be taken into account by means of the Bayes decision criterion. By Bayes theorem, we have⁵

$$(7) \quad p(\underline{\theta}|\underline{X}, I_T) \propto p(\underline{\theta}|I_T) p(\underline{X}|\underline{\theta})$$

where $p(\underline{\theta}|\underline{X}, I_T)$ is the *posterior* pdf for $\underline{\theta}$ given the sample data \underline{X} and the prior (nonsample) information I_T , \propto denotes proportionality, $p(\underline{\theta}|I_T)$ is the *prior* pdf for $\underline{\theta}$, and $p(\underline{X}|\underline{\theta})$ is the *likelihood function*. The posterior pdf conveys all prior (nonsample) and all sample information about $\underline{\theta}$ by means of the prior pdf and the likelihood function, respectively. Because the decision maker is uncertain about the true $\underline{\theta}$, the Bayes decision criterion states that the optimal decision must be made after integrating out the unknown parameters in $\mathbb{E}_{\underline{\theta}}(U)$ by applying the posterior pdf (7) to do so. That is, this criterion postulates that the appropriate objective function is

$$\begin{aligned} (8) \quad \max_{\underline{d} \in D} \mathbb{E}_{\underline{\theta}}[\mathbb{E}_{\underline{\theta}}(U)] &= \max_{\underline{d} \in D} \int_{\Theta} \left\{ \int_{\underline{Y}} U[r(\underline{d}, \underline{y})] p(\underline{y}|\underline{\theta}) d\underline{y} \right\} p(\underline{\theta}|\underline{X}, I_T) d\underline{\theta} \\ &= \max_{\underline{d} \in D} \int_{\underline{Y}} U[r(\underline{d}, \underline{y})] \left[\int_{\Theta} p(\underline{y}|\underline{\theta}) p(\underline{\theta}|\underline{X}, I_T) d\underline{\theta} \right] d\underline{y} \\ (8') \quad &= \max_{\underline{d} \in D} \int_{\underline{Y}} U[r(\underline{d}, \underline{y})] p(\underline{y}|\underline{X}, I_T) d\underline{y} \end{aligned}$$

where Θ is the domain of $\underline{\theta}$ and $p(\underline{y}|\underline{X}, I_T)$ is the *predictive* pdf of \underline{y} .⁶ In expression (8), $U[r(\underline{d}, \underline{y})]$ represents what the decision maker *wants*, $p(\underline{y}|\underline{\theta})$ denotes what he *knows*, and $p(\underline{\theta}|\underline{X}, I_T)$ is what he *believes* (Borch). The most important aspect of (8') compared with (5) is that it does not involve any unknown parameter, it only involves prior (nonsample) and sample information.

As pointed out by Bawa, Brown, and Klein, there are at least three reasons for employing the Bayes decision criterion (8) rather than the PCE (6), namely,

1. The Bayes criterion can be derived from the basic axioms advanced by von Neumann-Morgenstern and Savage, whereas the PCE approach has no axiomatic foundation.
2. The Bayes method takes into consideration all the relevant (sample as well as nonsample) information about $\underline{\theta}$ through the posterior pdf. In contrast, the PCE technique only uses the sample information contained in the point estimate $\hat{\underline{\theta}}$.
3. The Bayes model leads to decisions that have minimum average risk (or maximum average value).

The Multinormal Family of Distributions

If the particular functional forms of the pdf's involved in the Bayes criterion are not sufficiently tractable, the usefulness of this method to obtain specific solutions may be severely undermined. However, it is straightforward to obtain a simple predictive pdf if the elements in the $(k \times 1)$ vector \underline{y} are jointly normally distributed with unknown $(k \times 1)$ mean vector $\underline{\mu}$ and unknown $(k \times k)$ variance-covariance matrix Σ , i.e.,

$$(9) \quad p(\underline{y}|\underline{\theta}) = f_N^{(k)}(\underline{y}|\underline{\mu}, \Sigma)$$

⁵Recall that $p(a, e) = p(e) p(a|e) = p(a) p(e|a)$, and therefore $p(a|e) = p(a) p(e|a)/p(e) \propto p(a) p(e|a)$, where $p(a, e)$ is the joint pdf of any pair of random variables a and e , $p(a|e)$ and $p(e|a)$ are the conditional densities, and $p(a)$ and $p(e)$ are the marginal densities.

⁶The predictive pdf is obtained by employing the fact that $p(e) = \int_A p(a, e) da = \int_A p(a) p(e|a) da$, where A is the domain of a .

where $f_N^{(k)}(\cdot)$ is the k -variate normal density. To this end, let the prior pdf be multivariate normal-Wishart,⁷ so that

$$(10) \quad p(\underline{\theta} | I_T) = f_{NW}^{(k)}(\underline{\mu}, \underline{\Sigma}^{-1} | \underline{\mu}_0, \tau, \underline{\Sigma}_0, \nu) \\ = f_N^{(k)}(\underline{\mu} | \underline{\mu}_0, \tau^{-1} \underline{\Sigma}) f_W^{(k)}[\underline{\Sigma}^{-1} | (\nu \underline{\Sigma}_0)^{-1}, \nu], \tau > 0, \nu > k - 1$$

where $f_{NW}^{(k)}(\cdot)$ is the k -variate normal-Wishart density, $f_W^{(k)}(\cdot)$ is the k -variate Wishart density, $\underline{\mu}_0$ is a $(k \times 1)$ vector of prior means, $\underline{\Sigma}_0$ is a $(k \times k)$ prior variance-covariance matrix, and τ and ν are parameters that measure the strength of the prior beliefs in $\underline{\mu}_0$ and $\underline{\Sigma}_0$, respectively. The larger τ (ν) is, the stronger the decision maker's prior beliefs about $\underline{\mu}_0$ ($\underline{\Sigma}_0$) are.⁸ Given the pdf (9) and the prior pdf (10), the predictive pdf is k -variate Student- t (see Aitchison and Dunsmore):^{9,10}

$$(11) \quad p(\underline{y} | \underline{X}, I_T) = f_S^{(k)}(\underline{y} | \underline{\mu}_T, \underline{\Sigma}_T, m)$$

where: $f_S^{(k)}(\cdot) = k$ -variate Student- t density

$$\underline{\mu}_T = \omega_\tau \underline{\mu}_0 + (1 - \omega_\tau) \hat{\underline{\mu}}$$

$$m = \nu + T - 1 + \Delta(\tau)$$

$$\underline{\Sigma}_T = \left(1 + \frac{1}{\tau + T}\right) \left[1 - \frac{\Delta(\tau)}{m}\right] \{\omega_\nu \underline{\Sigma}_0 + (1 - \omega_\nu) \hat{\underline{\Sigma}} + \omega_\tau \left[\frac{T}{m - \Delta(\tau)}\right] (\hat{\underline{\mu}} - \underline{\mu}_0)(\hat{\underline{\mu}} - \underline{\mu}_0)'\}$$

$$\Delta(\tau) = \begin{cases} 0 & \text{if } \tau = 0 \\ 1 & \text{otherwise} \end{cases}$$

$$\hat{\underline{\mu}} = \underline{1}' \underline{X} / T$$

$$\hat{\underline{\Sigma}} = (\underline{X} - \underline{1} \hat{\underline{\mu}})' (\underline{X} - \underline{1} \hat{\underline{\mu}}) / (T - 1)$$

$$\omega_\tau = \tau / (\tau + T)$$

$$\omega_\nu = \nu / (\nu + T - 1)$$

$$\underline{1} = \text{vector of ones of dimension } (T \times 1)$$

Here, $\underline{\mu}_T$ and $\hat{\underline{\mu}}$ are $(k \times 1)$ vectors of posterior and sample means, whereas $\underline{\Sigma}_T$ and $\hat{\underline{\Sigma}}$ are $(k \times k)$ posterior and sample variance-covariance matrices. Note that ω_τ approaches 1 as the confidence on the prior means (represented by τ) increases relative to the number of sample observations T . Hence, as intuition would suggest, $\underline{\mu}_T$ is close to $\underline{\mu}_0$ when the decision maker has a high degree of confidence in his prior beliefs about the means or when the sample size is small, and $\underline{\mu}_T$ is close to the sample mean vector $\hat{\underline{\mu}}$ in the opposite situation. Similarly, $\underline{\Sigma}_T$ is closer to the prior variance-covariance matrix the higher the confidence on $\underline{\Sigma}_0$ (represented by ν) compared to T .

Upon substitution of (11) into (8'), the problem to solve is

$$(12) \quad \max_{\underline{d} \in D} \mathbb{E}_{\underline{\theta}}[\mathbb{E}_{\underline{y}|\underline{\theta}}(U)] = \max_{\underline{d} \in D} \int_Y U[r(\underline{d}, \underline{y})] f_S^{(k)}(\underline{y} | \underline{\mu}_T, \underline{\Sigma}_T, m) d\underline{y}$$

⁷The Wishart pdf is, loosely speaking, a multivariate generalization of the Chi-square pdf. The Wishart pdf applies when we have a $(k \times k)$ variance-covariance matrix instead of a scalar variance.

⁸For example, τ and ν are both large if the agent possess insider information about $\underline{\mu}$ and $\underline{\Sigma}$, whereas τ and ν tend to zero if the decision maker has little prior knowledge about $\underline{\mu}$ and $\underline{\Sigma}$.

⁹Note that in (11) the ranges of the prior parameters are relaxed to allow the limiting operations $\tau \rightarrow 0$ and $\nu \rightarrow -\Delta(\tau)$, which represent the absence of prior information about $\underline{\mu}$ and $\underline{\Sigma}$, respectively.

¹⁰If $T < k$, $\hat{\underline{\Sigma}} = 0$.

There remains the question of the particular functional form of the utility argument $r(\underline{d}, \underline{y})$. Unless $r(\underline{d}, \underline{y})$ is a tractable function of \underline{y} , the only way to solve the optimization problem is by numerical integration. But if $r(\underline{d}, \underline{y})$ is a linear combination of \underline{y} , the solution is straightforward. Let

$$(13) \quad r(\underline{d}, \underline{y}) = \alpha_0(\underline{d}) + \underline{\alpha}' \underline{y}$$

where $\alpha_0(\underline{d})$ is a function of the decision vector \underline{d} and $\underline{\alpha} \equiv [\alpha_1(\underline{d}), \dots, \alpha_k(\underline{d})]'$ is a $(k \times 1)$ vector of functions of \underline{d} . The crucial assumption here is that both $\alpha_0(\underline{d})$ and $\underline{\alpha}$ are independent of \underline{y} . If (i) $r(\underline{d}, \underline{y})$ is a linear combination of \underline{y} as depicted in (13), and (ii) the predictive pdf of \underline{y} is k -variate Student- t as in (11), then by standard statistical results $r(\underline{d}, \underline{y})$ follows a univariate Student- t distribution with mean $\mu_r = \alpha_0(\underline{d}) + \underline{\alpha}' \underline{\mu}_T$, variance $\sigma_r = \underline{\alpha}' \underline{\Sigma}_T \underline{\alpha}$, and m degrees of freedom (Judge et al., p. 242). Therefore, the decision problem (12) can be further simplified to yield

$$(14) \quad \max_{\underline{d} \in D} E_{\underline{y}|\underline{\theta}}[E_{\underline{y}|\underline{\theta}}(U)] = \int_R U(r) f_S(r|\mu_r, \sigma_r, m) dr$$

$$(14') \quad = \int_R U(\mu_r + \sigma_r t_m) f_S(t_m|0, \frac{m}{m-2}, m) dt_m$$

where R is the domain of $r(\underline{d}, \underline{y})$, $f_S(\cdot)$ is the univariate Student- t pdf, \mathbb{R} is the set of real numbers (which is the domain of t_m), and t_m is a random variable with a standardized univariate Student- t pdf with m degrees of freedom. Expression (14') follows from (14) because if $r(\underline{d}, \underline{y})$ is distributed as Student- t with mean μ_r , variance σ_r , and m degrees of freedom, then $t_m = (r - \mu_r)/\sigma_r$ is distributed as Student- t with mean 0, variance $m/(m-2)$, and m degrees of freedom. Expression (14') is the specification best suited to solve for the optimal decision vector \underline{d} , and is the one we will use to derive the optimum futures position.

The Optimum Futures Position under Estimation Risk

In this section we will apply the Bayes decision criterion to derive the optimum futures position in the presence of estimation risk. To remain consistent with the derivation of the optimum futures position in the absence of estimation risk [i.e., expression (4)], we will assume that the utility function is negative exponential [i.e., $U[r(\underline{d}, \underline{y})] = -\exp[-\lambda r(\underline{d}, \underline{y})]$], that its argument is terminal wealth [i.e., $r(\underline{d}, \underline{y}) = W_{T+1}$], and that cash and futures prices are jointly normally distributed.

Note that terminal wealth (1) is a linear combination of the random variables because it can be expressed as

$$(15) \quad W_{T+1} = \alpha_0(\underline{d}) + \underline{\alpha}' \underline{y}$$

$$\begin{aligned} \text{where: } \alpha_0(\underline{d}) &= W_T - c(Q) - f_T F \\ \underline{d} &= F \\ \underline{\alpha} &= (Q, F)' \\ \underline{y} &= (p_{T+1}, f_{T+1})' \end{aligned}$$

Hence, we can apply (14') to derive the optimum futures position. The objective function is

$$(16) \quad \max_F E[-\exp(-\lambda W_{T+1})] = \max_{\underline{d} \in D} \int_Y \{-\exp[-\lambda r(\underline{d}, \underline{y})]\} p(\underline{y}|\underline{X}, I_T) d\underline{y}$$

$$(16') \quad = \max_{\underline{d} \in D} \int_R \{-\exp[-\lambda (\mu_r + \sigma_r t_m)]\} f_S(t_m|0, \frac{m}{m-2}, m) dt_m$$

where (16) follows from the Bayes decision criterion and (16') is derived by using (14'). One problem with (16') is that the expectation assuming a Student- t distribution is not defined. However, if we approximate this distribution by means of the normal distribution with mean zero and variance $[m/(m-2)]$, the solution to (16') is straightforward because we can apply the mean-variance framework and obtain

$$(17) \quad \max_F \mathbb{E}[-\exp(-\lambda W_{T+1})] = \max_F [\mu_T - \frac{1}{2} \lambda (\frac{m}{m-2}) \sigma_\pi]$$

The FOC corresponding to (17) is

$$(18) \quad (\mu_{fT} - f_T) - \frac{1}{2} \lambda (\frac{m}{m-2}) (2Q \sigma_{pfT} + 2F \sigma_{ffT}) = 0$$

where: $\mu_{fT} = \omega_\tau \mu_{f0} + (1 - \omega_\tau) \hat{\mu}_f$

$$\sigma_{pfT} = (1 + \frac{1}{\tau + T}) [1 - \frac{\Delta(\tau)}{m}]$$

$$\sigma_{ffT} = (1 + \frac{1}{\tau + T}) [1 - \frac{\Delta(\tau)}{m}] \{ \omega_\nu \sigma_{ff0} + (1 - \omega_\nu) \hat{\sigma}_{ff} + \omega_\tau [\frac{T}{m - \Delta(\tau)}] (\hat{\mu}_p - \mu_{p0}) (\hat{\mu}_f - \mu_{f0}) \}$$

μ_{f0} and μ_{p0} are the sample and prior means of p_{T+1} ; μ_{fT} , $\hat{\mu}_p$ and μ_{f0} are the posterior, sample, and prior means of f_{T+1} ; σ_{pfT} , $\hat{\sigma}_{pf}$ and σ_{pf0} are the posterior, sample, and prior covariances between p_{T+1} and f_{T+1} ; and σ_{ffT} , $\hat{\sigma}_{ff}$ and σ_{ff0} are the posterior, sample, and prior variances of f_{T+1} .

Solving (18) for F yields the optimum futures position in the presence of estimation risk according to the Bayes decision criterion:

$$(19) \quad F_{BAY} = (\frac{m-2}{m}) \frac{(\mu_{fT} - f_T)}{\lambda \sigma_{ffT}} - \frac{\sigma_{pfT}}{\sigma_{ffT}} Q$$

In contrast, the standard approach used in empirical studies about the optimum futures position is the PCE method, which yields

$$(20) \quad F_{PCE} = \frac{(\hat{\mu}_f - f_T)}{\lambda \hat{\sigma}_{ff}} - \frac{\hat{\sigma}_{pf}}{\hat{\sigma}_{ff}} Q$$

It can be seen that $F_{BAY} \neq F_{PCE}$ in general. Moreover, it can also be observed that (19) is generally different from the optimum futures position in the absence of estimation risk (4). However, we will show below that (19) is a more general model than either (4) or (20) because the latter are nested in the former.

Consider first the speculative term. When the sample size is positive but the decision maker is much more confident about his prior beliefs than about the sample estimates of the futures expectation and variance, we have $\omega_\tau \rightarrow 1$ and $\omega_\nu \rightarrow 1$ and the speculative term in (19) simplifies to

$$(21) \quad (\frac{m-2}{m}) \frac{(\mu_{fT} - f_T)}{\lambda \sigma_{ffT}} = \frac{(\mu_{f0} - f_T)}{\lambda \sigma_{ff0}}$$

This expression is the speculative term in (4).

In contrast, if the decision maker holds "diffuse" prior beliefs about the mean and the variance (i.e., $\tau = \nu = 0$) the speculative term in (19) becomes

$$(22) \quad (\frac{m-2}{m}) \frac{(\mu_{fT} - f_T)}{\lambda \sigma_{ffT}} = \frac{T(T-3)}{(T^2-1)} \frac{(\hat{\mu}_f - f_T)}{\lambda \hat{\sigma}_{ff}}$$

Note that (22) is different from the speculative term obtained by means of the PCE method (20) by the factor of proportionality $[T(T-3)/(T^2-1)]$. This factor of proportionality will always lie between 0 and 1 for a finite number of observations and will get closer to 1 as the sample size grows. In other words, if the decision maker has diffuse prior beliefs, using the PCE will overstate the actual absolute value of the speculative futures position by the proportion $[3/(T-1/T)]$. For example, $[3/(T-1/T)] = 0.06$ with a sample of $T = 50$ observations.

If the sample size is so large that the quality of the sample estimates is infinitely greater than the quality of the decision maker's prior beliefs, then $\omega_\tau \rightarrow 0$ and $\omega_v \rightarrow 0$. In this instance, the speculative term in (19) collapses to $(\hat{\mu}_f - f_T)/(\lambda \hat{\sigma}_{ff})$, which is the speculative term calculated by means of the PCE approach (20). Note, however, that for this to happen it is necessary to have an infinitely large sample size.

A distinguishing feature of (19) is that it indicates an optimal way of blending prior (nonsample) and sample information. For example, consider the situation in which the decision maker has no information regarding the variance other than that provided by the sample, but he has nonsample information (or beliefs) that the futures price is positively biased (i.e., $\mu_{f0} < f_T$) and the sample mean indicates that the futures price is unbiased (i.e., $\hat{\mu}_f = f_T$). Then, the speculative term is negative according to (19), whereas it equals zero according to the PCE approach (20) and is not defined according to expression (4).

In Tables 2 and 3 we show the results of some simulations regarding the speculative term for the Bayes criterion, the PCE, and the case of perfect parameter information (PPI) [i.e., expressions (19), (20), and (4), respectively].¹¹ The speculative terms are measured in physical units of commodity traded in the futures market. There are noticeable differences among the three models. The absolute magnitude of the speculative term in (19) is negatively related to the sample variance ($\hat{\sigma}_{ff}$), whereas the sample size has a positive (although very small) effect on the speculative term. Also, the sample mean is a very important determinant of the speculative component obtained by means of the Bayes decision criterion. When the sample mean is unbiased (Table 2), the sample variance has little impact on the speculative term unless ω_v is small and ω_τ is large. Table 2 also shows that the speculative term in (19) can exceed the magnitude of that in (4); this happens when the relative strength of the prior mean (ω_p) is large but the relative strength of the prior variance (ω_v) is small and the sample variance ($\hat{\sigma}_{ff}$) is smaller than the prior variance (σ_{ff0}). Because of this particular scenario, the posterior mean-variance ratio is larger than the prior mean-variance ratio. In Table 2, the parameter exerting the greatest effect on the speculative term in (19) is the relative confidence in the prior mean (ω_p); this term is close to that in the PCE when ω_τ is small and close to that in the PPI when ω_τ is large. In contrast, ω_τ has little impact when the prior and the sample means coincide (Table 3). In this instance, the sample variance has a sizable effect on the speculative term derived by application of the Bayes decision criterion, and the same can be said about ω_v (particularly when $\hat{\sigma}_{ff} \neq \sigma_{ff0}$).

Consider now the hedge term. The minimum variance hedge ratio in (19) can be expressed as

$$(23) \quad \frac{\sigma_{pfT}}{\sigma_{ffT}} = \frac{\omega_v \sigma_{pf0} + (1 - \omega_v) \hat{\sigma}_{pf} + \omega_\tau \left[\frac{T}{m - \Delta(\tau)} \right] (\hat{\mu}_p - \mu_{p0}) (\hat{\mu}_f - \mu_{f0})}{\omega_v \sigma_{ff0} + (1 - \omega_v) \hat{\sigma}_{ff} + \omega_\tau \left[\frac{T}{m - \Delta(\tau)} \right] (\hat{\mu}_f - \mu_{f0})^2}$$

As it was discussed before, $\omega_\tau \rightarrow 1$ and $\omega_v \rightarrow 1$ when the decision maker is much more confident about his prior beliefs than about the sample estimates. In such circumstances, the minimum

¹¹The speculative term from expression (4) was calculated by substituting μ_{f0} and σ_{ff0} for $E(f_{T+1})$ and $\text{Var}(f_{T+1})$, respectively.

Table 2. Speculative terms for current futures price (f_T) equal to 1, prior futures mean (μ_{f0}) equal to 1.15, prior futures variance (σ_{ff0}) equal to 1, and sample futures mean ($\hat{\mu}_f$) equal to 1

Sample Variance	Sample Size	Relative Strength of Prior		Speculative Term Corresponding to		
		Mean	Variance	Bayes Crit.	PCE	PPI
0.5	30	0.25	0.25	$0.057/\lambda$	0	$0.15/\lambda$
0.5	30	0.25	0.75	$0.041/\lambda$	0	$0.15/\lambda$
0.5	30	0.75	0.25	$0.170/\lambda$	0	$0.15/\lambda$
0.5	30	0.75	0.75	$0.126/\lambda$	0	$0.15/\lambda$
0.5	200	0.25	0.25	$0.059/\lambda$	0	$0.15/\lambda$
0.5	200	0.25	0.75	$0.043/\lambda$	0	$0.15/\lambda$
0.5	200	0.75	0.25	$0.176/\lambda$	0	$0.15/\lambda$
0.5	200	0.75	0.75	$0.128/\lambda$	0	$0.15/\lambda$
1.0	30	0.25	0.25	$0.035/\lambda$	0	$0.15/\lambda$
1.0	30	0.25	0.75	$0.036/\lambda$	0	$0.15/\lambda$
1.0	30	0.75	0.25	$0.107/\lambda$	0	$0.15/\lambda$
1.0	30	0.75	0.75	$0.110/\lambda$	0	$0.15/\lambda$
1.0	200	0.25	0.25	$0.037/\lambda$	0	$0.15/\lambda$
1.0	200	0.25	0.75	$0.037/\lambda$	0	$0.15/\lambda$
1.0	200	0.75	0.25	$0.111/\lambda$	0	$0.15/\lambda$
1.0	200	0.75	0.75	$0.112/\lambda$	0	$0.15/\lambda$
1.5	30	0.25	0.25	$0.026/\lambda$	0	$0.15/\lambda$
1.5	30	0.25	0.75	$0.032/\lambda$	0	$0.15/\lambda$
1.5	30	0.75	0.25	$0.078/\lambda$	0	$0.15/\lambda$
1.5	30	0.75	0.75	$0.098/\lambda$	0	$0.15/\lambda$
1.5	200	0.25	0.25	$0.027/\lambda$	0	$0.15/\lambda$
1.5	200	0.25	0.75	$0.033/\lambda$	0	$0.15/\lambda$
1.5	200	0.75	0.25	$0.081/\lambda$	0	$0.15/\lambda$
1.5	200	0.75	0.75	$0.099/\lambda$	0	$0.15/\lambda$

Note: The parameter λ is the coefficient of absolute risk aversion.

Table 3. Speculative terms for current futures price (f_T) equal to 1, prior futures mean (μ_{f0}) equal to 1.15, prior futures variance (σ_{ff0}) equal to 1, and sample futures mean ($\hat{\mu}_f$) equal to 1.15

Sample Variance	Sample Size	Relative Strength of Prior		Speculative Term Corresponding to		
		Mean	Variance	Bayes Crit.	PCE	PPI
0.5	30	0.25	0.25	$0.228/\lambda$	$0.3/\lambda$	$0.15/\lambda$
0.5	30	0.25	0.75	$0.166/\lambda$	$0.3/\lambda$	$0.15/\lambda$
0.5	30	0.75	0.25	$0.232/\lambda$	$0.3/\lambda$	$0.15/\lambda$
0.5	30	0.75	0.75	$0.169/\lambda$	$0.3/\lambda$	$0.15/\lambda$
0.5	200	0.25	0.25	$0.238/\lambda$	$0.3/\lambda$	$0.15/\lambda$
0.5	200	0.25	0.75	$0.171/\lambda$	$0.3/\lambda$	$0.15/\lambda$
0.5	200	0.75	0.25	$0.239/\lambda$	$0.3/\lambda$	$0.15/\lambda$
0.5	200	0.75	0.75	$0.171/\lambda$	$0.3/\lambda$	$0.15/\lambda$
1.0	30	0.25	0.25	$0.143/\lambda$	$0.15/\lambda$	$0.15/\lambda$
1.0	30	0.25	0.75	$0.145/\lambda$	$0.15/\lambda$	$0.15/\lambda$
1.0	30	0.75	0.25	$0.145/\lambda$	$0.15/\lambda$	$0.15/\lambda$
1.0	30	0.75	0.75	$0.147/\lambda$	$0.15/\lambda$	$0.15/\lambda$
1.0	200	0.25	0.25	$0.149/\lambda$	$0.15/\lambda$	$0.15/\lambda$
1.0	200	0.25	0.75	$0.149/\lambda$	$0.15/\lambda$	$0.15/\lambda$
1.0	200	0.75	0.25	$0.149/\lambda$	$0.15/\lambda$	$0.15/\lambda$
1.0	200	0.75	0.75	$0.150/\lambda$	$0.15/\lambda$	$0.15/\lambda$
1.5	30	0.25	0.25	$0.104/\lambda$	$0.1/\lambda$	$0.15/\lambda$
1.5	30	0.25	0.75	$0.129/\lambda$	$0.1/\lambda$	$0.15/\lambda$
1.5	30	0.75	0.25	$0.105/\lambda$	$0.1/\lambda$	$0.15/\lambda$
1.5	30	0.75	0.75	$0.131/\lambda$	$0.1/\lambda$	$0.15/\lambda$
1.5	200	0.25	0.25	$0.108/\lambda$	$0.1/\lambda$	$0.15/\lambda$
1.5	200	0.25	0.75	$0.133/\lambda$	$0.1/\lambda$	$0.15/\lambda$
1.5	200	0.75	0.25	$0.109/\lambda$	$0.1/\lambda$	$0.15/\lambda$
1.5	200	0.75	0.75	$0.133/\lambda$	$0.1/\lambda$	$0.15/\lambda$

Note: The parameter λ is the coefficient of absolute risk aversion.

variance hedge ratio in (19) simplifies to $\sigma_{pf0}/\sigma_{ff0}$, which is the same as the minimum variance hedge ratio in the absence of estimation risk [i.e., expression (4)].

The opposite situation arises when the sample size is so large compared to the strength of the prior beliefs that the relative weights are $\omega_\tau \rightarrow 0$ and $\omega_v \rightarrow 0$. In this instance, the minimum variance hedge ratio in (19) collapses to $\hat{\sigma}_{pf}/\hat{\sigma}_{ff}$, i.e., it is equal to that under the PCE approach (20).

It is interesting to note that even if the decision maker's prior beliefs about the variance and covariance are such that $\omega_v \rightarrow 0$, the minimum variance hedge ratios for (19) and the PCE (20) may be different. In general, this will happen if the individual attaches some weight to his prior beliefs about the mean (i.e., $\omega_\tau > 0$). In such event, the minimum variance hedge ratio corresponding to (19) will be

$$(24) \quad \frac{\sigma_{pfT}}{\sigma_{ffT}} = \frac{\hat{\sigma}_{pf} + \omega_\tau \left(\frac{T}{T-1} \right) (\hat{\mu}_p - \mu_{p0}) (\hat{\mu}_f - \mu_{f0})}{\hat{\sigma}_{ff} + \omega_\tau \left(\frac{T}{T-1} \right) (\hat{\mu}_f - \mu_{f0})^2}$$

which is different from $\hat{\sigma}_{pf}/\hat{\sigma}_{ff}$.

In Tables 4 and 5, we illustrate the minimum variance hedge ratios corresponding to (4), (19), and (20) for a set of reasonable values for the parameters of the model. These tables reveal that there are important differences among the minimum variance hedge ratios in (19), (20), and (4) (unless $\hat{\sigma}_{pf} \equiv \sigma_{pf0}$). It can also be observed that the sample mean, the sample size, and the relative confidence in the prior mean have little effect on the minimum variance hedge ratio in (19). In contrast, the relative confidence in the prior variance and the size of the sample covariance ($\hat{\sigma}_{pf}$) are very important to determine the magnitude of the minimum variance hedge ratio in (19).

To summarize, the optimum futures position in the presence of estimation risk given by the Bayes decision criterion [expression (19)] embodies the two extreme scenarios of total lack of prior knowledge about the parameters [expression (20)] and perfect parameter information [expression (4)]. The advantage of expression (19) over either (20) or (4) is that, in addition to nesting the two extremes, it allows us to obtain the optimum futures position in the realistic scenario in which the decision maker has prior beliefs but is not completely certain about the quality of these priors. Expression (19) is useful because it can be directly applied to decision making. Another important application is for evaluating the robustness of the standard approach (20) with realistic counterexamples. For instance, one obvious prior would be to use the variance of futures implicit in the options price as σ_{ff0} . Other priors could be obtained by eliciting expert opinions about the mean and the variance of futures prices, the covariance between cash and futures, and the mean of cash prices (see Winkler).

Conclusions

The standard procedure used to estimate the optimum futures position is the parameter certainty equivalent (PCE) method. This method consists of directly substituting the sample estimates of the mean, variance, and covariance for the true unknown values in a formula derived under the assumption of perfect knowledge about these parameters. We show that the optimal futures position estimated by means of the PCE approach lacks normative value because it is generally suboptimal when there is uncertainty regarding the actual parameter values.

We provide a model that can be used to obtain an optimum futures position in the realistic situation where the decision maker has sample information and prior beliefs regarding the relevant parameters. This model is based on the Bayes decision criterion and nests both the theoretical model with perfect parameter information and the PCE formula. Our model yields the perfect parameter information paradigm when the decision maker is completely confident about his prior

Table 4. Minimum variance hedge ratios for prior futures mean (μ_{f0}) equal to 1.15, prior futures variance (σ_{ff0}) equal to 1, prior covariance (σ_{pf0}) equal to 1, sample futures mean ($\hat{\mu}_f$) equal to 1, and sample futures variance ($\hat{\sigma}_{ff}$) equal to 1

Sample Covariance	Sample Size	Relative Strength of Prior		Min. Variance Hedge Ratio Corresponding to		
		Mean	Variance	Bayes Crit.	PCE	PPI
0.5	30	0.25	0.25	0.622	0.5	1
0.5	30	0.25	0.75	0.874	0.5	1
0.5	30	0.75	0.25	0.617	0.5	1
0.5	30	0.75	0.75	0.871	0.5	1
0.5	200	0.25	0.25	0.622	0.5	1
0.5	200	0.25	0.75	0.874	0.5	1
0.5	200	0.75	0.25	0.617	0.5	1
0.5	200	0.75	0.75	0.871	0.5	1
1.0	30	0.25	0.25	0.996	1.0	1
1.0	30	0.25	0.75	0.999	1.0	1
1.0	30	0.75	0.25	0.987	1.0	1
1.0	30	0.75	0.75	0.996	1.0	1
1.0	200	0.25	0.25	0.996	1.0	1
1.0	200	0.25	0.75	0.999	1.0	1
1.0	200	0.75	0.25	0.987	1.0	1
1.0	200	0.75	0.75	0.996	1.0	1
1.5	30	0.25	0.25	1.369	1.5	1
1.5	30	0.25	0.75	1.123	1.5	1
1.5	30	0.75	0.25	1.357	1.5	1
1.5	30	0.75	0.75	1.120	1.5	1
1.5	200	0.25	0.25	1.369	1.5	1
1.5	200	0.25	0.75	1.123	1.5	1
1.5	200	0.75	0.25	1.358	1.5	1
1.5	200	0.75	0.75	1.120	1.5	1

Table 5. Minimum variance hedge ratios for prior futures mean (μ_{f0}) equal to 1.15, prior futures variance (σ_{ff0}) equal to 1, prior covariance (σ_{pf0}) equal to 1, sample futures mean ($\hat{\mu}_f$) equal to 1.15, and sample futures variance ($\hat{\sigma}_{ff}$) equal to 1

Sample Covariance	Sample Size	Relative Strength of Prior		Min. Variance Hedge Ratio Corresponding to		
		Mean	Variance	Bayes Crit.	PCE	PPI
0.5	30	0.25	0.25	0.625	0.5	1
0.5	30	0.25	0.75	0.875	0.5	1
0.5	30	0.75	0.25	0.625	0.5	1
0.5	30	0.75	0.75	0.875	0.5	1
0.5	200	0.25	0.25	0.625	0.5	1
0.5	200	0.25	0.75	0.875	0.5	1
0.5	200	0.75	0.25	0.625	0.5	1
0.5	200	0.75	0.75	0.875	0.5	1
1.0	30	0.25	0.25	1.000	1.0	1
1.0	30	0.25	0.75	1.000	1.0	1
1.0	30	0.75	0.25	1.000	1.0	1
1.0	30	0.75	0.75	1.000	1.0	1
1.0	200	0.25	0.25	1.000	1.0	1
1.0	200	0.25	0.75	1.000	1.0	1
1.0	200	0.75	0.25	1.000	1.0	1
1.0	200	0.75	0.75	1.000	1.0	1
1.5	30	0.25	0.25	1.375	1.5	1
1.5	30	0.25	0.75	1.125	1.5	1
1.5	30	0.75	0.25	1.375	1.5	1
1.5	30	0.75	0.75	1.125	1.5	1
1.5	200	0.25	0.25	1.375	1.5	1
1.5	200	0.25	0.75	1.125	1.5	1
1.5	200	0.75	0.25	1.375	1.5	1
1.5	200	0.75	0.75	1.125	1.5	1

information relative to the sample information. The PCE formula is nested within our model when the quality of the sample information is infinitely larger than the quality of the prior, and the sample size is infinite. Either case depicts a rather extreme state of affairs. In general, the decision maker will have relevant prior and sample information. In this instance, the model we advanced can be used to optimally blend both types of information.

References

- Aitchison, J. and I. R. Dunsmore. *Statistical Prediction Analysis*. New York: Cambridge University Press, 1975.
- Anderson, Ronald W. and Jean-Pierre Danthine. "Hedging and Joint Production: Theory and Illustrations." *J. Fin.* 35(1980): 487-98.
- Baillie, Richard T. and Robert J. Myers. "Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge." *J. Applied Econometrics* 6(1991): 109-124.
- Bawa, V., S. Brown, and R. Klein, eds. *Estimation Risk and Optimal Portfolio Choice*. Amsterdam: North-Holland Publishing Co., 1979.
- Benninga, S., R. Eldor, and I. Zilcha. "The Optimal Hedge Ratio in Unbiased Futures Markets." *J. Fut. Mkts.* 4(1984): 155-59.
- Borch, Karl H. *The Economics of Uncertainty*. Princeton, NJ: Princeton University Press, 1968.
- Cecchetti, S., R. Cumby, and S. Figlewski. "Estimation of Optimal Futures Hedge." *Rev. Econ. and Statist.* 70(1988): 623-30.
- Cox, John C. and Mark Rubinstein. *Options Markets*. Englewood Cliffs, NJ: Prentice Hall, Inc., 1985.
- DeGroot, Morris H. *Optimal Statistical Decisions*. New York: McGraw-Hill, 1970.
- Ederington, Louis H. "The Hedging Performance of the New Futures Markets." *J. Fin.* 34(1979): 57-70.
- Hill, Joanne and Thomas Schneeweis. "A Note on the Hedging Effectiveness of Foreign Currency Futures." *J. Futures Mkts.* 1(1981): 659-64.
- Johnson, Leland L. "The Theory of Hedging and Speculation in Commodity Futures." *Rev. Econ. Stud.* 27(1960): 139-51.
- Judge, George G., et al. *Introduction to the Theory and Practice of Econometrics*. 2nd edition, New York: John Wiley and Sons, 1988.
- Marshall, J. *Futures and Option Contracting--Theory and Practice*. Cincinnati: South-Western Publishing Co., 1989.
- Martin, Larry and Philip Garcia. "The Price-Forecasting Performance of Futures Markets for Live Cattle and Hogs: A Disaggregated Analysis." *Amer. J. Agr. Econ.* 63(1981): 209-15.
- Myers, Robert J. and Stanley R. Thompson. "Generalized Optimal Hedge Ratio Estimation." *Amer. J. Agr. Econ.* 71(1989): 858-68.
- Peck, Anne E. "Hedging and Income Stability: Concepts, Implications, and An Example." *Amer. J. Agr. Econ.* 57(1975): 410-19.
- Peck, Anne E., ed. *Selected Writings on Futures Markets*. Volume 2. Chicago: Chicago Board of Trade, 1977.
- Rausser, Gordon C. and Colin Carter. "Futures Market Efficiency in the Soybean Complex." *Rev. Econ. and Statist.* 65(1983): 469-78.
- Stein, Jerome L. "The Simultaneous Determination of Spot and Futures Prices." *Amer. Econ. Rev.* 60(1961): 1012-25.
- Stein, Jerome L. *The Economics of Futures Markets*. Oxford: Basil Blackwell, 1986.
- Winkler, Robert L. "Prior Information, Predictive Distributions, and Bayesian Model-Building." In *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*. A. Zellner (ed.), Chap. 9. Amsterdam: North-Holland Publishing Co., 1980.