The Empirical Minimum-Variance Hedge

Sergio H. Lence and Dermot J. Hayes

Decision making under unknown true parameters (estimation risk) is discussed along with Bayes' and parameter certainty equivalent (PCE) criteria. Bayes' criterion incorporates estimation risk in a manner consistent with expected utility maximization. The PCE method, which is the most commonly used, is not consistent with expected utility maximization. Bayes' criterion is employed to solve for the minimum-variance hedge ratio. Empirical application of Bayes' minimum-variance hedge ratio is addressed and illustrated. Simulations show that discrepancies between prior and sample parameters may lead to substantial differences between Bayesian and PCE minimum-variance hedges.

Key words: Bayes' criterion, estimation risk, minimum-variance hedge, risk management.

Estimation risk occurs when the joint probability density function (pdf) of a decision problem's random variables is not known with certainty. It is a common occurrence in economics. For example, parameters such as the marginal productivity of fertilizer, the elasticity of demand, and the regression coefficient of cash on futures prices are rarely known for sure. Agents making decisions in the presence of random variables generally are confronted with the additional uncertainty of less-than-perfect knowledge about the pdf governing the distribution of those variables.

Almost all studies involving decisions in the presence of estimation risk implicitly use the "plug-in" or "parameter certainty equivalent" (PCE) approach. The PCE consists of developing the theoretical decision model under the assumption that the pdf and its parameters are known with certainty. Once the optimal decision rule is derived, the empirical application proceeds by substituting sample estimates for the unknown parameters. Although it is intuitively appealing and empirically tractable, the PCE has no axiomatic foundations and is not consistent with expected utility maximization.

The shortcomings of the PCE approach can best be demonstrated with an example from price speculation. Assuming the true parameters are known, theory predicts that a risk-averse individual will speculate if the known futures mean is different from the current futures price. In order to determine empirically the optimal speculative position, the PCE method involves using the sample futures mean as a substitute for the true but unknown futures mean. Prior information, such as a possibly strong belief in the efficient market hypothesis, and sample information, such as the standard errors of the estimated parameters, are ignored. Taken to its extreme, the PCE would predict a long speculative position whenever the mean of recent futures prices is lower than the current futures price, and predict a short position when the opposite is true. This clearly is questionable speculative behavior.

When estimation risk is due to imperfect knowledge about the parameters of the joint pdf (given that the functional form of the pdf is known), Bayes' criterion is the method consistent with the expected utility paradigm (DeGroot, chapters 7 and 8). This criterion takes into account uncertainty regarding the unknown true parameters by assigning a pdf to these parameters and then integrating over the parameter space.

Bayes' criterion has been studied thoroughly in statistics (Raiffa and Schlaifer, DeGroot, Berger, Klein et al.). It has also been applied to solve important problems in finance, such as se-

---

Sergio H. Lence is assistant professor and Dermot J. Hayes is associate professor, Department of Economics, Iowa State University.

Journal Paper No. J-15025 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 2994.

Review coordinated by Steven Buccola.

curity market equilibrium (Bawa; Barry and Brown; Coles and Loewenstein), portfolio choice (Klein and Bawa; Bawa; Brown; Alexander and Resnick; Jorion; Frost and Savarino), and option pricing (Boyle and Ananthanarayanan). But until recently, Bayes' criterion was largely ignored in agricultural economics (Dixon and Barry; Pope and Ziemer; Collender and Zilberman; Collender; Chalfant, Collender, and Subramanian).

Dixon and Barry modeled an agricultural bank's allocation of funds among three assets and concluded that estimation risk influenced portfolio composition. Pope and Ziemer used second-degree stochastic dominance to examine the performance of alternative estimation methods. Using Monte Carlo simulations, they found that the plug-in approach generally performed no better than did the empirical distribution function, and that the empirical distribution generally led to more correct rankings under small sample sizes. Collender and Zilberman analyzed the optimal land allocation problem under alternative joint pdfs for crop returns, concluding that farmers with different opinions about the joint pdf of crop returns will both allocate and value land differently, even if they have the same absolute risk aversion and identical opinions about the mean and the variance of crop returns. Collender addressed the decision maker's ability to distinguish among different farm plans based on sample means and variances. Collender found, at reasonable significance levels, that it may be statistically impossible to distinguish among most estimated mean-variance combinations on the efficient frontier, even with large sample sizes. Chalfant, Collender, and Subramanian studied sampling properties of portfolio allocations based on the PCE approach. They showed that PCE allocation decisions are biased and inefficient, and proposed an alternative approach that would be unbiased and have lower variance.

We show in the present paper how to apply Bayes' criterion to calculate the minimum-variance hedge ratio. Collender and Zilberman similarly analyzed problems associated with using an incorrect functional form for the joint pdf, assuming perfect knowledge about the parameters. We are concerned instead with problems of having less-than-perfect knowledge about the parameters, assuming perfect knowledge about the functional form of the joint pdf.

Decisions under Uncertainty

The standard optimization problem under uncertainty can be represented by

$$\text{(1)} \quad \max_{\mathbf{d}\in D} \mathrm{E}_{\mathbf{y}|\theta}(U) = \max_{\mathbf{d}\in D} \int_Y U[R(\mathbf{d}, \mathbf{y})]\, p(\mathbf{y}|\boldsymbol{\theta})\, d\mathbf{y}$$

where $\mathrm{E}(\cdot)$ is the expectation operator, $U(\cdot)$ is a von Neumann-Morgenstern utility function, $R(\mathbf{d}, \mathbf{y})]$ is a function of a vector of decision variables $\mathbf{d}$ and a $(k \times 1)$ vector of future random variables $\mathbf{y} \equiv \mathbf{x}_{T+1}$ related to the decision problem, $p(\mathbf{y}|\boldsymbol{\theta})$ is the joint pdf of $\mathbf{y}$ given the vector of parameters $\boldsymbol{\theta}$, Y is the domain of $\mathbf{y}$, and D is the feasible decision set.

Decision problem (1) is the basic paradigm of expected utility theory (Hey). An important underlying assumption is that $p(\mathbf{y}|\boldsymbol{\theta})$ is perfectly known. However, in many real-world situations this assumption is not valid: there is estimation risk (Bawa, Brown, and Klein). Estimation risk may arise because of less-than-perfect knowledge about either (*i*) the functional form of $p(\mathbf{y}|\boldsymbol{\theta})$, or (*ii*) parameters contained in vector $\boldsymbol{\theta}$ (given that function $p(\mathbf{y}|\boldsymbol{\theta})$ is known with certainty). Although case (*i*) is relevant in certain situations (Bawa; Collender and Zilberman), we are concerned only with case (*ii*). In other words, we will define estimation risk as that in which the decision maker knows with certainty the functional form of joint pdf $p(\mathbf{y}|\boldsymbol{\theta})$, but has less-than-perfect knowledge about the parameters in $\boldsymbol{\theta}$. We will refer to the absence of estimation risk as a case of perfect parameter information (PPI).[1]

If $\boldsymbol{\theta}$ in (1) is not known with certainty, then $\mathrm{E}_{\mathbf{y}|\boldsymbol{\theta}}(U)$ is not known either because the expectation is a function of $\boldsymbol{\theta}$; therefore, $\mathrm{E}_{\mathbf{y}|\boldsymbol{\theta}}(U)$ cannot be maximized. Bayes' criterion provides a remedy to this situation in a manner consistent with the axioms of expected utility theory (DeGroot, chapters 7 and 8). The solution consists of taking into account the uncertainty about the parameters by postulating a joint pdf of $\boldsymbol{\theta}$ and integrating over the parameter space. That is, the decision problem is

$$\text{(2)} \quad \max_{\mathbf{d}\in D} \mathrm{E}_{\boldsymbol{\theta}}[\mathrm{E}_{\mathbf{y}|\boldsymbol{\theta}}(U)] = \max_{\mathbf{d}\in D} \int_{\Theta} \left\{ \int_Y U[R(\mathbf{d}, \mathbf{y})]\, p(\mathbf{y}|\boldsymbol{\theta})\, d\mathbf{y} \right\} p(\boldsymbol{\theta}|\mathbf{X}, I_T)\, d\boldsymbol{\theta}$$

[1] Decisions based on the PPI need not be similar to those based on the PCE. The former assumes perfect prior knowledge about the parameters, whereas the latter assumes perfect confidence in the quality of the sample information. Because there is no need for the sample information in one scenario to be identical to the prior information used in the other, the resulting decisions may be different.

where $p(\boldsymbol{\theta}|\mathbf{X}, I_T)$ is the posterior pdf of $\boldsymbol{\theta}$ given sample data matrix $\mathbf{X}$ and prior (nonsample) information I_T, and Θ is the domain of $\boldsymbol{\theta}$. Sample data matrix $\mathbf{X} \equiv (\mathbf{x}_1, \ldots, \mathbf{x}_T)'$ is a $(T \times k)$ matrix of T past realizations of $\mathbf{x}$.

The posterior pdf $p(\boldsymbol{\theta}|\mathbf{X}, I_T)$ contains all information available regarding the parameter vector $\boldsymbol{\theta}$ at decision time T. This pdf conveys all the sample and nonsample information about $\boldsymbol{\theta}$ because it is obtained by application of Bayes' theorem as follows:[2]

$$(3) \qquad p(\boldsymbol{\theta}|\mathbf{X}, I_T) = \frac{p(\boldsymbol{\theta}|I_T)\, p(\mathbf{X}|\boldsymbol{\theta})}{p(\mathbf{X}, I_T)}$$

where $p(\boldsymbol{\theta}|I_T)$ is the prior pdf of $\boldsymbol{\theta}$ and $p(\mathbf{X}|\boldsymbol{\theta})$ is the likelihood function. The prior pdf represents the decision maker's prior (nonsample) information about $\boldsymbol{\theta}$; this pdf reflects the probabilities which the agent assigns to different values of $\boldsymbol{\theta}$ based on his practical experience, knowledge, and beliefs. And according to the Likelihood Principle, all relevant experimental information about $\boldsymbol{\theta}$ after $\mathbf{X}$ is observed is contained in the likelihood function for the observed $\mathbf{X}$ (Berger, p. 28). By combining sample and nonsample information, the posterior pdf provides a better assessment about the unknown true parameter vector than does either the prior pdf or the likelihood function alone.

Assuming that $U[R(\mathbf{d}, \mathbf{y})]$ is independent of $\boldsymbol{\theta}$, (2) can be alternatively stated as

$$(4) \qquad \max_{\mathbf{d}\in D} \mathrm{E}_{\boldsymbol{\theta}}[\mathrm{E}_{\mathbf{y}|\boldsymbol{\theta}}(U)] = \max_{\mathbf{d}\in D} \int_Y U[R(\mathbf{d}, \mathbf{y})]\, p(\mathbf{y}|\mathbf{X}, I_T)\, d\mathbf{y}$$

where $p(\mathbf{y}|\mathbf{X}, I_T)$ is the predictive pdf of $\mathbf{y}$.[3] Expression (4) facilitates the comparison of Bayes' criterion with PPI case (1). The only difference between the right-hand sides of expressions (1) and (4) is that the joint pdf of $\mathbf{y}$ in the former is $p(\mathbf{y}|\boldsymbol{\theta})$, whereas in the latter it is $p(\mathbf{y}|\mathbf{X}, I_T)$. When parameter vector $\boldsymbol{\theta}$ is known with certainty, the sample $\mathbf{X}$ adds no information about the parameters; therefore, the decision maker can ignore $\mathbf{X}$ and proceed to make decisions based on the joint pdf $p(\mathbf{y}|\boldsymbol{\theta})$ as indicated in (1). In the more common situation characterized by imperfect knowledge about $\boldsymbol{\theta}$, however, it is unreasonable to ignore either prior or sample information. In such a case, the decision maker uses all the available information in an optimal manner by employing the predictive pdf. Bawa, Brown, and Klein have shown that, for any particular prior, Bayes' criterion yields the maximum expected utility.

Letting $\hat{\boldsymbol{\theta}}(\mathbf{X})$ denote the sample point estimate of the unknown parameter vector $\boldsymbol{\theta}$, the PCE method can be stated as

$$(5) \qquad \max_{\mathbf{d}\in D} \mathrm{E}_{\mathbf{y}|\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}(U) = \max_{\mathbf{d}\in D} \int_Y U[R(\mathbf{d}, \mathbf{y})]\, p[\mathbf{y}|\hat{\boldsymbol{\theta}}(\mathbf{X})]\, d\mathbf{y}$$

Simply put, in the PCE the sample point estimate $\hat{\boldsymbol{\theta}}(\mathbf{X})$ replaces the unknown vector $\boldsymbol{\theta}$ in (1); that is, parameter estimates are taken as if known with certainty. Solving the decision problem by means of the PCE generally is much easier than using Bayes' criterion, but the PCE has no axiomatic foundations. Klein et al. analyzed the necessary and sufficient conditions for the PCE approach to yield the optimal (i.e., Bayesian) solution. They show that the conditions are very restrictive and seldom fulfilled by the pdfs commonly used in economic studies. Moreover, they show that the utility loss from using the PCE rather than Bayes' criterion may be large.

Minimum-Variance Hedge Ratio

An important problem involving estimation risk is that of calculating the minimum-variance hedge ratio (MVH). The MVH is the ratio between the futures and the cash positions that minimizes income variance, given the agent's cash position. The MVH is an important paradigm in the theory of hedging and dominates the applied hedging literature.

Reduced to its essentials, the derivation of the MVH is as follows. Consider an agent at decision date T whose random terminal income π_{T+1} equals the returns from his cash and futures positions:

$$(6) \qquad \pi_{T+1} = p_{T+1}\, Q - (f_{T+1} - f_T)\, F$$

where p_{T+1} is the random cash price at date $T + 1$, Q is the amount of product sold at date $T + 1$, f_{T+1} is the random futures price prevailing at date $T + 1$ for delivery at some date $T + t \geq T + 1$, f_T is the current futures price for delivery at date $T + t$, and F is the amount sold

[2] Recall that $p(a, e) = p(e)\, p(a|e) = p(a)\, p(e|a)$, and therefore $p(a|e) = p(a)\, p(e|a)/p(e)$, 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.

[3] Expression (4) is obtained by reversing the order of integration of (2), noting that $U[R(\mathbf{d}, \mathbf{y})]$ is independent from $\boldsymbol{\theta}$, and using the fact that $p(\mathbf{y}|\mathbf{X}, I_T) = \int_\Theta p(\mathbf{y}|\boldsymbol{\theta})\, p(\boldsymbol{\theta}|\mathbf{X}, I_T)\, d\boldsymbol{\theta}$.

in the futures market at date T and purchased at date $T + 1$. The decision problem consists of selecting the hedge F that minimizes the variance of terminal income, given cash position Q:

$$(7) \quad \min_F \operatorname{var}_T(\pi_{T+1}) = \min_F [Q^2 \operatorname{var}_T(p_{T+1}) - 2QF \operatorname{cov}_T(p_{T+1}, f_{T+1}) + F^2 \operatorname{var}_T(f_{T+1})].$$

Subscripts in the variance and covariance operators denote that they are conditional on information at date T. First order condition corresponding to (7) is

$$(8) \quad \frac{\partial \operatorname{var}_T(\pi_{T+1})}{\partial F} = -2\, Q \operatorname{cov}_T(p_{T+1}, f_{T+1}) + 2\, F \operatorname{var}_T(f_{T+1}) = 0$$

which can be solved for the variance-minimizing hedge position[4]

$$(9) \quad F_{\text{PPI}} = \frac{\operatorname{cov}_T(p_{T+1}, f_{T+1})}{\operatorname{var}_T(f_{T+1})} Q.$$

The ratio $\operatorname{cov}_T(p_{T+1}, f_{T+1})/\operatorname{var}_T(f_{T+1})$ is the MVH. It has been shown that the MVH is the optimal hedge ratio if the current futures price f_T is an unbiased predictor of the posterior futures price f_{T+1}, regardless of the decision maker's absolute risk aversion (Benninga, Eldor, and Zilcha). In addition, the MVH is the optimal hedge ratio for extremely risk-averse decision makers, even if f_T is a biased predictor of f_{T+1} (Kahl). Because of these attributes, and also because of the apparent ease of the empirical calculation, MVH estimation has been the focus of numerous applied studies. Implicitly or explicitly, all such studies use the PCE approach; that is, they estimate F_{PPI} by means of F_{PCE}:

[4] The second order condition for a minimum is always satisfied because $\operatorname{var}_T(f_{T+1}) > 0$.

$$(10) \quad F_{\text{PCE}} = \frac{\hat{\sigma}_{pf}}{\hat{\sigma}_{ff}} Q$$

where $\hat{\sigma}_{pf}$ and $\hat{\sigma}_{ff}$ are the sample estimates of $\operatorname{cov}_T(p_{T+1}, f_{T+1})$ and $\operatorname{var}_T(f_{T+1})$, respectively.

Alternative methods have been applied to obtain the MVH estimate $\hat{\sigma}_{pf}/\hat{\sigma}_{ff}$. A popular technique consists of regressing cash on futures prices with historical data, then taking the futures price regression coefficient as the estimated MVH. Examples of this approach (or some variation of it) are Ederington; Myers and Thompson; and Viswanath. Other authors advocate GARCH models (Baillie and Myers) and conditional forecasts (Peck) to estimate the MVH. Large differences in estimated MVHs obtained by different authors with the same commodities suggest that MVH estimation risk is important. But if estimation risk exists, the sample estimate of the MVH need not lead to the optimal decision. Moreover, the stated properties of the MVH (i.e., optimality under unbiased futures prices or under extreme risk aversion) hold under PPI conditions, but need not hold in the presence of estimation risk.

Using Bayes' criterion to Derive the MVH in the Presence of Estimation Risk

In the MVH model, utility is given by $U[R(\mathbf{d}, \mathbf{y})] = -\operatorname{var}_T(\pi_{T+1})$, with $R(\mathbf{d}, \mathbf{y}) = \pi_{T+1}$, $\mathbf{d} = F$, and $\mathbf{y} = (p_{T+1}, f_{T+1})'$. This utility function is special in that it depends on parameters of the joint pdf of $\mathbf{y}$ because $\operatorname{var}_T(p_{T+1})$, $\operatorname{var}_T(f_{T+1})$, and $\operatorname{cov}_T(p_{T+1}, f_{T+1})$ belong to parameter set $\boldsymbol{\theta}$. Substituting this utility function into (2) and proceeding analogously to the previous derivation of F_{PPI}, the optimal hedge under Bayes' criterion can be shown to equal

$$(11) \quad F_{\text{BAY}} = \frac{\int_{\Theta} \left\{ \int_{Y} \operatorname{cov}_T(p_{T+1}, f_{T+1})\, p(\mathbf{y}|\theta)\, d\mathbf{y} \right\} p(\theta|\mathbf{X}, I_T)\, d\theta}{\int_{\Theta} \left\{ \int_{Y} \operatorname{var}_T(f_{T+1})\, p(\mathbf{y}|\theta)\, d\mathbf{y} \right\} p(\theta|\mathbf{X}, I_T)\, d\theta} Q$$

$$(11') \quad = \frac{\int_{\Theta} \operatorname{cov}_T(p_{T+1}, f_{T+1})\, p(\theta|\mathbf{X}, I_T)\, d\theta}{\int_{\Theta} \operatorname{var}_T(f_{T+1})\, p(\theta|\mathbf{X}, I_T)\, d\theta} Q$$

$$(11'') \qquad = \frac{\displaystyle\int_{\Theta} \mathrm{cov}_T(p_{T+1}, f_{T+1})\, p(\theta|I_T)\, p(\mathbf{X}|\theta)\, d\theta}{\displaystyle\int_{\Theta} \mathrm{var}_T(f_{T+1})\, p(\theta|I_T)\, p(\mathbf{X}|\theta)\, d\theta} Q.$$

In deriving (11′) from (11), we use the fact that neither $\mathrm{cov}_T(p_{T+1}, f_{T+1})$ nor $\mathrm{var}_T(f_{T+1})$ depend on a particular value of $\mathbf{y}$. Hence, we can take them outside the inner integrals on the right-hand side of (11). The resulting inner integrals are readily solved because $\int_Y p(\mathbf{y}|\boldsymbol{\theta})\, d\mathbf{y} = 1$ by the properties of pdfs. To derive (11″) from (11′), we employ the definition of the posterior pdf $p(\boldsymbol{\theta}|\mathbf{X}, I_T)$ given by (3) and the fact that the denominator on the right-hand side of (3) is independent of $\boldsymbol{\theta}$.

Expression (11″) is the most convenient base upon which to perform numerical integrations, and is therefore the most suitable expression for calculating F_{BAY} in practical applications. However, (11′) provides the best insight because it shows that the numerator is the expected covariance and the denominator is the expected futures variance, where these expectations are obtained by integrating with respect to the posterior pdf $p(\boldsymbol{\theta}|\mathbf{X}, I_T)$. Alternatively, the numerator (denominator) of (11′) is the value of the covariance (futures variance) that can be expected by combining both prior and sample information about the unknown true parameters.

It is important to note that F_{BAY} nests F_{PPI} as a limiting situation. Bayes' MVH simplifies to the PPI MVH when the decision maker knows $\mathrm{cov}_T(p_{T+1}, f_{T+1})$ and $\mathrm{var}_T(f_{T+1})$ with certainty. In that instance, prior and posterior pdfs for $\mathrm{cov}_T(p_{T+1}, f_{T+1})$ and $\mathrm{var}_T(f_{T+1})$ are identical. Furthermore, these pdfs have their mass concentrated at single points, i.e., at the certain values of $\mathrm{cov}_T(p_{T+1}, f_{T+1})$ and $\mathrm{var}_T(f_{T+1})$. Hence, the posterior means of $\mathrm{cov}_T(p_{T+1}, f_{T+1})$ and $\mathrm{var}_T(f_{T+1})$ are the same as the certain values assigned by the agent, in which case F_{BAY} in (11′) is identical to the expression for F_{PPI} (9).

In contrast, F_{PCE} is not a special case of F_{BAY}. The only possible case where Bayes' MVH might collapse to the PCE MVH is where there is no prior knowledge about parameter vector $\boldsymbol{\theta}$. With such a diffuse prior, the posterior is determined by the sample information through the likelihood function. However, the means of $\mathrm{cov}_T(p_{T+1}, f_{T+1})$ and $\mathrm{var}_T(f_{T+1})$ with respect to the likelihood function will in general be different from the maximum likelihood point estimates of $\mathrm{cov}_T(p_{T+1}, f_{T+1})$ and $\mathrm{var}_T(f_{T+1})$. Therefore, F_{BAY} and F_{PCE} will almost always differ from one another even in the extreme scenario in which the agent has diffuse priors.

It is also worth emphasizing that F_{BAY} provides a means for calculating the minimum-variance hedge when there are no sample data. Examples of such a scenario include the opening of a futures contract for a new commodity or the occurrence of a major event that changes the market structure and therefore price behavior. Under these circumstances, F_{PCE} cannot, because of lack of sample data, be estimated and F_{PPI} cannot be calculated unless the decision maker knows the parameters with certainty. However, F_{BAY} can be calculated as long as the agent has a nondiffuse prior (nonsample information) about the parameters.

Implementation

The basic element required to calculate an MVH, be it Bayes' or PCE, is the joint pdf $p(\mathbf{y}|\boldsymbol{\theta})$ of cash and futures prices. Assume the joint pdf is given by[5]

$$(12) \qquad \ln(p_{T+1}) - \ln(p_T) = \mu_p + \beta[\ln(f_T) - \ln(p_T)] + u_{pT+1}$$

$$(13) \qquad \ln(f_{T+1}) - \ln(f_T) = \mu_f + u_{fT+1}$$

$$(14) \qquad \mathbf{u}_{T+1} \text{ i.i.bn. } (\mathbf{0}, \boldsymbol{\Sigma})$$

where $\mathbf{u}_{T+1} = [u_{pT+1}, u_{fT+1}]'$

$$\mathbf{0} = [0, 0]'$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_p^2 & \rho\sigma_p\sigma_f \\ \rho\sigma_p\sigma_f & \sigma_f^2 \end{bmatrix}, \sigma_p > 0, \quad \sigma_f > 0, -1 < \rho < 1$$

and i.i.bn. means identically independently bivariate normally distributed. The vector of parameters whose values are not known with certainty is $\boldsymbol{\theta} = [\mu_p, \mu_f, \beta, \sigma_p, \sigma_f, \rho]'$.

[5] The joint pdf depicted by (12) through (14) is an example we used to show how to calculate Bayes' MVH. For example, if it is believed that cash prices exhibit seasonality, shifters should be added to (12). Alternatively, if it is believed that prices follow a bivariate t rather than a bivariate normal distribution, the distributional assumption in (14) should be changed to bivariate t.

The joint pdf represented by (12) through (14) is chosen to render the example both realistic and intuitive. Following most studies of asset price behavior, natural logarithms are used instead of price levels. This transformation reduces any positive skewness in price levels and eliminates the possibility of negative prices. The first-difference specification reflects the fact that price series usually are found to be integrated of order one. An error correction term $[\ln(f_T) - \ln(p_T)]$ appears in the equation for cash prices but not in the equation for futures; this formulation implies that the price discovery process occurs in the futures market and that cash prices adjust towards futures prices (Garbade and Silber). Coefficient β reflects the speed at which cash prices adjust towards futures prices. Finally, the bivariate normal pdf is adopted because it is standard in both empirical and theoretical studies.

The joint pdf is essential to calculating both the PCE and Bayes' MVH. In the former, the joint pdf provides the means to obtain the sample parameter estimates. In the latter, the joint pdf defines the likelihood function $p(\mathbf{X}|\boldsymbol{\theta})$. To see this point, note that the likelihood function in general is equal to

$$(15) \quad p(\mathbf{X}|\boldsymbol{\theta}) = p(p_T, f_T, p_{T-1}, f_{T-1}, \ldots, f_2, f_2, p_1, f_1|\boldsymbol{\theta})$$

$$(15') \quad = p(p_T, f_T|p_{T-1}, f_{T-1}, p_{T-2}, f_{T-2}, \ldots, p_2, f_2, p_1, f_1, \boldsymbol{\theta})\, p(p_{T-1}, f_{T-1}|p_{T-2}, f_{T-2}, p_{T-3}, f_{T-3}, \ldots, p_2, f_2, p_1, f_1, \boldsymbol{\theta}) \ldots p(p_2, f_2|p_1, f_1, \boldsymbol{\theta})$$

where (15′) follows from (15) by the properties of pdfs. But (15′) can be simplified to

$$(15'') \quad p(\mathbf{X}|\boldsymbol{\theta}) = p(p_T, f_T|p_{T-1}, f_{T-1}, \boldsymbol{\theta})\, p(p_{T-1}, f_{T-1}|p_{T-2}, f_{T-2}, \boldsymbol{\theta}) \ldots p(f_2, f_2|p_1, f_1, \boldsymbol{\theta})$$

because under the assumed joint pdf (12) through (14), prices at t depend only on prices at $t-1$. Had we assumed instead that prices follow a VAR process of order n, the right-hand side of (15″) would consist of the product of terms like $p(p_t, f_t|p_{t-1}, f_{t-1}, p_{t-2}, f_{t-2}, \ldots, p_{t-n}, f_{t-n}, \boldsymbol{\theta})$.

Under the additional assumption of bivariate normality, (15″) further specializes to

$$(15''') \quad p(\mathbf{X}|\boldsymbol{\theta}) = \prod_{t=2}^{T} \frac{1}{2\,\pi\,|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}\,\mathbf{u}_t'\,\Sigma^{-1}\,\mathbf{u}_t\right)$$

where

$$\mathbf{u}_t = \begin{bmatrix} \ln(p_t) - \ln(p_{t-1}) - \beta\,[\ln(f_{t-1}) - \ln(p_{t-1})] - \mu_p \\ \ln(f_t) - \ln(f_{t-1}) - \mu_f \end{bmatrix}$$

exp(·) is the base of natural logarithms, and $|\boldsymbol{\Sigma}|$ is the determinant of $\boldsymbol{\Sigma}$. Other joint pdfs, such as the multivariate t-distribution, lead to different functional forms for likelihood function (15‴). To derive these, simply replace the normal distribution with the distribution of choice in (15″).

Two aspects of the likelihood function are notable. First, for the given sample data (p_t, f_t; $t = 1, \ldots, T$), $p(\mathbf{X}|\boldsymbol{\theta})$ is a function only of the vector of unknown parameters $\boldsymbol{\theta} = [\mu_p, \mu_f, \beta, \sigma_p, \sigma_f, \rho]'$. Second, $p(\mathbf{X}|\boldsymbol{\theta})$ is also used (either explicitly or implicitly) to calculate the PCE MVH.

The other key input needed to calculate Bayes' MVH is prior pdf $p(\boldsymbol{\theta}|I_T)$. The prior represents the decision maker's beliefs, expert knowledge, and/or other type of nonsample information about parameter vector $\boldsymbol{\theta}$. This nonsample information must be expressed in terms of pdfs. For example, the agent may know there will be a shortage of railroad cars in the next few days which will drive down local cash prices. Intuitively, such information is important but is neglected by the PCE MVH because it is nonsample information. In contrast, Bayes' MVH incorporates the nonsample information through the prior pdf of either μ_p or β or both. To show how this can be done, assume for example that the prior is depicted by the normal distribution for μ_p, i.e., $p(\mu_p|I_T) = \phi[(\mu_p - \mu_{\mu_p})/\sigma_{\mu_p}]$, where $\phi(\cdot)$ is the standard normal pdf, μ_{μ_p} is the prior mean of μ_p, and σ_{μ_p} is the prior standard deviation of μ_p. Knowledge that cash prices will fall as a result of the railroad shortage means that μ_{μ_p} must be negative; furthermore, the greater the expected shortage (or the greater the response of local prices to transportation conditions) the more negative μ_{μ_p} should be. Prior standard deviation σ_{μ_p} is negatively related to the quality of the nonsample information. Hence, σ_{μ_p} should be smaller the higher the confidence placed on this information.

Any pdf can be used to represent the decision maker's nonsample knowledge, provided it satisfies both the pdf properties and the restrictions of the parameter space. To meet the pdf properties, the function selected to represent the prior must be positively valued and must integrate to unity.[6] To satisfy the parameter space restrictions, the function must equal zero at values that parameters cannot take. For example, the prior pdf should be zero at nonpositive variances because the latter must be positive. Similarly, the prior pdf of the correlation coefficient must be zero outside interval $(-1, 1)$ because the correlation coefficient cannot exceed one in absolute value.

To summarize, the additional requirements to calculate Bayes' MVH are (*i*) specifying the prior pdf of unknown parameters $p(\boldsymbol{\theta}|I_T)$, and (*ii*) solving the integrals in (11″).

Solving the Integrals in Bayes' MVH

The most difficult step in calculating Bayes' MVH is to solve the integrals on the right-hand side of (11″). In general, these integrals have no analytical solution and must be solved numerically. Software packages such as Mathematica™ or Mathcad® can do so without need of programming skills. There remains the problem, however, that numerical methods may fail to converge when solving multiple integrals such as those involved with F_{BAY}. An alternative procedure that avoids nonconvergence problems is sensitivity analysis. To this end, numerical integration is performed over one or a few parameters while maintaining the other parameters fixed at reasonable values, given both sample and nonsample information. Integrations are repeated for different combinations of values of the fixed parameters. The results obtained are then used to assess the MVH on the basis of the weight attached to each fixed parameter combination.

Another alternative to numerical integration is to use either importance sampling or the rejection method to approximate the integrals in (11′) or (11″). For practical applications, the latter techniques may be favored over the former because they require only a random number generator, a standard feature in statistical packages such as SHAZAM, TSP, SAS, MicroTSP, and RATS. The intuition behind the rejection method, and an explanation of how to use it, can be found in Smith and Gelfand. Kloek and van Dijk provides a good reference for importance sampling.

Application to Soybean Hedging

To illustrate the proposed method with a real-world situation, consider the case of a decision maker in North Central Iowa who, on 3 December 1992, wants to hedge soybeans on the nearby futures contract (i.e., January 1993) for one week. Assume also that the joint pdf $p(\mathbf{y}|\boldsymbol{\theta})$ of cash and futures prices is given by (12) through (14).

Sample information set ($\mathbf{X}$) consists of weekly data on cash and futures prices of soybeans since the January 1993 contract began trading in the Chicago Board of Trade (CBOT). Cash prices are North Central Iowa prices on Thursdays, as reported by the Iowa State University Market News. Futures are settlement prices on Thursdays for the January 1993 contract, and were obtained from the CBOT DataBank. The sample spans 17 October 1991 through 3 December 1992, for a total of 60 observations.

For the sake of completeness, the sample data are used to verify whether the pdf of (12) through (14) is reasonable. First, Augmented Dickey-Fuller tests are performed to investigate whether the sample data are consistent with the assumption of a single unit root in the logarithms of both cash and futures prices. Following the methodology proposed by Dolado, Jenkinson, and Sosvilla-Rivero, the null hypothesis of two unit roots is rejected but the null of one unit root cannot be rejected for both series at standard significance levels (see table 1).[7] Hence, the first-difference specification in (12) and (13) is supported by the sample data.

Second, the postulated pdf is fitted by full information maximum likelihood, yielding

---

[6] An exception to this rule arises when the decision maker has very little information about a parameter, i.e., the prior is diffuse (Zellner, pp. 41–53). In the preceding example, a diffuse prior could be modeled by letting σ_{μ_p} be very large. However, it is more convenient to just use $p(\mu_p|I_T) = 1$ if the parameter μ_p can take any positive or negative value. When a parameter can only be positive, as in the case of the standard deviation σ_p, the diffuse prior to use is $p(\sigma_p|I_T) = 1/\sigma_p$. But the diffuse priors $p(\mu_p|I_T) = 1$ and $p(\sigma_p|I_T) = 1/\sigma_p$ are not pdfs because they do not integrate to one.

[7] The initial model used to test for unit roots in table 1 is $(w_t - w_{t-1}) = \beta_0 + \beta_1 \text{ time} + \gamma_{1\ t\ 1} + \gamma_{21}(w_{t-1} - w_{t-2}) + \gamma_{22} (w_{t-2} - w_{t-3}) + \gamma_{23} (w_{t-3} - w_{t-4}) + e_t$ However, the null hypothesis of $\gamma_{21} = \gamma_{22} = \gamma_{23} = 0$ cannot be rejected for any of the series. For the $[\ln(p_t) - \ln(p_{t-1})]$ and $[\ln(f_t) - \ln(f_{t-1})]$ series, the null hypothesis of $\gamma_1 = 0$ in model (A) is rejected, allowing us to conclude that these series are stationary (Dolado, Jenkinson, and Sosvilla-Rivero). For the $\ln(p_t)$ and $\ln(f_t)$ series, in contrast, the null hypothesis of $\gamma_1 = 0$ in model (A) cannot be rejected, making it necessary to test whether β_0 and β_1 are significantly different from zero. Following the steps suggested by Dolado, Jenkinson, and Sosvilla-Rivero, it is found that neither β_0 nor β_1 are significantly different from zero. Hence, model (B) is the appropriate one to test for unit roots in the $\ln(p_t)$ and $\ln(f_t)$ series.

$$\ln(p_t) - \ln(p_{t-1}) = \underset{[-1.12]}{-0.0049} + \underset{[1.44]}{0.061}\,[\ln(f_{t-1}) - \ln(p_{t-1})] + u_{pt} \quad (16)$$

$$\ln(f_t) - \ln(f_{t-1}) = \underset{[-0.38]}{-0.0009} + u_{ft} \quad (17)$$

$$\hat{\Sigma} = \begin{bmatrix} 0.00031123 & 0.00027187 \\ 0.00027187 & 0.00029400 \end{bmatrix}$$

where t-statistics are reported within brackets below the corresponding coefficients, and $\hat{\Sigma}$ is the sample estimate of Σ. The log of the likelihood function is -359.318. Neither μ_p nor μ_f are significantly different from zero. As expected, the error correction term has a positive effect; however, this effect is not statistically significant. The lack of significance of some of the estimated parameters emphasizes one of the primary advantages of Bayes' method. Rather than ignore these high standard errors, as we would with the PCE, we can use the information to reduce our relative confidence in the sample data. Addition of lagged dependent variables does not increase the model's explanatory power. Estimated errors u_{pt} and u_{ft} behave consistently with the model's underlying assumptions (see table 2). Based on the Ljung-Box modified-Q statistics, the null hypothesis of no autocorrelation cannot be rejected for either series of estimated errors. Similarly, the test for autoregressive conditional heteroskedasticity (ARCH) proposed by Engle shows no evidence of ARCH(1) effects in the estimated errors. Finally, the null hypothesis of normality for each series of estimated errors cannot be rejected. In summary, the pdf depicted by (12) through (14) seems to represent the joint behavior of cash and futures prices well enough for the expository purposes of this paper.[8]

The use of logarithms as opposed to levels adds realism but also complexity because $\mathrm{cov}_T(p_{T+1}, f_{T+1}) \neq \mathrm{cov}_T(u_{pT+1}, u_{fT+1})$ and $\mathrm{var}_T(f_{t+1}) \neq \mathrm{var}_T(u_{fT+1})$. Under bivariate normality, Press (p. 149) has shown that the variance and covariance of any two series v_i ($i = 1, 2$) and the moments of their logarithmic transforms satisfy

$$\mathrm{cov}(v_1, v_2) = \exp\{\mu_{\ln(v_1)} + \mu_{\ln(v_2)} + [\sigma^2_{\ln(v_1)} + \sigma^2_{\ln(v_2)}]/2\}\{\exp[\rho_{12}\sigma_{\ln(v_1)}\,\sigma_{\ln(v_2)}] - 1\} \quad (18)$$

$$\mathrm{var}(v_i) = \exp[2\mu_{\ln(v_i)} + \sigma^2_{\ln(v_i)}]\{\exp[\sigma^2_{\ln(v_i)}] - 1\},\ i = 1, 2 \quad (19)$$

where μ's denote means, σ's represent standard deviations, and ρ_{12} is the correlation coefficient between $\ln(v_1)$ and $\ln(v_2)$. This parameter transformation is a consequence of using price logarithms rather than levels, and is common to both Bayes' approach and the PCE.

The fitted model was used to perform simulations on the PPI ratio (9), the PCE ratio (10), and the Bayesian ratio (11″). F_{BAY} is obtained by substituting (15‴), (18), (19), and the prior pdf into (11″), and solving the integrals. For purposes of comparison, F_{PCE} is obtained by plugging (18) and (19), and the parameter estimates shown in (16) and (17) into (10).[9] This procedure yields $F_{\mathrm{PCE}} = 0.874\ Q$.

The results of sensitivity analysis on correlation coefficient ρ are shown in figure 1.[10] This figure was built by integrating over the correlation coefficient while fixing all other parameters at the values estimated from the sample. The prior pdf of the correlation coefficient was assumed to be

$$r(\rho)\ \text{i.i.n.}\ [\mu_{r(\rho)}, \sigma^2_{r(\rho)}],\quad r(\rho) \equiv \ln[(1 + \rho)/(1 - \rho)]^{1/2} \quad (20)$$

Table 1. Augmented Dickey-Fuller Tests for Unit Roots

Model for $[\ln(p_t) - \ln(p_{t-1})]$ and $[\ln(f_t) - \ln(f_{t-1})]$ series:

(A) $(w_t - w_{t-1}) = \beta_0 + \beta_1\,\text{time} + \gamma_1 w_{t-1} + e_t$

null hypothesis: $\gamma_1 = 0$
alternative hypothesis: $\gamma_1 \neq 0$

Model for $\ln(p_t)$ and $\ln(f_t)$ series:

(B) $(w_t - w_{t-1}) = \gamma_1 w_{t-1} + e_t$

null hypothesis: $\gamma_1 = 0$
alternative hypothesis: $\gamma_1 \neq 0$

Series (w_t)	Evaluation Statistic	Critical Value 5%	Critical Value 1%
$[\ln(p_t) - \ln(p_{t-1})]$	$\hat{\tau}_t = -7.62$	-3.49	-4.14
$[\ln(f_t) - \ln(f_{t-1})]$	$\hat{\tau}_t = -8.14$	-3.49	-4.14
$\ln(p_t)$	$\hat{\tau} = 0.12$	-1.95	-2.62
$\ln(f_t)$	$\hat{\tau} = -0.47$	-1.95	-2.62

[8] Choice of the pdf may have nontrivial effects on the calculated MVH. Bayes' MVH under alternative pdfs $p(\mathbf{y}|\boldsymbol{\theta})$ can be calculated with the same techniques discussed in this article. Such simulations are omitted because of space constraints and because the focus of the paper is on the effect of unknown parameter values.

[9] Recall that $F_{\mathrm{PCE}} \neq 0.92$ ($= 0.00027187/0.000294$) Q because of log-normality.

[10] Equivalent results for σ_p and σ_f are available on request.

Table 2. Tests for Normality, Autocorrelation, and Autoregressive Conditional Heteroskedasticity (ARCH) of Estimated Errors

Test For	Estimated Error	Evaluation Statistic	p Value
Autocorrelation	u_{pt}	$Q'(1) = 0.65$	0.42
	u_{pt}	$Q'(3) = 1.03$	0.79
	u_{ft}	$Q'(1) = 0.04$	0.85
	u_{ft}	$Q'(3) = 1.96$	0.58
ARCH(1)	u_{pt}	LMA = 0.0002	0.99
	u_{ft}	LMA = 1.91	0.17
Normality	u_{pt}	LMN = 1.54	0.46
	u_{ft}	LMN = 1.08	0.58

Note: $Q'(i)$ is the Ljung-Box portmanteau test or modified-Q statistic for i-order autocorrelation (Ljung and Box). LMA is the Lagrange multiplier test for first order autoregressive conditional heteroscedasticity (Engle). LMN is the Lagrange multiplier test for normality (Greene, p. 329).

where $\mu_{r(\rho)}$ is the prior mean of $r(\rho)$ and $\sigma_{r(\rho)}$ is the prior standard deviation of $r(\rho)$. Prior (20) is based on Fisher's classical approximation. We chose this prior because it is identical to the asymptotic distribution used for hypothesis testing and for calculating confidence intervals about the correlation coefficient (Cox, pp. 119–21). In addition, the normal prior requires only two prior parameters (the prior mean and prior standard deviation) and facilitates interpretation of these prior parameters.

The horizontal axis in figure 1 depicts the range of the prior mean of $r(\rho)$. To avoid unreasonable prior means, the range was set equal to the 95% confidence interval calculated from the sample. The horizontal axis ranges from $\mu_{r(\rho)} = 1.204$ to $\mu_{r(\rho)} = 1.728$, because the 95% sample confidence interval of $r(\rho)$ is (1.204, 1.728).[11] The midpoint of the horizontal axis denotes the sample point estimate of $r(\rho)$ (e.g., $\mu_{r(\rho)} = 1.466$). The sample point estimate of correlation coefficient ρ is 0.899.[12]

The curves labeled high, medium, and low confidence denote degree of confidence in prior information relative to sample information. The low-confidence scenario is such that the prior standard deviation is four times the sample standard deviation of $r(\rho)$. Analogously, medium confidence depicts the case in which the prior and sample standard deviations are the same, and high confidence is such that the prior standard deviation is one-fourth of the sample standard deviation. The rationale for linking confidence in the prior with the prior standard deviation is that, as discussed earlier, the former is inversely related to the latter.

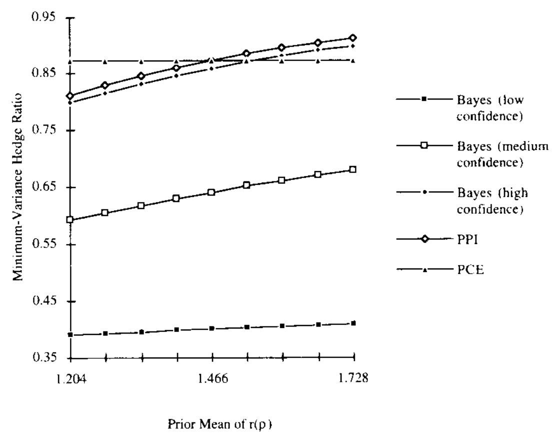

Figure 1. Minimum-variance hedge ratios for alternative priors about the coefficient of correlation

Given the sample data used, it is unlikely that the hedger could have had high confidence in the estimated value of the correlation coefficient between cash and futures prices. If there is uncertainty about this relationship, the decision maker will choose not to hedge fully (compare Bayes' medium-confidence MVH with the PCE). The underhedging makes intuitive sense. Risk averse hedgers will find it optimal to hedge all their output if markets are unbiased, transaction costs are zero, and the futures market can provide an effective hedge. If there is uncertainty about the degree of correlation between futures and cash prices, a full hedge may increase risk. The expected-utility-maximizing solution is to reduce the use of the futures market as uncertainty about the effectiveness of this market for hedging purposes increases. One of the more

[11] Note that $\rho = [\exp(2\,r) - 1]/[\exp(2\,r) + 1]$ because $r(\rho) = \ln[(1+\rho)/(1-\rho)]^{1/2}$. Therefore, the confidence interval (1.204, 1.728) for $r(\rho)$ is equivalent to the confidence interval (0.835, 0.939) for ρ.

[12] See footnote 11.

frustrating aspects of calculating optimal hedges with the PCE approach is that slight changes in the data set can induce major impacts on the "optimal" hedge. Results presented here show that this is an aberration caused by placing too much confidence in the estimated parameters.

Summary and Conclusions

Decision models generally assume perfect parameter information (PPI), that is that the true parameters characterizing joint probability density function (pdf) of the relevant random variables are known. In most applications, however, the true parameters are not known; there is estimation risk.

Bayes' criterion provides a way of dealing with estimation risk in a manner consistent with expected utility maximization. The approach assigns a pdf of the unknown true parameters based on sample and prior information, and uses this pdf to integrate the original objective function over the parameter space. Optimization is then performed over the resulting integral. Bayes' criterion has been used in statistics and finance but has been relatively neglected in agricultural economics. The standard technique employed in agricultural economics is the parameter certainty equivalent (PCE). The PCE consists of substituting sample estimates of the unknown true parameters into the PPI decision rule. The PCE approach is easier to implement than Bayes' criterion but is not consistent with expected utility maximization. Moreover, PCE decision rules generally differ from the Bayesian decision rules.

The minimum-variance hedge ratio is the ratio of futures to cash positions minimizing the variance of income, given a particular cash position. Empirical estimation of the minimum-variance hedge ratio has been the subject of many studies employing the PCE approach and is a clear example of a problem involving estimation risk. Simulation of a practical soybean hedge reveals that estimation risk may lead to substantial differences between PCE and Bayesian solutions of the variance-minimizing hedge ratio. The Bayesian hedge ratio also departs substantially from the PPI hedge ratio when parameters are not known with certainty. Such discrepancies highlight the superiority of Bayes' criterion over the PCE or PPI approaches in that neither of the latter yield decision rules which combine sample and prior information.

[Received July 1992; final revision received August 1993.]

References

Alexander, G. J., and B. G. Resnick. "More on Estimation Risk and Simple Rules for Optimal Portfolio Selection." *J. Finan.* 40(March 1985):125–33.

Baillie, R. T., and R. J. Myers. "Bivariate GARCH Estimation of the Optimal Commodity Futures Hedge." *J. Applied Econometrics* 6(April-June 1991):109–24.

Barry, C. B., and S. J. Brown. "Differential Information and Security Market Equilibrium." *J. Finan. and Quant. Analysis* 20(December 1985):407–22.

Bawa, V. S. "Portfolio Choice and Capital Market Equilibrium with Unknown Distributions." *Estimation Risk and Optimal Portfolio Choice*, eds. Bawa, V., S. Brown, and R. Klein, chapter 7. Amsterdam: North-Holland Publishing Co., 1979.

Bawa, V., S. Brown, and R. Klein. "Estimation Risk: An Introduction." *Estimation Risk and Optimal Portfolio Choice*, eds. Bawa, V., S. Brown, and R. Klein, chapter 1. Amsterdam: North-Holland Publishing Co., 1979.

Benninga, S., R. Eldor, and I. Zilcha. "The Optimal Hedge Ratio in Unbiased Futures Markets." *J. Futures Mkts.* 4(Summer 1984):155–59.

Berger, J. O. *Statistical Decision Theory and Bayesian Analysis, 2nd edition.* New York: Springer-Verlag, 1985.

Boyle, P. P., and A. L. Ananthanarayanan. "The Impact of Variance Estimation in Option Valuation Models." *J. Finan. Econ.* 5(December 1977):375–87.

Brown, S. J. "Estimation Risk and Optimal Portfolio Choice: The Sharpe Index Model." *Estimation Risk and Optimal Portfolio Choice*, eds. Bawa, V., S. Brown, and R. Klein, chapter 9. Amsterdam: North-Holland Publishing Co., 1979.

Chalfant, J. A., R. N. Collender, and S. Subramanian. "The Mean and Variance of the Mean-Variance Decision Rule." *Amer. J. Agr. Econ.* 72(November 1990):966–74.

Coles, J. L., and U. Loewenstein. "Equilibrium Pricing and Portfolio Composition in the Presence of Uncertain Parameters." *J. Finan. Econ.* 22(December 1988):279–303.

Collender, R. N. "Estimation Risk in Farm Planning under Uncertainty." *Amer. J. Agr. Econ.* 71(November 1989):996–1002.

Collender, R. N., and D. Zilberman. "Land Allocation under Uncertainty for Alternative Specifications of Return Distributions." *Amer. J. Agr. Econ.* 67(November 1985):779–86.

Cox, C. P. *A Handbook of Introductory Statistical Methods.* New York: John Wiley & Sons, 1987.

DeGroot, M. H. *Optimal Statistical Decisions.* New York: McGraw-Hill, 1970.

Dixon, B. L., and P. J. Barry. "Portfolio Analysis Considering Estimation Risk and Imperfect Markets." *West. J. Agri. Econ.* 8(December 1983): 103–11.

Dolado, J., T. Jenkinson, and S. Sosvilla-Rivero. "Cointegration and Unit Roots." *J. Econ. Surveys* 4(1990): 249–73.

Ederington, L. H. "The Hedging Performance of the New Futures Markets." *J. Finan.* 34(March 1979):157–70.

Engle, R. F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica* 50(July 1982):987–1007.

Frost, P. A., and J. E. Savarino. "An Empirical Bayes Approach to Efficient Portfolio Selection." *J. Finan. and Quant. Analysis* 21(September 1986):293–305.

Garbade, K. D., and W. L. Silber. "Price Movements and Price Discovery in Futures and Cash Markets." *Rev. Econ. and Statist.* 65(May 1983):289–97.

Greene, W. H. *Econometric Analysis.* New York: Macmillan Publishing Co., 1990.

Jorion, P. "International Portfolio Diversification with Estimation Risk." *J. Bus.* 58(July 1985):259–78.

Kahl, K. H. "Determination of the Recommended Hedging Ratio." *Amer. J. Agr. Econ.* 65(August 1983):603–05.

Klein, R. W., and V. S. Bawa. "The Effect of Limited Information and Estimation Risk on Optimal Portfolio Diversification." *J. Finan. Econ.* 5(August 1977):89–111.

Klein, R. W., L. C. Rafsky, D. S. Sibley, and R. D. Willig. "Decisions with Estimation Uncertainty." *Econometrica* 46(November 1978):1363–87.

Kloek, T., and H. K. van Dijk. "Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo." *Econometrica* 46(January 1978):1–19.

Ljung, G. M., and G. E. P. Box. "On a Measure of Lack of Fit in Time Series Models." *Biomètrika* 65(August 1978):297–303.

Myers, R. J., and S. R. Thompson. "Generalized Optimal Hedge Ratio Estimation." *Amer. J. Agr. Econ.* 71(November 1989):858–68.

Peck, A. E. "Hedging and Income Stability: Concepts, Implications, and An Example." *Amer. J. Agr. Econ.* 57(August 1975):410–19.

Pope, R. D., and R. F. Ziemer. "Stochastic Efficiency, Normality, and Sampling Errors in Agricultural Risk Analysis." *Amer. J. Agr. Econ.* 66(February 1984):31–40.

Press, J. S. *Applied Multivariate Analysis: Using Bayesian and Frequentist Methods of Inference, 2nd. edition.* Malabar, Florida: Robert E. Krieger Publishing Company, 1982.

Raiffa, H., and R. Schlaifer. *Applied Statistical Decision Theory.* Boston: Harvard University Press, 1961.

Smith, A. F. M., and A. E. Gelfand. "Bayesian Statistics without Tears: A Sampling-Resampling Perspective." *Amer. Statistician* 46(May 1992):84–88.

Viswanath, P. V. "Efficient Use of Information, Convergence Adjustments, and Regression Estimates of Hedge Ratios." *J. Futures Mkts.* 13(February 1993):43–53.

Zellner, A. *An Introduction to Bayesian Inference in Econometrics.* New York: John Wiley and Sons, Inc., 1971.