THE JOURNAL OF FINANCE • VOL. XLIX, NO. 1 • MARCH 1994

Parameter-based Decision Making under Estimation Risk: An Application to Futures Trading

SERGIO H. LENCE and DERMOT J. HAYES*

ABSTRACT

This study shows how the standard portfolio model of futures trading should be modified when there is less than perfect information about the relevant parameters (estimation risk). The standard and the optimal decision rules for futures trading in the presence of estimation risk are compared and discussed. An operational model of futures trading for use under estimation risk is advanced. In the presence of relevant prior and sample information, the model can be used to optimally blend both types of information.

MANY OF THE CONTRIBUTIONS of economic theory to the study of financial markets involve the use of parameters that are assumed known or estimable by 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 directly substituted 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 that requires some relatively straightforward analysis of the relevant data (e.g., Marshall (1989), chapters 6 and 7, Cox and Rubinstein (1985), chapter 6).

The problems with the procedure described above are twofold. First, the theory has been slow to acknowledge that the practical application of the model almost always involves parameter estimates. Generally, the additional risk of not knowing the true parameters is completely neglected. Second, practitioners (or the market participants) will often have some type of nonsample information (such as insider information or subjective priors) that seems relevant for decision making. However, this potentially important information is ignored altogether when only sample data are used to obtain the parameter estimates.

The problem of decision making under imperfect knowledge about the relevant parameters has been recognized for a long time. Decision making in

* The authors are from the Department of Economics, Iowa State University, Ames. This is Journal Paper No. J-14929 of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa. Project No. 2994.

the presence of estimation risk, as it is usually called, has been studied by Raiffa and Schlaifer (1961), DeGroot (1970), and Zellner (1971), among others. In the financial literature, the issue of estimation risk has been analyzed with regard to option pricing (Boyle and Ananthanarayanan (1977)), portfolio choice, and security market equilibrium. Bawa, Brown, and Klein (1979) review most of the developments in the last two areas prior to 1978. More recent studies on security market equilibrium include Barry and Brown (1985), Coles and Loewenstein (1988), and Coles, Loewenstein, and Suay (1991). For work related to portfolio choice, see Chen and Brown (1983), Alexander and Resnick (1985), Jorion (1985), Frost and Savarino (1986), and Cheung and Kwan (1988).

Futures trading is an area in which relaxing the assumption of perfect knowledge about the relevant parameters is potentially important. A popular paradigm in the futures literature is the portfolio model, according to which the number of futures contracts purchased depends on the tradeoff between return and risk (Johnson (1960), Stein (1961), Anderson and Danthine (1980)). This 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 sample estimates for the true but unknown parameters. Therefore, the standard approach to estimate the optimal futures position has the two problems mentioned earlier, i.e., the decision rule is not necessarily optimal under the given circumstances (the theory itself may be wrong) and the procedure offers no guidance as to how to proceed when there is relevant nonsample information.

This article (1) shows how the standard portfolio model of futures trading should be modified when there is less than perfect information about the relevant parameters, (2) discusses the differences between the standard and the optimal decision rules for futures trading, and (3) advances a practical model of futures trading for use in the presence of estimation risk. The operational rule for futures trading under estimation risk is marginally more complicated than the traditional rule, and as such it can be easily adopted by practitioners and researchers.

The plan of the article is as follows. In Section I, we discuss the problems that arise when the relevant parameters are unknown and review the theory of optimal decision making under estimation risk. We then review the standard model of the optimum futures position in Section II. In Section III, we combine the material presented in the first two sections to derive the optimum futures position in the presence of estimation risk, and discuss and compare this method with the standard approach. In Section IV, we summarize the main conclusions of the study.

I. Optimal Decisions under Estimation Risk

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.[1] 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

$$\underline{d}_{PPI} = \underset{\underline{d} \in D}{\operatorname{argmax}} E_{\underline{y}|\underline{\theta}}\left\{U\left[r\left(\underline{d}, \underline{y}\right)\right]\right\} = \underset{\underline{d} \in D}{\operatorname{argmax}} \int_Y U\left[r\left(\underline{d}, \underline{y}\right)\right] p\left(\underline{y}|\underline{\theta}\right) d\underline{y}, \quad (1)$$

where D is the feasible decision set, $E(\cdot)$ is the expectation operator, and Y is the domain of $\underline{y}$. The subscript PPI in (1) is the acronym for "perfect parameter information." But in empirical applications, we are generally faced with incomplete knowledge about $p(\underline{y}|\underline{\theta})$, or estimation risk (Bawa, Brown, and Klein (1979)). Estimation risk may be attributable 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.

If the decision maker does not know the true values of the parameters in $\underline{\theta}$, the problem as represented in (1) cannot be solved because $E_{\underline{y}|\underline{\theta}}(U)$ is a function of these unknown parameters and therefore 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 $\underline{X} \equiv (\underline{x}_1, \ldots, \underline{x}_T)'$ is a $(T \times k)$ matrix of T past realizations of $\underline{x}$. After the substitution, we have

$$\begin{aligned} \underline{d}_{PCE} &= \underset{\underline{d} \in D}{\operatorname{argmax}} E_{\underline{y}|\underline{\theta}=\hat{\underline{\theta}}}\left\{U\left[r\left(\underline{d}, \underline{y}\right)\right]\right\} \\ &= \underset{\underline{d} \in D}{\operatorname{argmax}} \int_Y U\left[r\left(\underline{d}, \underline{y}\right)\right] p\left[\underline{y}|\hat{\underline{\theta}}(\underline{X})\right] d\underline{y}. \quad (2) \end{aligned}$$

This technique is sometimes called the parameter certainty equivalent (PCE) method (Bawa, Brown, and Klein (1979)) to emphasize that parameters are taken as if they are 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 Bayes' decision criterion. By using Bayes' theorem, we have $p(\underline{\theta}|\underline{X}, I_T) \propto p(\underline{\theta}|I_T)p(\underline{X}|\underline{\theta})$, 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 $\underline{\theta}$, Bayes' decision criterion states that the optimal decision is made after integrating out the unknown parameters in

[1] For a detailed exposition of decision making under estimation risk, see Raiffa and Schlaifer (1961) or DeGroot (1970).

$E_{\underline{y}|\underline{\theta}}(U)$ by employing the posterior pdf. Bayes' criterion postulates that the optimum decision is given by

$$\underline{d}_{BAY} = \underset{\underline{d} \in D}{\operatorname{argmax}} E_{\underline{\theta}} \left\{ E_{\underline{y}|\underline{\theta}} \left[U\left(r\left(\underline{d}, \underline{y} \right) \right) \right] \right\} \tag{3}$$

$$= \underset{\underline{d} \in D}{\operatorname{argmax}} \int_{\Theta} \left\{ \int_{Y} U\left[r\left(\underline{d}, \underline{y} \right) \right] p\left(\underline{y} | \underline{\theta} \right) d\underline{y} \right\} p(\underline{\theta} | \underline{X}, I_T)\, d\underline{\theta}$$

$$= \underset{\underline{d} \in D}{\operatorname{argmax}} \int_{Y} U\left[r\left(\underline{d}, \underline{y} \right) \right] \left[\int_{\Theta} p\left(\underline{y} | \underline{\theta} \right) p(\underline{\theta} | \underline{X}, I_T)\, d\underline{\theta} \right] d\underline{y}$$

$$= \underset{\underline{d} \in D}{\operatorname{argmax}} \int_{Y} U\left[r\left(\underline{d}, \underline{y} \right) \right] p\left(\underline{y} | \underline{X}, I_T \right) d\underline{y}, \tag{3$'$}$$

where Θ is the domain of $\underline{\theta}$ and $p(\underline{y}|\underline{X}, I_T)$ is the *predictive* pdf of $\underline{y}$.[2] The most important aspects of (3′) compared with (1) are that (3′) does not involve any unknown parameter and that it comprises both prior (nonsample) and sample information.

As pointed out by Bawa, Brown, and Klein (1979), there are at least three reasons for employing Bayes' decision criterion (3) rather than the PCE (2). First, Bayes' criterion can be derived from the basic axioms advanced by von Neumann-Morgenstern and Savage, whereas the PCE approach has no axiomatic foundation (DeGroot (1970), chapters 7 and 8). Second, Bayes' method takes into consideration all the relevant (sample as well as nonsample) information about $\underline{\theta}$ through the posterior pdf (Berger (1985), chapters 1 and 4). In contrast, the PCE technique only uses the sample information contained in the point estimate $\hat{\underline{\theta}}$. Finally, Bayes' criterion leads to decisions that have minimum average risk (or maximum average value) for the specified prior (Berger (1985), chapter 4).

II. The Standard Model of the Optimum Futures Position

In its most popular form, the portfolio model of futures trading yields the following optimal futures position at the decision date T:[3]

$$F_{PPI} = \frac{(\mu_f - f_T)}{\lambda \sigma_{ff}} - \frac{\sigma_{pf}}{\sigma_{ff}} Q, \tag{4}$$

where F_{PPI} is the amount bought in the futures market at date T and sold at date $T + 1$, μ_f denotes the expectation of f_{T+1}, 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 $T + h$, λ is the decision maker's degree of absolute risk aversion, σ_{ff} is the variance of f_{T+1}, σ_{pf} is the

[2] The predictive pdf is obtained by using 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.

[3] See Anderson and Danthine (1980) for a thorough analysis of this model.

covariance between f_{T+1} and p_{T+1}, p_{T+1} is the random cash price at date $T+1$, and Q is the amount of product sold in the cash market at date $T+1$. According to expression (4), the optimum futures position is made of two components (Anderson and Danthine (1980)). The first is the *speculative* component and is represented by $(\mu_f - f_T)/(\lambda \sigma_{ff})$, whereas the second is the *hedge* component and is given by $(-\sigma_{pf}/\sigma_{ff} Q)$. The ratio σ_{pf}/σ_{ff} is usually called the minimum variance hedge ratio.

In empirical applications, the usual approach is to find the optimum futures position using

$$F_{PCE} = \frac{(\hat{\mu}_f - f_T)}{\lambda \hat{\sigma}_{ff}} - \frac{\hat{\sigma}_{pf}}{\hat{\sigma}_{ff}} Q, \tag{5}$$

where $\hat{\mu}_f$, $\hat{\sigma}_{ff}$, and $\hat{\sigma}_{pf}$ are the sample estimates of μ_f, σ_{ff}, and σ_{pf}, respectively. Despite its intuitive appeal, the trading rule described in (5) has the clear shortcomings of the PCE approach discussed in the previous section. In addition, there is no clear consensus regarding the values of the parameter estimates for any particular commodity or financial futures (e.g., Ederington (1979), Hill and Schneeweis (1981), Myers and Thompson (1989), Baillie and Myers (1991), Cecchetti, Cumby, and Figlewski (1988)), indicating that estimation risk is not negligible in the selection of the optimal futures position. Also, the PCE model (5) cannot handle the incorporation of nonsample information. However, it is difficult to rule out the existence of nonsample information such as insider information and expert opinions, especially given the findings by Stein (1986, chapter 5) and Rausser and Carter (1983).

III. The Optimum Futures Position in the Presence of Estimation Risk

In this section, we will derive the optimal futures position using Bayes' criterion (F_{BAY}). The key assumptions underlying the closed-form solution F_{PPI} in (4) are: (i) the utility function is negative exponential, (ii) the function $r(\underline{d}, \underline{y})$ is wealth at date $T+1$ (π_{T+1}), and (iii) cash and futures prices are jointly normally distributed with mean vector $\underline{\mu}$ and covariance matrix $\underline{\Sigma}$:[4]

$$p(\underline{y}|\underline{\theta}) = N_2(\underline{y}|\underline{\mu}, \underline{\Sigma}), \tag{6}$$

where $\underline{y} = (p_{T+1}, f_{T+1})'$ and $N_k(\cdot)$ is the k-variate normal pdf. To obtain a specific solution using Bayes' criterion, in addition to assumptions (i) through (iii), it is necessary to postulate the form of the prior pdf for the unknown true parameters. We will hypothesize a normal-Wishart prior for the true but

[4] The 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, the latter characterization is generally preferred because it is more straightforward and yields basically the same results.

unknown mean vector $\underline{\mu}$ and covariance matrix $\underline{\Sigma}$ because it is generally the most convenient prior in the case of a normal joint pdf:

$$p(\underline{\theta}|I_T) = NW_2\left(\underline{\mu}, \underline{\Sigma}^{-1}|\underline{\mu}_0, \tau, \underline{\Sigma}_0, \nu\right)$$
$$= N_2\left(\underline{\mu}|\underline{\mu}_0, \tau^{-1}\underline{\Sigma}\right)W_2\left[\underline{\Sigma}^{-1}\middle|(\nu\underline{\Sigma}_0)^{-1}, \nu\right], \qquad \tau > 0, \quad \nu > 1, \quad (7)$$

where $NW_k(\cdot)$ is the k-variate normal-Wishart pdf, $W_k(\cdot)$ is the k-variate Wishart pdf, $\underline{\mu}_0$ is the vector of prior means, $\underline{\Sigma}_0$ is the prior 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 are the decision maker's prior beliefs about $\underline{\mu}_0$ $(\underline{\Sigma}_0)$.[5]

Under assumptions (i) through (iii) and the normal-Wishart prior (7), expression (3′) for the hedging model can be shown to equal[6]

$$F_{BAY} = \underset{F}{\operatorname{argmax}}\left\{\int_{-\infty}^{+\infty} -\exp[-\lambda(\mu_\pi + \sigma_\pi t_m)]S(t_m)\,dt_m\right\}, \qquad (8)$$

where

$$\begin{aligned}
&\mu_\pi = \pi_T - c(Q) - f_T F + (Q, F)\underline{\mu}_T\\
&\sigma_\pi = [(Q, F)\underline{\Sigma}_T(Q, F)']^{1/2}\\
&c(Q) = \text{cost of producing (or storing) } Q\\
&\underline{\mu}_T = \omega_\tau\underline{\mu}_0 + (1 - \omega_\tau)\underline{\hat{\mu}}\\
&\underline{\Sigma}_T = [1 + 1/(\tau + T)](1 - \Delta/m)\\
&\cdot[\omega_\nu\underline{\Sigma}_0 + (1 - \omega_\nu)\underline{\hat{\Sigma}} + \omega_\tau T/(m - \Delta)(\underline{\hat{\mu}} - \underline{\mu}_0)(\underline{\hat{\mu}} - \underline{\mu}_0)']\\
&\omega_\tau = \tau/(\tau + T)\\
&\omega_\nu = \nu/(\nu + T - 1)\\
&m = \nu + T - 1 + \Delta\\
&\Delta = \begin{cases} 0 & \text{if } \tau = 0\\ 1 & \text{otherwise}\end{cases}\\
&\underline{\hat{\mu}} = \underline{\iota}'\underline{X}/T\\
&\underline{\hat{\Sigma}} = (\underline{X} - \underline{\iota}\underline{\hat{\mu}}')'(\underline{X} - \underline{\iota}\underline{\hat{\mu}}')/(T - 1)\\
&\underline{\iota} = \text{vector of ones of dimension } (T \times 1)
\end{aligned}$$

and $S(t_m)$ is the standardized univariate Student's t pdf with m degrees of freedom. The vector $\underline{\mu}_T$ comprises the posterior means and the covariance matrix $\underline{\Sigma}_T$ comprises the posterior variances and covariances. 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

[5] For example, τ and ν are both large if the agent possesses 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 $\underline{\Sigma}$.

[6] A bivariate normal-Wishart prior and a bivariate normal joint pdf yield a bivariate Student's t predictive pdf for p_{T+1} and f_{T+1} (Aitchison and Dunsmore (1975), chapter 2). But the predictive pdf for π_{T+1} is univariate Student's t because π_{T+1} is a linear combination of p_{T+1} and f_{T+1} (Zellner (1971), appendix B). Expression (8) is finally obtained by standardizing this univariate Student's t pdf for π_{T+1}.

degree of confidence in his or her prior beliefs about the means or when the sample size is small relative to the quality of the prior information, 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 covariance matrix the higher the confidence on $\underline{\Sigma}_0$ (represented by ν) compared to the sample size T.

The solution to (8) must be obtained by numerical integration because there is no closed-form expression for F_{BAY} under the stated assumptions.[7] When $m > 20$, however, it is possible to derive a very accurate closed-form approximation for F_{BAY}. For $m > 20$, the standardized univariate Student's t pdf with m degrees of freedom can be accurately approximated by the normal pdf with zero mean and variance $m/(m - 2)$ (Johnson and Kotz (1970), p. 101). By employing this result in (8), it is straightforward to derive a closed-form solution for F_{BAY} because the mean-variance framework can be applied to obtain

$$F_{BAY} \cong (1 - 2/m)\frac{(\mu_{fT} - f_T)}{\lambda\sigma_{ffT}} - \frac{\sigma_{pfT}}{\sigma_{ffT}}Q. \tag{9}$$

Although (9) is an approximation, it is illuminating for comparing F_{BAY} with the PPI and PCE solutions (4) and (5), respectively.

Consider first the speculative term. When the sample size is positive but the decision maker is much more confident about his or her prior beliefs than about the sample estimates of the futures mean and variance, we have $\omega_\tau \to 1$ and $\omega_\nu \to 1$ and the speculative term in (9) simplifies to

$$(1 - 2/m)\frac{(\mu_{fT} - f_T)}{\lambda\sigma_{ffT}} \to \frac{(\mu_{f0} - f_T)}{\lambda\sigma_{ff0}}. \tag{10}$$

Therefore, if the prior parameters are the true parameters (which is reasonable in this scenario of infinite confidence in the prior), the speculative term in F_{BAY} collapses to that in F_{PPI}.

In contrast, if the decision maker holds "diffuse" prior beliefs about the mean and the variance (i.e., $\tau \to 0$ and $\nu \to 0$), the speculative term in (9) becomes

$$(1 - 2/m)\frac{(\mu_{fT} - f_T)}{\lambda\sigma_{ffT}} \to \frac{T(T-3)}{(T^2-1)}\frac{(\hat{\mu}_f - f_T)}{\lambda\hat{\sigma}_{ff}}. \tag{11}$$

The right-hand side of (11) is different from the speculative term in F_{PCE} 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

[7] Strictly speaking, the integral in (8) does not exist because it amounts to the moment-generating function of the Student's t pdf (Kendall and Stuart (1969), p. 61); the value of this integral tends to minus infinity as $t_m \to -\infty$. For all practical purposes, however, this is inconsequential because an F_{BAY} can be found that maximizes the integral for any finite lower bound for t_m, and the F_{BAY}s thus obtained are quite insensitive to the choice of this lower bound.

get closer to 1 as the sample size grows. In other words, if the decision maker has diffuse prior beliefs, using the PCE method rather than Bayes' criterion will overstate the absolute value of the speculative futures position by the proportion $3/(T - 1/T)$. For example, $3/(T - 1/T)$ equals 6 percent with a sample size (T) of 50 observations. As the sample size tends to infinity, $3/(T - 1/T)$ tends to one and the speculative term in F_{BAY} equals that in F_{PCE}.

If the decision maker has nondiffuse prior beliefs but the sample size is so large that $\omega_\tau \to 0$ and $\omega_\nu \to 0$, the speculative term in F_{BAY} collapses to $(\hat{\mu}_f - f_T)/(\lambda \hat{\sigma}_{ff})$, which is the speculative term in F_{PCE}. Note, however, that for this to happen it is necessary to have an infinitely large sample size.

Similar comments can be made about the pure hedge term. As the confidence in the prior information tends to infinity, the posterior minimum variance hedge ratio $\sigma_{pfT}/\sigma_{ffT}$ tends to $\sigma_{pf0}/\sigma_{ff0}$, which is the PPI minimum variance hedge ratio (again, assuming that the prior parameters are the true parameters). In the opposite instance of diffuse prior information, the posterior ratio simplifies to the PCE ratio $\hat{\sigma}_{pf}/\hat{\sigma}_{ff}$.

A distinguishing feature of (9) 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 has nonsample information (or beliefs) that the futures price is positively biased and the sample mean indicates that the futures price is unbiased (i.e., $\mu_{f0} < f_T$ and $\hat{\mu}_f = f_T$, respectively). Then, the speculative term is negative according to (9), whereas it equals zero according to the PCE approach (5) and is not defined according to expression (4).

To summarize, the (approximate) optimum futures position in the presence of estimation risk given by expression (9) nests the two extreme scenarios of total lack of prior knowledge about the parameters (expression (5)) and perfect parameter information (expression (4)). The advantage of expression (9) over either (5) 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 (9) is useful because it can be directly applied to decision making. Another important application is for evaluating the robustness of the standard approach (5) 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 (1980)).

In Tables I and II, we show the results of some simulations regarding the futures positions obtained by means of Bayes' criterion (8), the PCE approach (5), and the PPI case (4) (the latter assuming that the priors equal the true parameters). In the interest of space, we only report the results for (8) since the solutions for (8) and (9) are almost identical, because $m \geq 50$. For the

Table I

Simulation Results for the Speculative Component of the Alternative Decision Criteria

The simulated values of the speculative term of Bayes' futures position were obtained by setting the cash position equal to zero ($Q = 0$) in the expression $F_{BAY} = \text{argmax}_F\{\int - \exp[-\lambda(\mu_\pi + \sigma_\pi t_m)]S(t_m)\,dt_m\}$, where λ is the coefficient of absolute risk aversion, μ_π (σ_π) is the posterior mean (standard deviation) of wealth, and $S(t_m)$ is the standardized univariate Student's t pdf with m degrees of freedom. The simulated values for the PCE and PPI speculative terms are calculated as $F_{PCE} = (\hat{\mu}_f - f_T)/(\lambda\hat{\sigma}_{ff})$ and $F_{PPI} = (\mu_{f0} - f_T)/(\lambda\sigma_{ff0})$, where $\hat{\mu}_f(\hat{\sigma}_{ff})$ is the sample mean (variance) of futures price, f_T is the current futures price, and μ_{f0} (σ_{ff0}) is the prior mean (variance) of futures price. It is assumed that $f_T = 10$ and that the decision maker has the same confidence in the prior information as in the sample information ($\omega_\tau = \omega_\nu = 0.5$).

Futures Mean		Futures Variance		Sample	Speculative Term Corresponding to		
Prior (μ_{f0})	Sample ($\hat{\mu}_f$)	Prior (σ_{ff0})	Sample ($\hat{\sigma}_{ff}$)	Size (T)	Bayes (F_{BAY})	PCE (F_{PCE})	PPI (F_{PPI})
12	12	4	4	25	$0.470/\lambda$	$0.5/\lambda$	$0.5/\lambda$
				100	$0.492/\lambda$	$0.5/\lambda$	$0.5/\lambda$
			3	25	$0.534/\lambda$	$0.667/\lambda$	$0.5/\lambda$
				100	$0.562/\lambda$	$0.667/\lambda$	$0.5/\lambda$
12	10	4	4	25	$0.190/\lambda$	0	$0.5/\lambda$
				100	$0.197/\lambda$	0	$0.5/\lambda$
			3	25	$0.210/\lambda$	0	$0.5/\lambda$
				100	$0.219/\lambda$	0	$0.5/\lambda$

scenarios reported in Table I, the solutions from (9) differ from those obtained by means of (8) by less than 2.6 percent when $T = 25$, and by less than 0.6 percent when $T = 100$. Similarly, the solutions from (9) in the scenarios of Table II differ by less than 0.01 percent from the solutions calculated by means of (8).

Both the speculative and the hedge components in Tables I and II are measured in units of physical commodity traded in the futures market. The means and variances used in the simulations imply a coefficient of variation for futures prices of around 18 percent and differences between prior and futures means (when any) of approximately one standard deviation of futures prices. The hypothesized sample sizes (25 and 100) are well within the range of sample sizes commonly used in empirical studies.

The speculative components from the simulations are reported in Table I. There are noticeable differences among the three models. The absolute magnitude of the speculative term in (8) 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. The most important determinant of the speculative component obtained by means of Bayes' decision criterion is the difference between the sample and prior futures means. For example, a real-world interpretation of the scenario presented in row 5 is that the sample information indicates that the futures markets are unbiased but the decision maker has insider information that predicts a 20 percent apprecia-

Table II

Simulation Results for the Hedging Component of the Alternative Decision Criteria

The simulated values of the hedging term of Bayes' futures position are obtained by setting the current futures price equal to the posterior futures mean ($f_T = \mu_{fT}$) in the expression $F_{BAY} = \text{argmax}_F\{\int -\exp[-\lambda(\mu_\pi + \sigma_\pi t_m)]S(t_m)\,dt_m\}$, where λ is the coefficient of absolute risk aversion, μ_π (σ_π) is the posterior mean (standard deviation) of wealth, and $S(t_m)$ is the standardized univariate Student's t pdf with m degrees of freedom. The simulated values for the PCE and PPI speculative terms are calculated as $F_{PCE} = \hat{\sigma}_{pf}/\hat{\sigma}_{ff}Q$ and $F_{PPI} = \sigma_{pf0}/\sigma_{ff0}Q$, where $\hat{\sigma}_{pf}$ (σ_{pf0}) is the sample (prior) covariance between cash and futures prices, $\hat{\sigma}_{ff}$ (σ_{ff0}) is the sample (prior) variance of futures price, and Q is the cash position. It is assumed that the prior and sample cash means are equal ($\mu_{p0} = \hat{\mu}_p$) and that the decision maker has the same confidence in the prior information as in the sample information ($\omega_\tau = \omega_\nu = 0.5$).

Futures Mean		Futures Variance		Covariance		Sample	Hedge Term Corresponding to		
Prior (μ_{f0})	Sample ($\hat{\mu}_f$)	Prior (σ_{ff0})	Sample ($\hat{\sigma}_{ff}$)	Prior (σ_{pf0})	Sample ($\hat{\sigma}_{pf}$)	Size (T)	Bayes (F_{BAY})	PCE (F_{PCE})	PPI (F_{PPI})
12	12	4	4	3.8	3.8	25	−0.950 Q	−0.950 Q	−0.95 Q
						100	−0.950 Q	−0.950 Q	−0.95 Q
					2.95	25	−0.844 Q	−0.738 Q	−0.95 Q
						100	−0.844 Q	−0.738 Q	−0.95 Q
			3	3.8	3.8	25	−1.086 Q	−1.267 Q	−0.95 Q
						100	−1.086 Q	−1.267 Q	−0.95 Q
					2.95	25	−0.964 Q	−0.983 Q	−0.95 Q
						100	−0.964 Q	−0.983 Q	−0.95 Q
12	10	4	4	3.8	3.8	25	−0.754 Q	−0.950 Q	−0.95 Q
						100	−0.758 Q	−0.950 Q	−0.95 Q
					2.95	25	−0.669 Q	−0.738 Q	−0.95 Q
						100	−0.674 Q	−0.738 Q	−0.95 Q
			3	3.8	3.8	25	−0.837 Q	−1.267 Q	−0.95 Q
						100	−0.843 Q	−1.267 Q	−0.95 Q
					2.95	25	−0.743 Q	−0.983 Q	−0.95 Q
						100	−0.748 Q	−0.983 Q	−0.95 Q

tion in the futures price. Under PCE, the decision maker ignores the private information arguing that it is impossible to beat the efficient market. Under PPI, he or she takes the private information as perfect and takes a futures position that would be optimal only if there were no sample information available. The Bayes method indicates a moderate and more intuitive response.

It is interesting to note that the speculative term in F_{BAY} is different from that in F_{PCE} even if prior and sample means and variances are the same. The reason for this result is that the posterior variance is larger than the weighted average of the prior and sample variances by the proportion $[1 + 1/(\tau + T)](1 - \Delta/m)$; this proportion will tend to one as the sample size tends to infinity.

Table II summarizes the simulation results regarding the hedge components. Here, the sample size exerts a negligible effect on the Bayes' hedge term. In contrast to the speculative component, the hedge terms in F_{BAY} and F_{PCE} are the same when all prior and sample parameters are equal. This result occurs because both the posterior variance and the posterior covariance are multiplied by the proportion $[1 + 1/(\tau + T)](1 - \Delta/m)$, so that the multiplicative effect cancels out. Again, consider row 5. Here the prior and sample means and covariances agree but the sample variance is smaller than the prior variance. Under PPI, the individual uses the prior variance and underhedges; in contrast, under PCE, the individual overhedges by 26.7 percent. Again, Bayes' criterion provides a more moderate and intuitive response.

Another observation from Table II is that the hedge components in F_{BAY} and F_{PCE} may be different even if the sample and prior variances are equal and the sample and prior covariances are equal. This difference will occur whenever the prior futures mean is different from the sample futures mean. This result is attributable to the additional variability stemming from the difference between sample and prior means, which is reflected as a larger posterior variance (see the last component in the expression for $\underline{\Sigma}_T$ in (8)).

The simulations reported in Tables I and II reveal that the futures positions resulting from the three alternative approaches can be quite different, suggesting that the monetary value of the alternative rules can also be quite large. Indeed, simulations not reported here show that this observation holds true when the monetary values of the three alternative futures positions are measured by their certainty equivalent returns.

IV. Conclusions

The standard procedure used to estimate the optimum futures position is the PCE method. This method consists of directly substituting the sample estimates of the mean, variance, and covariance for the true but 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 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 or her prior 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 advanced can be used to optimally blend both types of information. The paper contains simulations that demonstrate the sensitivity of the optimum futures position to the method that is used. Given that Bayes' criterion is the method that maximizes expected utility, the noticeable differences in the optimal futures position imply a large monetary value to investors using the proposed method.

REFERENCES

Aitchison, J., and I. R. Dunsmore, 1975, *Statistical Prediction Analysis* (Cambridge University Press, New York).

Alexander, Gordon J., and Bruce G. Resnick, 1985, More on estimation risk and simple rules for optimal portfolio selection, *Journal of Finance* 40, 125–133.

Anderson, Ronald W., and Jean-Pierre Danthine, 1980, Hedging and joint production: Theory and illustrations, *Journal of Finance* 35, 487–498.

Baillie, Richard T., and Robert J. Myers, 1991, Bivariate GARCH estimation of the optimal commodity futures hedge, *Journal of Applied Econometrics* 6, 109–124.

Barry, Christopher B., and Stephen J. Brown, 1985, Differential information and security market equilibrium, *Journal of Financial and Quantitative Analysis* 20, 407–422.

Bawa, V., S. Brown, and R. Klein, eds., 1979, *Estimation Risk and Optimal Portfolio Choice* (North-Holland Publishing Co., Amsterdam).

Berger, James O., 1985, *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. (Springer-Verlag, New York).

Boyle, Phelim P., and A. L. Ananthanarayanan, 1977, The impact of variance estimation in option valuation models, *Journal of Financial Economics* 5, 375–387.

Cecchetti, S., R. Cumby, and S. Figlewski, 1988, Estimation of optimal futures hedge, *Review of Economics and Statistics* 70, 623–630.

Chen, Son-Nan, and Stephen J. Brown, 1983, Estimation risk and simple rules for optimal portfolio selection, *Journal of Finance* 38, 1087–1093.

Cheung, C. Sherman, and Clarence C. Y. Kwan, 1988, A note on simple criteria for optimal portfolio selection, *Journal of Finance* 43, 241–245.

Coles, Jeffrey L., and Uri Loewenstein, 1988, Equilibrium pricing and portfolio composition in the presence of uncertain parameters, *Journal of Financial Economics* 22, 279–303.

——— and Jose Suay, 1991, On equilibrium pricing under parameter uncertainty, Unpublished manuscript, School of Business, University of Utah, Salt Lake City.

Cox, John, and Mark Rubinstein, 1985, *Options Markets* (Prentice Hall, Englewood Cliffs, N.J.).

DeGroot, Morris H., 1970, *Optimal Statistical Decisions* (McGraw-Hill, New York).

Ederington, Louis H., 1979, The hedging performance of the new futures markets, *Journal of Finance* 34, 57–70.

Frost, Peter A., and James E. Savarino, 1986, An empirical Bayes approach to efficient portfolio selection, *Journal of Financial and Quantitative Analysis* 21, 293–305.

Hill, Joanne, and Thomas Schneeweis, 1981, A note on the hedging effectiveness of foreign currency futures, *Journal of Futures Markets* 1, 659–664.

Johnson, Leland L., 1960, The theory of hedging and speculation in commodity futures, *Review of Economic Studies* 27, 139–151.

Johnson, Norman L., and Samuel Kotz, 1970, *Continuous Univariate Distributions*, Vol. 2 (John Wiley and Sons, New York).

Jorion, Philippe, 1985, International portfolio diversification with estimation risk, *Journal of Business* 58, 259–278.

Kendall, Maurice G., and Alan Stuart, 1969, *The Advanced Theory of Statistics*, Vol. 1: Distribution Theory, 3rd ed. (Hafner Publishing Co., New York).

Marshall, J., 1989, *Futures and Option Contracting* (South-Western Publishing Co., Cincinnati).

Myers, Robert J., and Stanley R. Thompson, 1989, Generalized optimal hedge ratio estimation, *American Journal of Agricultural Economics* 71, 858–868.

Raiffa, Howard, and Robert Schlaifer, 1961, *Applied Statistical Decision Theory* (Harvard University Press, Boston).

Rausser, Gordon C., and Colin Carter, 1983, Futures market efficiency in the soybean complex, *Review of Economics and Statistics* 65, 469–478.

Stein, Jerome L., 1961, The simultaneous determination of spot and futures prices, *American Economic Review* 60, 1012–1025.

———, 1986, *The Economics of Futures Markets* (Basil Blackwell, Oxford).

Winkler, Robert L., 1980, Prior information, predictive distributions, and Bayesian model-building, in Arnold Zellner, ed.: *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys* (North-Holland Publishing Co., Amsterdam).

Zellner, Arnold, 1971, *An Introduction to Bayesian Inference in Econometrics* (John Wiley and Sons, New York).