NCCC-134
APPLIED COMMODITY PRICE ANALYSIS, FORECASTING AND MARKET RISK MANAGEMENT

The Relative Performance of In-Sample and Out-of-Sample Hedging Effectiveness Indicators

by

Roger A. Dahlgran

Suggested citation format:

Dahlgran., R. A. 2009. "The Relative Performance of In-Sample and Out-of-Sample Hedging Effectiveness Indicators." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.uiuc.edu/nccc134].

The Relative Performance of In-Sample and Out-of-Sample Hedging Effectiveness Indicators

Roger A. Dahlgran [*]

Paper presented at the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management
St. Louis, Missouri, April 20-21, 2009

Copyright 2008 by Roger Dahlgran. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

[*] Roger Dahlgran (dahlgran@u.arizona.edu) is an Associate Professor in the Department of Agricultural and Resource Economics, 403C Chavez Bldg, University of Arizona, Tucson, Arizona 85721-0023.

The Relative Performance of In-Sample and Out-of-Sample Hedging Effectiveness Indicators

Practitioner's Abstract

Hedging effectiveness is the proportion of price risk removed through hedging. Empirical hedging studies typically estimate a set of risk minimizing hedge ratios, estimate the hedging effectiveness statistic, apply the estimated hedge ratios to a second group of data, and examine the robustness of the hedging strategy by comparing the hedging effectiveness for this "out-of-sample" period to the "in-sample" period. This study focuses on the statistical properties of the in-sample and out-of-sample hedging effectiveness estimators. Through mathematical and simulation analysis we determine the following: (a) the R^2 for the hedge ratio regression will generally overstate the amount of price risk reduction that can be achieved by hedging, (b) the properly computed hedging effectiveness in the hedge ratio regression will also generally overstate the true amount of risk reduction that can be achieved, (c) hedging effectiveness estimated in the out-of-sample period will generally understate the true amount of risk reduction that can be achieved, (d) for equal numbers of observations, the overstatement in (b) is less that the understatement in (c), (e) both errors decline as more observations are used, and (f) the most accurate approach is to use all of the available data to estimate the hedge ratio and effectiveness and to not hold any data back for hedge strategy validation. If structural change in the hedge ratio model is suspected, tests for parameter equality have a better statistical foundation that do tests of hedging effectiveness equality.

Keywords: out-of-sample, post sample, hedging, effectiveness, forecasts, simulation.

Introduction

Hedging studies typically proceed by posing a price risk minimization problem for a specific commodity, collecting data, and estimating hedge ratios with regression analysis. The regression R square is reported as the proportion of price risk eliminated by hedging. To estimate the price risk reduction expected from future hedging, these studies then apply the estimated hedge ratios to out-of-sample data and compare the variance of unhedged outcomes to the variance of hedged outcomes. This paper focuses on the last step.

Applying estimated hedge ratios to nonsample data is intuitively appealing. Claims about the robustness of a particular hedging strategy have merit as do claims that the effectiveness of estimated hedge ratios applied to nonsample data constitute a forecast of the effectiveness that can generally be expected from the hedging strategy. But, given the prevalence of this practice, its fundamental assumptions merit closer scrutiny.

For example, the comparison of in-sample and out-of-sample hedging effectiveness implicitly assumes that in-sample effectiveness is a biased estimator of out-of-sample results. While the bias of the in-sample estimator remains to be seen, a comparison of a single observation for in-sample and out-of-sample effectiveness is not sufficient to determine either the presence of, or the magnitude and direction of bias.

Second, the comparison of in-sample and out-of-sample hedging effectiveness fails to address the notion that both are random variables, each with its own variance. Differences are to be expected. The precision of each estimate is more telling than the magnitude of their difference as the comparison gives no indication of whether the difference is significant.

Third, the notion of the robustness of the estimated hedging strategy is tied to the assumption that the cash-futures price relationship did not change between the in sample and out-of-sample periods. While such a structural change would render the estimated hedging strategy less effective under the new regime and hence less robust, this notion is better tested by re-estimating the regression over the in-sample and out-of-sample periods and testing for parameter equality over both periods. This parameter equality test is girded with better understood statistical theory than is a test for effectiveness equality.

Finally, this procedure of applying hedge ratios to out-of-sample data does not address the optimal allocation of data to the estimation period and the out-of-sample period. With a fixed number of data points, using more of the data for estimation improves the precision of the estimated hedge ratios and estimated effectiveness but reduces the precision of the out-of-sample effectiveness forecast because fewer observations are available to do this. As data are scarce the optimal allocation between the in-sample and out-of-sample periods should be considered.

The objectives of this paper are to examine the distributional properties of the hedging effectiveness statistic. In particular we will explore whether in-sample hedging effectiveness is an unbiased estimator for out-of-sample results and how sample size influences the bias and precision of the effectiveness estimators. This study will utilize simulation analysis in which thousands of random samples of various sizes are drawn. For each sample, we will compute the hedge ratio and the corresponding hedging effectiveness. We also draw random samples to which the estimated hedge ratios are applied so that we can examine out-of-sample hedging effectiveness

Theoretical Background

Hedging behavior assumes that an agent seeks to minimize the price risk of holding a necessary spot (or cash) market position by taking an attendant futures market position (Johnson, Stein). The profit outcome (π) of these combined positions is

(1) $$\pi = x_s (p_1 - p_0) + x_f (f_1 - f_0),$$

where x_s is the agent's necessary cash market position, p is the commodity's cash price, x_f is the agent's discretionary futures market position, f is the futures contract's price, and subscripts 1 and 0 refer to points in time. Risk is minimized by selecting the x_f (x_f^*) that minimizes the variance of π ($V(\pi)$) giving

$$x_f^*/x_s = -\sigma_{\Delta p,\Delta s} / \sigma^2_{\Delta f}.$$

This risk minimizing hedge ratio (x_f^*/x_s) is estimated by the least-squares estimator $\hat{\beta}$ in the regression

(2) $\Delta^H p_t = \alpha + \beta \Delta^H f_{Mt} + \varepsilon_t, t = 1, 2, \ldots T$

where, in addition to the previous definitions, f_{Mt} represents the M-maturity futures contract's price at time t, Δ^H represents differencing over the hedging interval[1], ε_t represents stochastic error at time t, and T represents the number of observations. The risk minimizing futures position is $x_f^* = -\hat{\beta} x_s$.

Anderson and Danthine (1980, 1981) generalized this approach to accommodate positions in multiple futures contracts. In this case, x_f and $(f_1 - f_0)$ in (1) are replaced by vectors of length k and hedge ratio estimation involves fitting the multiple regression model

(3) $\Delta p_t = \alpha + \sum_{j=1}^{k} \beta_j \Delta f_{jt} + \varepsilon_t$, t = 1, 2, 3, … T,

where Δf_{jt} is the change in the price of futures contract j over the hedge period, and $\hat{\beta}_j$ is the estimated hedge ratio indicating the number of units in futures contract j per unit of spot position.

Other generalizations of this model include applications to soybean processing (Dahlgran, 2005; Fackler and McNew; Garcia, Roh, and Leuthold; and Tzang and Leuthold), cattle feeding (Schafer, Griffin and Johnson), hog feeding (Kenyon and Clay), and cottonseed crushing (Dahlgran, 2005; Rahman, Turner, and Costa). In this studies, the profit objective is

$$\pi = y\, p_{y,1} - x\, p_{x,0} + x_f (f_1 - f_0)$$

where inputs (x) and outputs (y) are connected by product transformation,

$$y = \kappa x.$$

Hedge ratio estimation for this model involves fitting the regression

(4) $p_{y,t} - \kappa p_{x,t-H} = \alpha + \sum_{j=1}^{k} \beta_j \Delta f_{jt} + \varepsilon_t$, t = 1, 2, 3, … T.

The hedge ratio regressions in (2), (3), and (4) can all be represented by the general regression model $\mathbf{Y = X\beta + \varepsilon}$, with T observations and K (= k+1) explanatory variables in $\mathbf{X}$. $\boldsymbol{\beta}$ is estimated with $\hat{\boldsymbol{\beta}} = (\mathbf{X'X})^{-1}\mathbf{X'Y}$.

Ederington defined hedging effectiveness (e) as the proportion of price risk eliminated by hedging. More specifically,

(5) $e = [\, V(\pi_u) - V(\pi_h) \,] / V(\pi_u)$

[1] H refers to the length of the assumed hedging period. Henceforth, Δ^H will be represented more succinctly with Δ where H is assumed.

where V is the variance operator, π_u the agent's unhedged outcome and π_h is the agent's hedged outcome. The regression R^2 serves generally as an estimator of the coefficient of determination and serves in hedge ratio estimation as an estimator of hedging effectiveness.

The coefficient of determination is defined by Marchand as follows. Let **[Y : X] = [Y : X_1, X_2, …,X_k]** be distributed as a k+1-variate normal with covariance matrix Σ and let S be the covariance matrix obtained from a sample of size T where T > k > 1. Partition Σ and S as $\Sigma = \begin{bmatrix} \sigma_{YY} & \sigma_{YX} \\ \sigma_{XY} & \Sigma_{XX} \end{bmatrix}$ $S = \begin{bmatrix} S_{YY} & S_{YX} \\ S_{XY} & S_{XX} \end{bmatrix}$ where σ_{YY} and S_{YY} are scalars. The multiple correlation coefficient between **Y** and **[X_1, X_2,…,X_k]** is defined as $\rho = (\sigma_{YY}^{-1}\sigma_{YX}\Sigma_{XX}^{-1}\sigma_{XY})^{1/2}$ and ρ^2 is the coefficient of determination. The analogous sample quantities are $R = (S_{YY}^{-1}S_{YX}S_{XX}^{-1}S_{XY})^{1/2}$ and R^2. Marchand goes on to state (p. 173) "It is well known that, on average, R^2 overestimates ρ^2." Consequently, R^2 is a biased estimator of ρ^2, $E(R^2) > \rho^2$, and as $T \to \infty$, $E(R^2) = \rho^2$. Thus, the in-sample hedging effectiveness estimator overstates the true value of hedging effectiveness but this bias diminishes as sample size increases. Determining the magnitude of the bias requires the probability distribution of the effectiveness statistic.

The distribution of R^2 can be derived from the distribution of the regression F statistic. Specifically, for regressions (2), (3), or (4)

$$(6) \qquad F = \frac{SSR/k}{SSE/T-k-1} = \frac{R^2}{1-R^2}\left(\frac{T-k-1}{k}\right)$$

where F is the regression F-statistic, SSR is the regression sum of squares and SSE is the error sum of squares. While the regression F statistic is used to test whether the noncentrality parameter of the numerator chi square is zero, [i.e., $\lambda = (\boldsymbol{\beta}'\mathbf{X}\mathbf{X}'\boldsymbol{\beta} - \boldsymbol{\mu}'\boldsymbol{\mu})/\sigma^2$ if $\boldsymbol{\beta} = \mathbf{0}$ in (2), (3) or (4)] this assumption negates the hedging motive. Consequently, we assume that F in (6) follows a noncentral F distribution so

$$(7a) \qquad \Pr\left\{ F_{n_2}^{n_1,\lambda} < f_{n_2}^{n_1,\lambda}(\alpha) \right\} = \alpha ,$$

where $F_{n_2}^{n_1,\lambda}$ is the noncentral F random variable with n_1 (numerator) and n_2 (denominator) degrees of freedom, and noncentrality parameter $\lambda = (\boldsymbol{\beta}'\mathbf{X}\mathbf{X}'\boldsymbol{\beta} - T\,E(Y)^2)/\sigma^2$, and $f(\alpha)$ is the numerical value for which the probability of a smaller value of the F random variable is α. The corresponding cumulative probability distribution for R^2 is

$$(7b) \qquad \Pr\left\{ \frac{R^2}{1-R^2}\left(\frac{T-k-1}{k}\right) < f_{T-k-1}^{k,\lambda}(\alpha) \right\} = \Pr\left\{ R^2 < \frac{k\, f_{T-k-1}^{k,\lambda}(\alpha)}{(T-k-1) + k\, f_{T-k-1}^{k,\lambda}(\alpha)} \right\} = \alpha .$$

Chattamvelli clarifies this result. *"If $\chi_{n_1}^2$ and $\chi_{n_2}^2$ are independent central chi squared random variables with n_1 and n_2 degrees of freedom, then $F = (\chi_{n_1}^2/n_1)/(\chi_{n_2}^2/n_2)$ has an F distribution*

and $B=n_1 F / (n_2 + n_1 F) = \chi^2_{n_1} / (\chi^2_{n_1} + \chi^2_{n_2})$ has a beta distribution. When both of the χ^2 are noncentral, F has a doubly noncentral F distribution. When only one of the χ^2 is noncentral, F has a (singly) noncentral F distribution. Analogous definitions hold for the beta case."

As (6) is composed of the requisite independent chi square random variables, the regression R^2 follows a singly noncentral beta distribution with n_1= k = K-1 and n_2 = T-K = T-k-1 degrees of freedom, and λ as defined above. The numeric values of the beta random variable corresponding to probability α are computable from the noncentral f values in the second form of the probability statement in (7b).

Pe and Drygas (p. 313) provide an expression for finding the moments of the doubly noncentral beta. They state "if X_1 and X_2 are independently distributed as noncentral χ^2 with n_i degrees of freedom and noncentrality parameters λ_i (i = 1, 2), then $Z = X_1 / (X_1+X_2)$ is distributed independently from X_1+X_2 as a doubly noncentral β_1 distribution with parameters $n_1/2$, $n_2/2$, and λ_1, λ_2 respectively" then the r^{th} moment about the origin is

(8a) $$E(Z^r) = e^{-\frac{1}{2}(\lambda_1+\lambda_2)} \sum_{i=0}^{\infty} \frac{\left(\lambda_1/2\right)^i}{i!} \sum_{j=0}^{\infty} \left(\lambda_2/\lambda_1\right)^j \binom{i}{j} \frac{(\frac{n_1}{2}+i+j)_{r-2j}}{(\frac{n_1}{2}-j)_{2j}(i+\frac{n_1+n_2}{2})_r},$$

where $(\theta)_k = \frac{\Gamma(\theta+k)}{\Gamma(\theta)}$ and $\Gamma(\theta) = (\theta\text{-}1)\,\Gamma(\theta\text{-}1) = (\theta\text{-}1)!$ for integer θ and $\Gamma(1/2) = \sqrt{\pi}$ if θ is half integer. When applied to the statement for R^2 in (7b), λ_2=0, n_1=k, and n_2=T-k-1 so (8a) reduces to

(8b) $$E(R^{2r}) = e^{-\lambda/2} \sum_{i=0}^{\infty} \frac{\left(\lambda/2\right)^i}{i!} \frac{(\frac{k}{2}+i)_r}{(\frac{k}{2})_0 (i+\frac{T-1}{2})_r}.$$

This expression allows us to derive the mean (r = 1) as well as the variance of R^2.

The previous discussion applies to the regression R^2, and while we next argue that the regression R^2 is an incomplete estimator of hedging effectiveness, this previous discussion provides a useful background. First, hedging effectiveness is defined more explicitly as

(9) $$e = \frac{E\{[\Delta p_t - E(\Delta p_t)]^2\} - E\{[\Delta p_t - \hat{\beta}\Delta f_{Mt} - E(\Delta p_t - \hat{\beta}\Delta f_{Mt})]^2\}}{E\{[\Delta p_t - E(\Delta p_t)]^2\}}.$$

This definition establishes that the variances apply to differences between actual and expected outcomes or more simply, that the variances are for unanticipated outcomes.

If the hedge ratio regression displays systematic behavior such as seasonality or serial correlation, then hedging effectiveness must be defined so that these systematic components are part of the expected outcome, whether or not hedging occurs. To represent this, the hedge ratio regression in (2), (3), or (4) is expressed as

$$\mathbf{Y} = \mathbf{X_1}\boldsymbol{\beta}_1 + \mathbf{X_2}\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$$

where the K columns of $\mathbf{X}$ have been partitioned into k_1 deterministic and known components contained in $\mathbf{X_1}$, and k_2 stochastic components contained in $\mathbf{X_2}$. In addition to the column of ones for the intercept, $\mathbf{X_1}$ might also contain dummy variables or estimated lagged errors which account for serial correlation. $\mathbf{X_2}$ contains the futures price changes and other random components. Because the elements of $\mathbf{X_1}$ are systematic, they form anticipations so the unanticipated outcome is

$$\mathbf{y} \equiv \mathbf{Y} - \mathbf{X_1}\boldsymbol{\beta}_1 = \mathbf{X_2}\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}.$$

By Marchand's definition the unknown hedging effectiveness parameter is

$$\eta = \sigma_{yy}^{-1}\sigma_{yX_2}\Sigma_{X_2X_2}^{-1}\sigma_{X_2y},$$

and the in-sample hedging effectiveness estimator is

$$(10) \qquad e_1 = \frac{(\mathbf{Y} - \mathbf{X_1}\hat{\boldsymbol{\beta}}_1)'(\mathbf{Y} - \mathbf{X_1}\hat{\boldsymbol{\beta}}_1) - (\mathbf{Y} - \mathbf{X_1}\hat{\boldsymbol{\beta}}_1 - \mathbf{X_2}\hat{\boldsymbol{\beta}}_2)'(\mathbf{Y} - \mathbf{X_1}\hat{\boldsymbol{\beta}}_1 - \mathbf{X_2}\hat{\boldsymbol{\beta}}_2)}{(\mathbf{Y} - \mathbf{X_1}\hat{\boldsymbol{\beta}}_1)'(\mathbf{Y} - \mathbf{X_1}\hat{\boldsymbol{\beta}}_1)}$$

$$= \frac{\mathbf{Y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{X_1}(\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}']\mathbf{Y}}{\mathbf{Y}'[\mathbf{I} - \mathbf{X_1}(\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}']\mathbf{Y}}$$

This expression demonstrates that the regression R^2 and e_1 are the same only when $\mathbf{X_1}$ consists solely of a column of ones. Otherwise, e_1 has the characteristics of the regression R^2 in that it is bounded by zero and one but the regression R^2 overstates hedging effectiveness by assigning too many degrees of freedom to the numerator (i.e., K-1 instead of K-k_1), thereby overstating the numerator sum of squares, and assigning too few degrees of freedom to the denominator (i.e., T-1 instead of T-k_1), thereby understating the denominator sum of squares. So in addition to R^2 being an upwardly biased estimator for ρ^2 it also overstates hedging effectiveness.

The statistical properties of the hedging effectiveness estimator follow from analysis of variance definitions. Let SST = $\mathbf{Y}'\mathbf{Y}$, SSR($\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$) = $\hat{\mathbf{Y}}'\hat{\mathbf{Y}} = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$, and SSE is the sum of squared errors (i.e., SSE = $\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} = \mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}$) so SST = SSR($\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$) + SSE. Searle (p. 247) shows (a) that SSR($\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$) = SSR($\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1$) + SSR($\boldsymbol{\beta}_1$), where SSR($\boldsymbol{\beta}_1$) = $\mathbf{Y}'\mathbf{X_1}(\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}\mathbf{Y}$, (b) that SSR($\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1$) = $\mathbf{Y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} - \mathbf{X_1}(\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}']\mathbf{Y}$, and (c) that SSR($\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1$) / σ^2 has a noncentral χ^2 distribution and is independent of both SSR($\boldsymbol{\beta}_1$) and SSE.

Applying these definitions establishes that

$$(11a) \qquad F = \frac{SSR(\hat{\boldsymbol{\beta}}_2 \mid \hat{\boldsymbol{\beta}}_1)/k_2}{SSE/(T-K)} = \frac{\mathbf{Y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{X_1}(\mathbf{X_1}'\mathbf{X_1})^{-1}\mathbf{X_1}']\mathbf{Y}/(K-k_1)}{\mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}/(T-K)}$$

is distributed as a noncentral F random variable with k_2 numerator and T-K denominator degrees of freedom, and $\lambda = [\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} - \mathbf{E}(\mathbf{Y}'\mathbf{X}_1)\,\mathbf{E}(\mathbf{X}_1'\mathbf{X}_1)^{-1}\,\mathbf{E}(\mathbf{X}_1'\mathbf{Y})] \,/\, \sigma^2$, and

$$\text{(11b)} \quad e_1 = \frac{SSR(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1)}{SSE + SSR(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1)} = \frac{SSR(\boldsymbol{\beta}_2 \mid \boldsymbol{\beta}_1)}{SST - SSR(\boldsymbol{\beta}_1)}$$

is distributed as a singly noncentral beta random variable with degrees of freedom and λ corresponding to (11a).

The cumulative probability distribution of e_1 is derived from the noncentral F random variable in (11a). First, dividing both the numerator and denominator of (11a) by $\mathbf{Y}'[\mathbf{I}-\mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1]\mathbf{Y}$ expresses (11a) in terms of e_1 as

$$\text{(12a)} \quad F = \left(\frac{e_1}{1-e_1}\right)\left(\frac{T-K}{k_2}\right)$$

so that the probability statement

$$\text{(12b)} \quad \Pr\left\{\left(\frac{e_1}{1-e_1}\right)\left(\frac{T-K}{k_2}\right) < f_{T-K}^{k_2,\lambda}(\alpha)\right\} = \Pr\left\{e_1 < \frac{k_2\, f_{T-K}^{k_2,\lambda}(\alpha)}{(T-K) + k_2\, f_{T-K}^{k_2,\lambda}(\alpha)}\right\} = \alpha$$

defines the cumulative probability distribution for the in-sample hedging effectiveness estimator.

Methods

Simulation analysis is used to explore (a) the relationship between the regression R^2 and e_1, (b) the impact of sample size on the precision of e_1, and (c) bias and efficiency tradeoffs between in-sample and out-of-sample effectiveness estimates. In this analysis, 10,000 samples of size T will be drawn for the model

$$\text{(13)} \quad \Delta p_t = \alpha + \delta D_t + \beta \Delta f_t + \gamma (\Delta p_{t-1} - \alpha - \delta D_{t-1} - \beta \Delta f_{t-1}) + \varepsilon_t,$$

where, in addition to the terms defined in (2), D_t represents a generic dummy variable ($D_t = 1$ if t even, 0 otherwise), $\varepsilon_t \sim \text{NID}(0, \sigma_\varepsilon^2)$, and $(\Delta p_{t-1} - \alpha - \delta D_{t-1} - \beta \Delta f_{t-1})$ represents first order autoregressive effects when $\gamma \neq 0$. This model encompasses features common to hedge ratio regressions such as (a) an intercept to account for a long-term spot price trend, (b) seasonal spot price variation that is accounted for by the dummy variable, and (c) serial correlation due to non-instantaneous spot price equilibration. Estimation of the parameters in (13) requires the inversion of a 3x3 matrix $\mathbf{X}'\mathbf{X}$ which, as opposed to more comprehensive models, is not computationally prohibitive. This consideration is especially important in light of the number of samples drawn and the potential correction for autocorrelation.

Simulated data are generated subject to assumptions regarding (a) the structural parameters α, β, δ, and γ, (b) the variables Δf_t and ε_t where $\Delta f_t \sim N(0, \sigma_{\Delta f}^2)$, $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$,

(c) ϕ, the correlation between Δf_t and ε_t, and (d) T, the size of each sample. While arbitrary values are selected for the parameters, the results are illustrative.

Once parameter values for α, δ, β, γ, $\sigma_{\Delta f}^2$, σ_ε^2, and ϕ are selected, a sample of size T is generated by (13) and the parameters and hedging effectiveness (e_1) are estimated. Then another sample of size T is drawn and the estimated parameters are applied to these data and the out-of-sample hedging effectiveness (e_2) is computed from the variances of the unhedged and hedged outcomes (i.e., $e_2 = 1 - \frac{\hat{V}(\pi_h)}{\hat{V}(\pi_u)}$). This process is repeated 10,000 times for each sample size. The estimates from each sample are used to form the empirical cumulative distribution functions (CDFs) for the regression R^2, the in-sample hedging effectiveness estimates (e_1), and out-of-sample hedging effectiveness estimates (e_2). The empirical CDFs are compared to the theoretical CDFs specified by (7b) and (12b). Also, because the population parameters are known, R^2 can be compared to ρ^2, and in-sample and out-of-sample effectiveness estimates (e_1 and e_2) can be compared to η. The sampling distributions for R^2, e_1, and e_2 are reported via cumulative probability plots.

Parameter values of $\alpha = 0$, $\delta = -2$, $\beta = 1$ and $\sigma_{\varepsilon\varepsilon} = 1$ are assumed throughout. $\alpha = 0$ implies no trend in the spot price. $\beta = 1$ represents a direct hedging application. $\delta = -2$ gives sufficient magnitude to the dummy variable to illustrate the distinction between R^2 and hedging effectiveness. $\sigma_{\varepsilon\varepsilon} = 1$ makes ε_t a unit normal random variable. Initially, we assume that $\gamma = 0$ to eliminate serial correlation.

Beyond this, four alternative sets of assumptions are considered. Case 1 adopts the assumptions of the standard regression model as Δf_t is assumed fixed in repeated samples. Under this condition the distributions of the resulting statistics are known and serve as a benchmark for the empirical distributions of the other simulations. In case 2, we assume futures price changes are generated by the stochastic process $\Delta f_t \sim N(0, \sigma_{xx})$ where $\sigma_{xx} = 1$. In case 3 assume serial correlation ($\gamma \neq 0$). This condition frequently occurs in empirical work when the time differences are short (1 to 2 weeks). When $\gamma = 0$, case 3 is identical to case 2. To maintain comparability across cases and values of γ, η remains set at ½. As α, δ, β $\sigma_{\varepsilon\varepsilon}$ are also set to predetermined values, σ_{xx} must adjust.

Table 1 shows how ρ^2 and η depend on our parameters.

Table 1. Parametric values of ρ^2 and η.[a]

Case	ρ^2	η	Other
1. X fixed	$\dfrac{\frac{\delta^2}{4}-\frac{\delta\beta a}{2}+\beta^2}{\frac{\delta^2}{4}-\frac{\delta\beta a}{2}+\beta^2+\sigma_{\varepsilon\varepsilon}}$	$\dfrac{\beta^2\left(1-\frac{a^2}{4}\right)}{\beta^2\left(1-\frac{a^2}{4}\right)+\sigma_{\varepsilon\varepsilon}}$	$a=\sqrt{\dfrac{12}{T^2-1}}$
2. X random	$\dfrac{\frac{\delta^2}{4}+\beta^2\sigma_{xx}}{\frac{\delta^2}{4}+\beta^2\sigma_{xx}+\sigma_{\varepsilon\varepsilon}}$	$\dfrac{\beta^2\sigma_{xx}}{\beta^2\sigma_{xx}+\sigma_{\varepsilon\varepsilon}}$	
3. Serial correlation $\gamma \neq 0$	$\dfrac{\frac{\delta^2}{4}+\beta^2\sigma_{xx}}{\frac{\delta^2}{4}+\beta^2\sigma_{xx}+\frac{\sigma_{\varepsilon\varepsilon}}{1+\gamma^2}}$	$\dfrac{\beta^2\sigma_{xx}}{\beta^2\sigma_{xx}+\frac{\sigma_{\varepsilon\varepsilon}}{1+\gamma^2}}$	$\sigma_{xx}=\dfrac{\sigma_{\varepsilon\varepsilon}}{\beta^2(1+\gamma^2)}\left(\dfrac{\eta}{1-\eta}\right)$

a/ $\rho^2=\Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}-T^{-1}E(Y^{'}X_1)[E(X_1^{'}X_1)]^{-1}E(X_1^{'}Y)$ where $\mathbf{X_1}=(1-\gamma)\,\mathbf{1}$, while $\eta=\Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy}-T^{-1}E(Y^{'}X_1)[E(X_1^{'}X_1)]^{-1}E(X_1^{'}Y)$ where $\mathbf{X_1}=[(1-\gamma)\,\mathbf{1},\ \mathbf{D}-\gamma\,\Delta\mathbf{D}_L]$

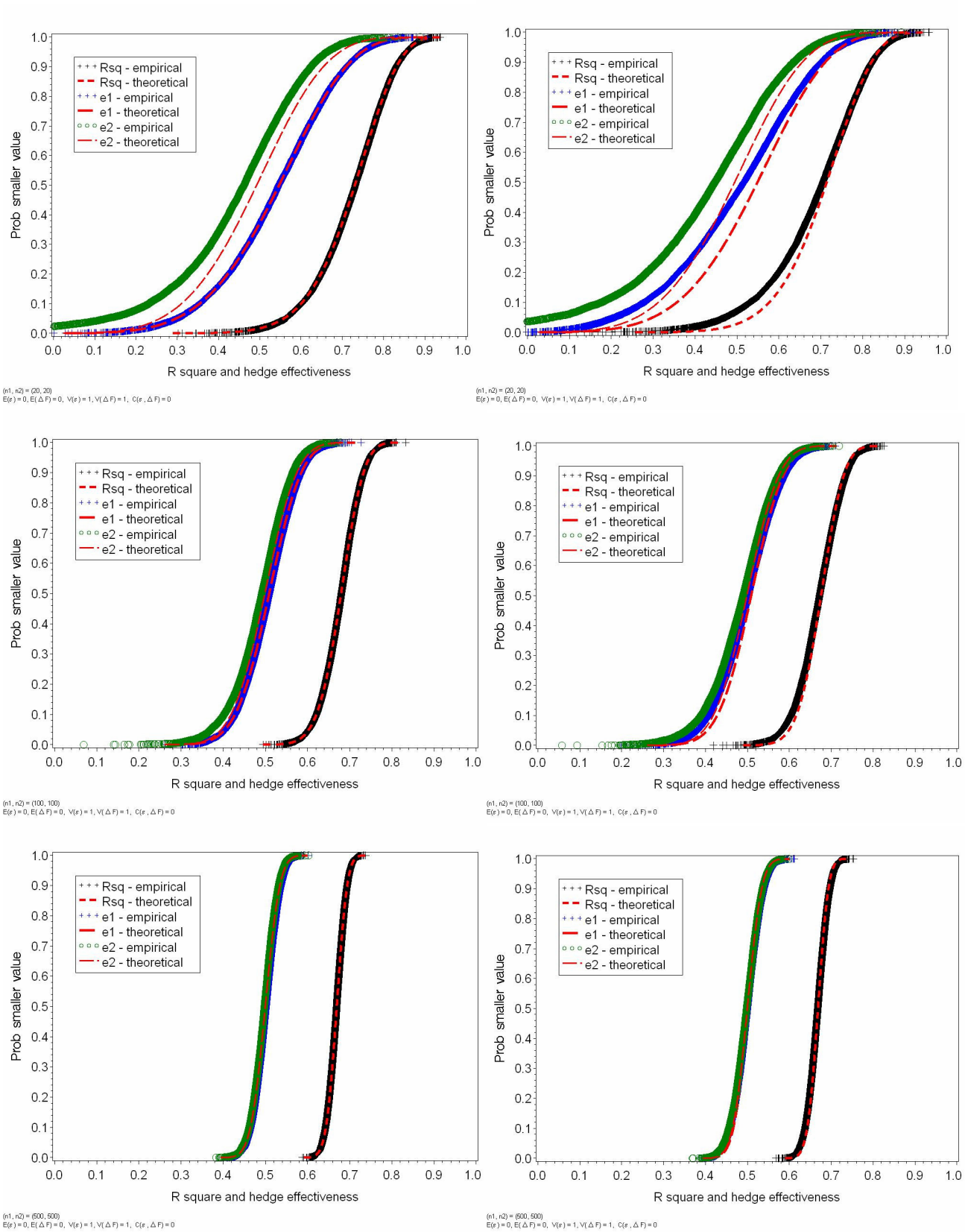


Figure 1. Theoretical and empirical cumulative distribution functions for fixed X.

Results

Case 1 - Fixed regressors: This simulation establishes a benchmark against which the mathematical results presented above can be compared. The standard linear regression model assumes non-stochastic explanatory variables (i.e., fixed in repeated samples) and the distributional results that underpin hypothesis testing and inference follow from the fixed-regressor assumption. For this case, the elements of $X_t = [\ 1 : D_t\ ;\ \Delta f_t\ ]$ are assigned as $D_t = 1$ if t even, 0 otherwise, and Δf_t is evenly spaced with $\Sigma\ \Delta f_t = 0$ and $T^{-1}\ \Sigma\ \Delta f_t^2 = 1$.[2] Samples of size (T) 40, 100, and 250 were selected to represent small medium and large samples in the context of sample sizes used in hedging studies.

By Marchand's definition $\rho^2 = \sigma_{yy}^{-1}\Sigma_{yx}\Sigma_{xx}^{-1}\Sigma_{xy} = \dfrac{E[Y'X(X'X)^{-1}X'Y - Y'X_1(X_1'X_1)^{-1}X_1'Y]}{E(Y'Y) - E[Y'X_1(X_1'X_1)^{-1}X_1'Y]}$. If $\mathbf{X_1 = 1}$, then we get $\rho^2 = \dfrac{\delta^2/4 + \beta^2 - \delta\beta a/2}{\sigma_{\varepsilon\varepsilon} + \delta^2/4 + \beta^2 - \delta\beta a/2}$, whereas if $\mathbf{X_1 = [\ 1, D\ ]}$, then we get $\eta = \dfrac{\beta^2(4-a^2)/4}{\sigma_{\varepsilon\varepsilon} + \beta^2(4-a^2)/4}$ where $a = \sqrt{12/(T^2-1)}$. Because of *a* both ρ^2 and η to depend on the sample size (T). Under our parametric assumptions, ρ^2 converges to ⅔ while η converges to ½. Thus, $\rho^2 \neq \eta$.

The cumulative distribution functions (CDFs) of R^2 and hedging effectiveness are shown in the left-hand panels of figure 1 for samples of size 40 (top), 100 (middle) and 250 (lower) observations, respectively. Each panel shows the empirical CDF for the regression R^2, in-sample effectiveness, and out-of-sample effectiveness. Each panel also shows the theoretical CDFs for R^2 and in-sample effectiveness as defined by (7b) and (12b), respectively. The simulated and theoretical CDFs match for R^2 and in-sample effectiveness but not for out-of-sample effectiveness. We also see that the theoretical and simulated CDFs for R^2 lies to the right of the CDFs for in-sample effectiveness. This corresponds to the result above that ρ^2 converges to ⅔ while η converges to ½. Both of these sets of CDFs lie to the right of the out-of-sample effectiveness distribution. As we move down the graphs in figure 1 the sample size increases and the out-of-sample effectiveness CDF converges to the in-sample effectiveness CDF.

Figure 1 also reveals that while in-sample effectiveness is always between zero and one, it is possible that application of the estimated hedge ratios out-of-sample might result in greater price risk. This occurs approximately ten percent of the time when hedge ratios are estimated from a sample of size 10 and occurred occasionally (though rarely) with samples of size 50. This illustrates that hedging can increase price risk if the parameter estimates differ substantially from

---

[2] For case 1, Δf_t is represented by an arithmetic sequence. Let i = 1,2, 3, … N. Then $\sum_{i=1}^{N} i = \dfrac{N(N+1)}{2}$ and $\sum_{i=1}^{N} i^2 = \dfrac{N(N+1)(2N+1)}{6}$. The sum is centered with $\sum_{i=1}^{N}(i-c) = 0$ so c = (N+1)/2 and standardized as $\sum_{i=1}^{N}[a(i-c)]^2 = N$ so a = $[12/(N^2-1)]^{1/2}$

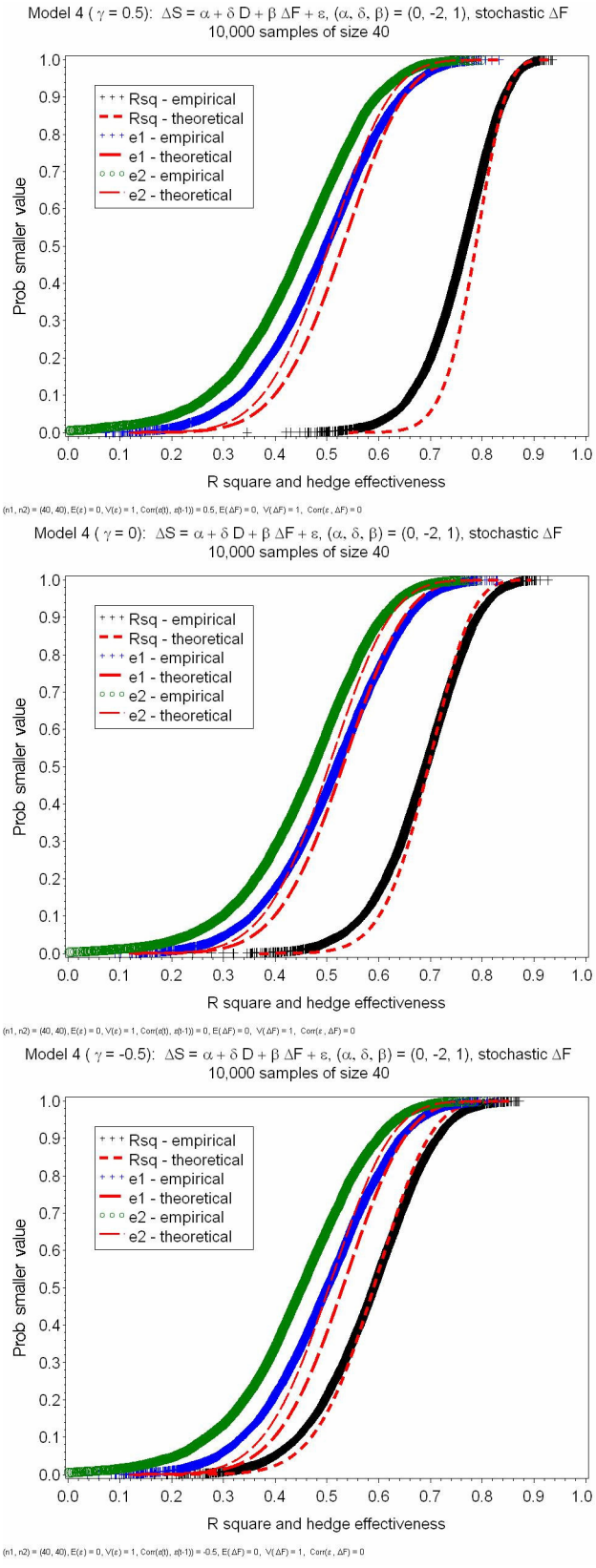


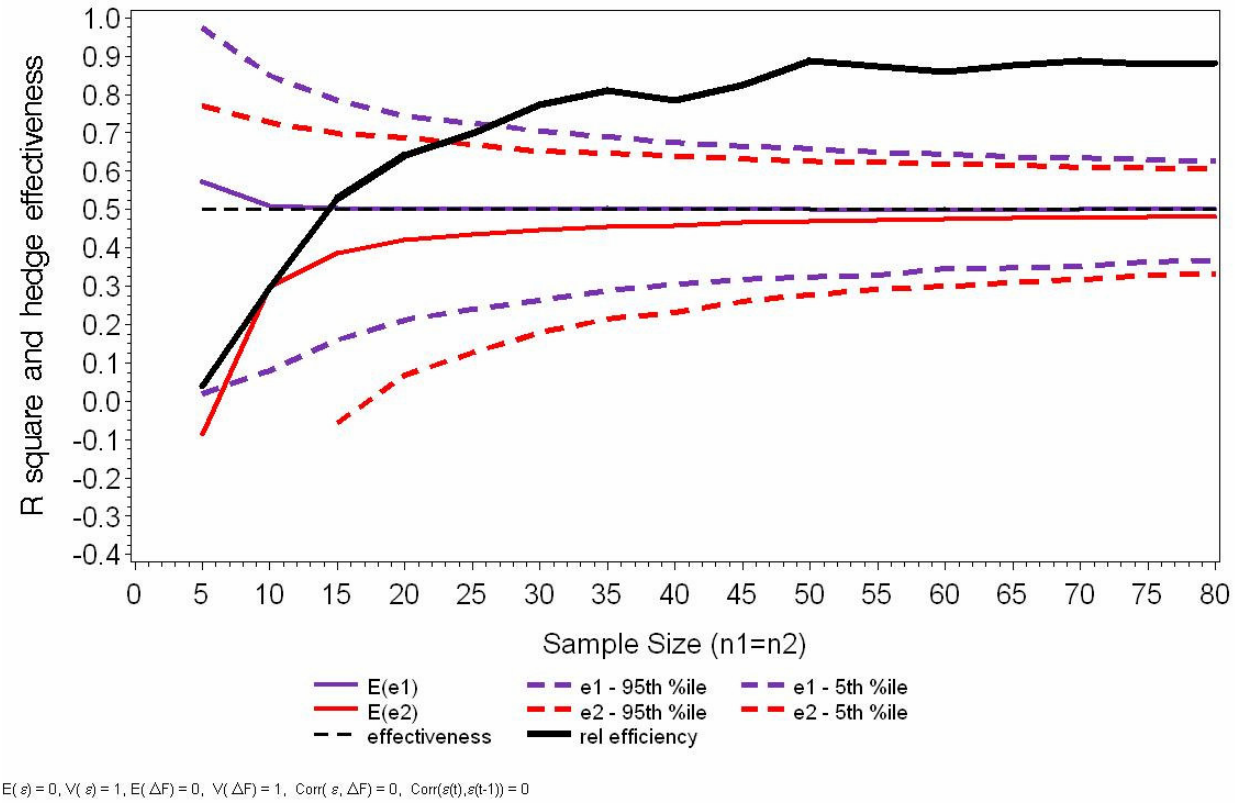
Figure 2. Theoretical and empirical cumulative distribution functions for serially correlated errors.

the true parameter values. When these unusual estimates were applied out-of-sample, the variance of hedged outcomes exceeds the variance of unhedged outcomes. As a result hedging effectiveness is negative indicating that hedging has increased price risk. The unit upper bound on out-of-sample effectiveness remains as the removal of all price variation through hedging is the best that can be achieved.

Case 2 - Stochastic regressors: The near exact match between the theoretical and empirical CDFs under the fixed-regressor assumption of the previous case validates our computational methods. When we replace the fixed-regressor assumption with the assumption that $\Delta f_t \sim N(0, 1)$, we observe that the non-centrality parameter λ takes random values depending on $\mathbf{\Delta f}$. To illustrate the theoretical CDFs shown in the right-hand panels of figure 1, we have used the expected value of λ instead of its random value for each sample. Thus, the theoretical CDFs in the left and right panels of figure1 are identical. We observe in the right-hand panels of figure 1 that the empirical CDFs do not match the theoretical CDFs when the regressors are stochastic. The difference is less pronounced for larger sample sizes but even with 250 observations, the difference is apparent. While the out-of-sample effectiveness is inadequately represented by a beta distribution, the difference diminishes as the sample size increases. From these results we see that attempts to draw inferences about the true value of hedging effectiveness will require appeal to large sample properties of our estimator.

Case 3 – Serial correlation: Figure 2 expands on our consideration of stochastic regressors and adds the assumption of serial correlation. R^2, η, and λ depend on σ_{xx} and $\sigma_{\varepsilon\varepsilon}$ (table 1) and $\sigma_{\varepsilon\varepsilon}$ has already been assumed to be one. To continue with comparisons of the CDFs across cases and values of γ, we maintain $\eta = ½$. With η fixed and $\sigma_{xx} = \frac{\sigma_{\varepsilon\varepsilon}}{\beta^2(1+\gamma^2)}\left(\frac{\eta}{1-\eta}\right)$ we see that as γ increases in absolute value, σ_{xx} must decrease. This causes the R^2 to decline as γ increases (figure2).

We also know that as T increases, the precision of our η and ρ^2 estimators also increase, so figure 2 is used to illustrate the effect of serial correlation. Figure 2 shows simulation results for $\gamma = 0.5$, 0, and -0.5 when T = 40. We observe the following. First the ordering of the CDFs is the same as in the previous two cases with the CDF for R^2 lying to the right of the CDF for e_1 which lies to the right of the CDF for e_2. Second, we observe that as γ increases, the theoretical and empirical CDFs for R^2 shift to the left.



E(ε) = 0, V(ε) = 1, E(ΔF) = 0, V(ΔF) = 1, Corr(ε, ΔF) = 0, Corr(ε(t),ε(t-1)) = 0

Figure 3. Variances of in-sample and out-of-sample hedging effectiveness by sample size.

We next focus on the relationship between effectiveness (η) and its in-sample estimator (e_1) and its out-of-sample estimator (e_2). Figure 3 shows the relationship between sample size and the dispersion of e_1 and e_2 under the assumptions of case 2. This figure demonstrates the following: First, the expected values of e_1 and e_2 bracket the true value of η (½). The upward bias of the in-sample effectiveness estimator [$E(e_1) - ½$] is apparent only for small samples and quickly disappears. Contrary to this, the downward bias of the out-of-sample estimator is larger and persists over larger sample sizes. Overall we see that the expected value of both estimators depend on the sample size and for both the bias approaches zero as the sample size increases.

The 5% and 95% probability bands indicate the precision of the two estimators. Here we see that as N increases the variances of both e_1 and e_2 decrease. Together with the disappearing biases we conclude that e_1 and e_2 are both consistent estimators for η. Finally, figure 2 plots the relative efficiency of e_1 versus e_2 [i.e., $Var(e_1) / Var(e_2)$]. This ratio is less than 1 and approaches one from below (figure3). Hence we conclude that if equal observations are allocated to the estimation of e_1 and e_2, e_1 will have the smaller variance and bias.

Figure 3 also reveals a more subtle but important relationship. With a fixed number of observations, using more for estimating the hedge ratio and computing e_1 means fewer are available for computing e_2. So while the bias and variance of e_1 are driven down as more observations are used for hedge ratio estimation, the bias and variance of e_2 are driven in the opposite direction as fewer observations are left over. We might also imagine that because e_1 is biased upward and e_2 is biased downward, an average of e_1 and e_2 might outperform either.

Our last simulation explores these issues. It proceeds as follows. Suppose you have T observations. Let the proportion of the observations allocated to the estimation of the hedge ratio and e_1 be p so that (1-p) T observation are left for the estimation of e_2. Through repeated sampling, we can compute the mean square error [$MSE(\hat{\eta}) = Var(\hat{\eta}) + Bias(\hat{\eta})^2$] for each estimator [$e_1$, e_2, and $\bar{e} = (e_1 + e_2)/2$] and pick the estimator with the smallest MSE.

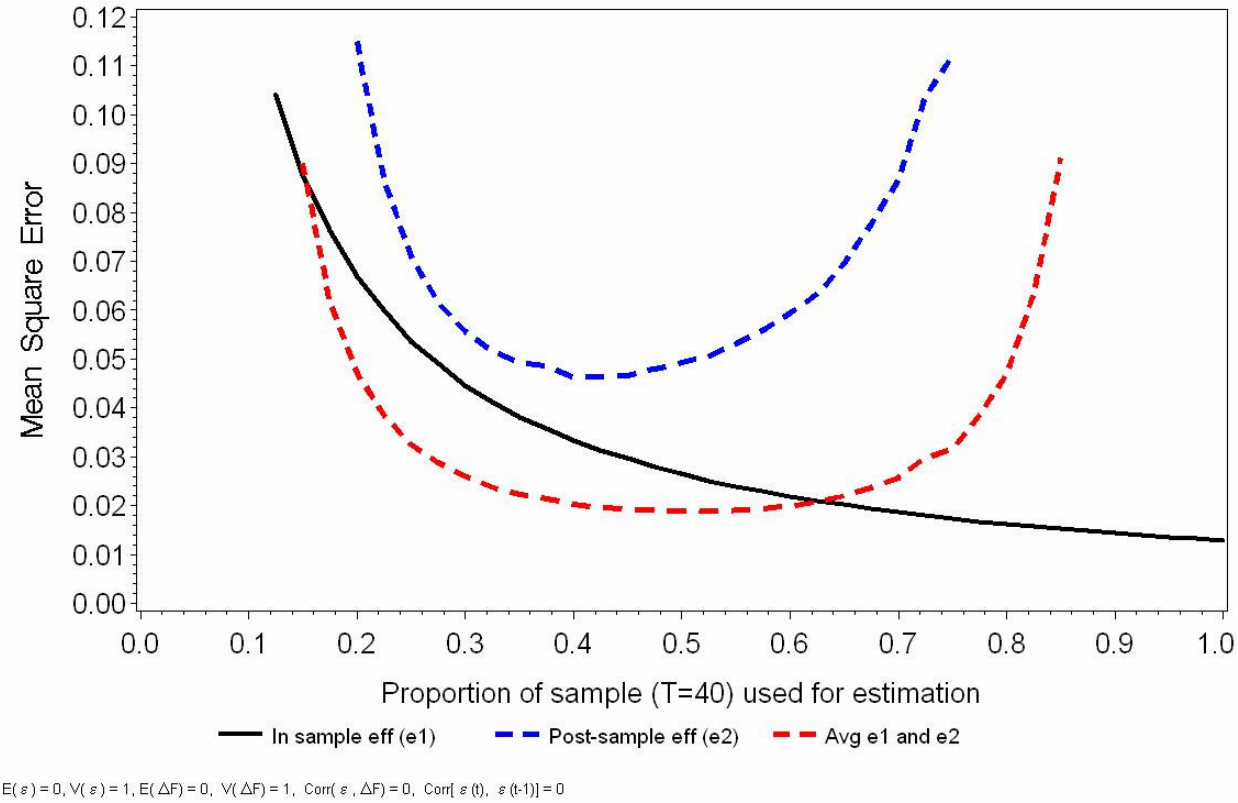


Figure 4. MSE of e_1, e_2 and $\bar{e}$ by allocation of sample to hedge ratio estimation, T=40.

Figure 4 presents the findings of this analysis assuming 40 observations. We see that the mean square error of e_1 declines continuously as more observations are used in the estimation of the hedge ratio and the computation of in-sample effectiveness. It is minimized (0.01286) when all 40 observations are allocated to its estimation. Figure 4 also indicates that the MSE of e_2 reaches a minimum (0.04630) when 40% of the observations are used to estimate the model parameters which are then applied out of sample to compute e_2. Finally, because $E(e_1)$ and $E(e_2)$ bracket the true value of η, the mean square error of $\bar{e}$ is also plotted in figure 4. The MSE of this estimator reaches a minimum (0.01884) when 52.5% of the sample is used for model estimation and the computation of e_1 and the remaining 47.5% of the sample is used for the computation of e_2. By comparing the minimized MSEs we conclude that the minimum MSE estimator is e_1 with the entire sample allocated to its computation.

The imprecision of the graphical analysis of figure 4 and the myriad of sample sizes to be examined require that the graphical approach of figure 4 be abandoned but the results depicted are summarized in table 2. These results indicate that for all sample sizes the minimum MSE for e_1 occurs when all available observations are used in its computation. The $\bar{e}$ estimator is

minimized when roughly half of the observations are used to compute e_1 and the other half used to compute e_2. This conclusion also holds for all sample sizes considered. Finally, table 3 indicates that the e_2 estimator achieves a minimum MSE when less than half of the observations are used in the computation of the hedge ratio. The optimal split of the observations between those used for the computation of the hedge ratio and those used for the computation of e_2 is such that the larger the sample, the smaller the proportion used for hedge ratio estimation. The most noteworthy feature of table 2 is that the minimum MSE estimator is e_1 when all observations are allocated to its computation regardless of the sample size. Thus, the practice of holding back observations for the estimation of out-of-sample effectiveness increases the MSE of e_1, and the computed e_2 has a higher MSE than the e_1 estimator that e_2 is being compared to.

Table 2. Allocation of observations in estimating effectiveness in-sample effectiveness, out-of-sample, average of in-sample and out-of-sample effectiveness.[a]

Estimator[a]	e_1		e_2		$\bar{e}$	
Number of Obs	Minimized MSE($\times 10^{-2}$)	Sample Allocation[b]	Minimized MSE($\times 10^{-2}$)	Sample Allocation[b]	Minimized MSE($\times 10^{-2}$)	Sample Allocation[b]
40	1.286	100%	4.630	40%	1.884	52%
100	0.511	100%	1.111	36%	0.581	49%
250	0.201	100%	0.340	28%	0.212	48%
500	0.100	100%	0.150	20%	0.103	49%

a/ e_1 is computed from the same observations as used to estimate the model parameters. e_2 uses the estimated model parameters to estimate effectiveness out-of-sample. $\bar{e}$ is the sample average of e_1 and e_2.

b/ Percentage of the sample observations used to estimate model parameters and compute e_1. 100 minus this percentage is used to compute e_2.

Conclusions

This paper reports on a continuing investigation into forecasting how a hedging strategy will perform. Nonetheless, we have established some definitive conclusions. First, we have established that the R^2 for the hedge ratio regression is an incomplete measure of hedging effectiveness and is appropriate only when the spot price displays neither systematic effects nor serial correlation. When spot prices are characterized by seasonality, serial correlation, day of the week effects, or relationships with other conditioning variables such as inventory levels or planted acreage, these systematic effects should be modeled as part of the hedge regression and hedging effectiveness should not include these variables' effect on the spot price.

Second, we have established that even after accounting for these systematic effects, the expected value of in-sample hedging effectiveness exceeds expected value of out-of-sample effectiveness. This occurs because of upward bias in measured hedging effectiveness due to the selection of parameter estimates that maximize R^2 and its subcomponent e_1. This maximization does not apply when the estimated hedge ratios are applied out of sample, so the expected in-sample effectiveness will exceed the expected out-of-sample effectiveness. In addition, because e_2 is computed as $1 - \frac{\hat{V}(\pi_h)}{\hat{V}(\pi_u)}$, e_2 is skewed to the left. In cases where there is little variation in π_u the e_2 statistic can take large negative values. This is contrary to e_1 which is never negative.

We have also established that the expected difference between the in-sample effectiveness estimator and out-of-sample effectiveness estimator diminishes with larger sample sizes. In light of this finding, we expect that the standard practice of fitting a hedging strategy to out-of-sample data will result in a lower out-of-sample effectiveness than measured in-sample. This result should not be interpreted as a lack of robustness of the hedging strategy or that structural change has occurred but instead that the in-sample effectiveness estimator naturally overstates what will be experienced out of sample. We have seen that the lower out-of-sample effectiveness estimate is a biased estimate of true hedging effectiveness.

Related to issues of robustness of our estimated hedging strategy and/or structural change, we have argued that a comparison of in-sample and out-of-sample hedging effectiveness is not the best way to test for these conditions. A procedure that is better grounded in statistics and probability is to test for parameter equality across the sample period and the out-of-sample period. A rejection of the hypothesis of parameter equality means that the hedging strategy estimated in the in-sample period is not appropriate for the out-of-sample period.

We also have shown that for a common sample size, the out-of-sample estimator is more variable than the in-sample estimator and that the variance of both estimators fall as the number of observations increase. When a given sample must be allocated to either the in-sample estimate or the out-of-sample estimate the minimum MSE estimator is e_1 with all available data used for its estimation.

This study leaves many issues unaddressed giving rise to further questions. For example, many model specifications have not been examined in enough detail. In particular,

a. Suppose the hedge ratio regression displays serial correlation. Does this affect our conclusion that the minimum MSE estimator is e_1 using all available data?

b. We considered only three estimators, e_1, e_2, and $\bar{e}$ (the simple average of e_1 and e_2 regardless of the sample allocation). Does there exist another estimator, $\tilde{e}$, that is a weighted average of e_1 and e_2, where the weights are not ½ and ½ that has a lower MSE than e_1?

c. Can the estimated parameters be used to estimate λ, and $\hat{\lambda}$ used in conjunction with the known numerator and denominator degrees of freedom to specify a beta distribution. Given

the estimated beta distribution, can we form an interval that correctly predicts the frequency of future hedging effectiveness?

d. And finally, only one specification with $\eta = ½$ was studied. This level of risk reduction would be considered low under many direct hedging applications while in cross hedging applications it might be something that managers could only hope for. Are our conclusions affected by alternative values of λ?

These topics will be studied as the scope of this paper is increased.

References

Anderson, R.W., and J. Danthine. 1980. "Hedging and Joint Production: Theory and Illustrations." *Journal of Finance* 35:487-501.

Anderson, R.W., and J. Danthine. 1981. "Cross Hedging." *Journal of Political Economy* 89:1182-1196.

Chattamvelli, R. "On the doubly noncentral F distribution." Computational Statistics and Data Analysis 20(1995):481-489.

Dahlgran, R.A. 2005. "Transaction Frequency and Hedging in Commodity Processing." *Journal of Agricultural and Resource Economics* 30:411-430.

Dahlgran, R.A. 2000. "Cross-Hedging the Cottonseed Crush: A Case Study." *Agribusiness* 16:141-158.

Ederington, L.H. 1979. "The Hedging Performance of the New Futures Markets." *Journal of Finance* 34:157-170.

Fackler, P.L., and K.P. McNew. 1993. "Multiproduct Hedging: Theory, Estimation, and an Application." *Review of Agricultural Economics* 15:521-535.

Garcia, P., J.S. Roh, and R.M. Leuthold. 1995. "Simultaneously Determined, Time-Varying Hedge Ratios in the Soybean Complex." *Applied Economics* 27:1127-1134.

Johnson, L.L. 1960. "The Theory of Hedging and Speculation in Commodity Futures." *Review of Economic Studies* 27:131-159.

Kenyon, D., and J. Clay. "Analysis of Profit Margin Hedging Strategies for Hog Producers." *Journal of Futures Markets* 7(1987):183-202.

Marchand, E.. "On moments of beta mixtures, the noncentral beta distribution and the coefficient of determination." Journal of Statistical Computation and Simulation 59(1997):161-178.

Pe, T., and H. Drygas. "An alternative representation of the noncentral beta and F distributions." Statistical papers 47(2006):311-318.

Rahman, S.M., S.C. Turner, and E.F. Costa. 2001. "Cross-Hedging Cottonseed Meal." *Journal of Agribusiness* 19:163-171.

Searle, S.R. Linear Models. John Wiley and Sons, Inc. New York. 1971.

Shafer, C.E., W.L Griffin, and L.D. Johnson "Integrated Cattle Feeding Hedging Strategies, 1972-1976." *Southern Journal of Agricultural Economics* 10(1978):35-42.

Tzang, D., and R.M. Leuthold. 1990. "Hedge Ratios under Inherent Risk Reduction in a Commodity Complex." *Journal of Futures Markets* 10:497-504.

Stein, J.J. 1961. "The Simultaneous Determination of Spot and Futures Prices." *American Economic Review* 51:1012-1025.