Hedge Effectiveness Forecasting

by

Roger A. Dahlgran and Xudong Ma

Hedge Effectiveness Forecasting

Roger A. Dahlgran, and

Xudong Ma[*]

_____

[*] Roger Dahlgran (dahlgran@u.arizona.edu) and Xudong Ma (max@email.arizona.edu) are
Associate Professor and Graduate Research Assistant, respectively, in the Department of
Agricultural and Resource Economics, 403C Chavez Bldg, University of Arizona, Tucson,
Arizona 85721-0023.

Hedge Effectiveness Forecasting

This study focuses on hedging effectiveness defined as the proportionate price risk reduction created by hedging. By mathematical and simulation analysis we determine the following: (a) the regression R^2 in 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 amount of risk reduction that can be achieved by hedging, (c) the overstatement in (b) declines as the sample size increases, (d) application of estimated hedge ratios to non sample data results in an unbiased estimate of hedging effectiveness, (e) application of hedge ratios computed from small samples presents a significant chance of actually increasing price risk by hedging, and (f) comparison of in sample and out of sample hedging effectiveness is not the best method for testing for structural change in the hedge ratio regression.

Introduction

Hedging studies typically proceed by posing a price risk minimization problem, 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 last step is the focus of this paper.

Applying estimated hedge ratios to non sample 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. However, the fundamental assumptions of this procedure need closer scrutiny.

For example, the comparison of in sample and out of sample hedging effectiveness implicitly assumes that the in sample effectiveness estimator is unbiased for out of sample results. While unbiasedness remains to be seen, a comparison of single observations for in sample and out of sample effectiveness is insufficient to determine either the absence 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 measures are random variables, each with its own variance, and differences are to be expected. The precision of each of the estimates is more telling than the magnitude of their difference as the comparison gives no indication of when 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

1

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 hedge ratios but reduces the precision of the out of sample forecast. 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 and influences effectiveness bias and precision. 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 $\hat{\beta}$ 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 (possibly with serial correlation) at time t, and T represents the number of observations. The risk minimizing futures position is $x_f^* = -\hat{\beta} \, x_s$.

---

[1]   All price changes occur during the assumed hedging period. Henceforth, Δ^H will be represented more succinctly with Δ where H is assumed.

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) $\qquad \Delta p_t = \alpha + \sum_{j=1}^{k} \beta_j \Delta f_{jt} + \varepsilon_t$, 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 case, the profit objective is

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

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

$\qquad y = \gamma \, x$.

Hedge ratio estimation for this model involves fitting the regression

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

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

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

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

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 as an estimator for e as well as the coefficient of determination.

Marchand defines the coefficient of determination as follows. Let $[\, \mathbf{Y} : \mathbf{X} \,] = [\, \mathbf{Y} : \mathbf{X_1, X_2, …, X_k} \,]$ 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

$$\Sigma = \begin{bmatrix} \sigma_{YY} & \sigma_{YX} \\ \sigma_{XY} & \Sigma_{XX} \end{bmatrix} \qquad S = \begin{bmatrix} S_{YY} & S_{YX} \\ S_{XY} & S_{XX} \end{bmatrix}$$

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.

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

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

While the regression F statistic is used to test whether the noncentrality parameter of the numerator chi square random variable is zero, (i.e., whether **β = 0** in (2), (3) or (4)) this assumption negates the hedging motive. Consequently, we recognize the noncentrality parameter and assume a noncentral F distribution for (6) so that

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

where F is the noncentral F random variable with n_1 (numerator) and n_2 (denominator) degrees of freedom, noncentrality parameter λ, and $f(\alpha)$ is the numerical value for which the probability of a smaller value of F is α. The corresponding cumulative probability distribution for R^2 is

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

where $\lambda = (\boldsymbol{\beta'XX'\beta} - T\overline{Y}^2)/\sigma^2$.

Chattamvelli provides an alternative approach.

> *"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) = $\chi_{n_1}^2$ /($\chi_{n_1}^2$ + $\chi_{n_2}^2$) 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 $\lambda = (\boldsymbol{\beta'XX'\beta} - T\overline{Y}^2)/\sigma^2$. The values of the beta random variable are apparent in the second form of the probability statement in (7b).

Pe and Drygas (p. 313) 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 rth moment about the origin is

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

where $(\theta)_k = \dfrac{\Gamma(\theta + k)}{\Gamma(\theta)}$ and $\Gamma(n) = (n-1)\,\Gamma(n-1) = (n-1)!$ for integer n and $\Gamma(1/2) = \sqrt{\pi}$ if n is half integer. When applied to R^2, $\lambda_2 = 0$, $n_1 = k$, and $n_2 = T-k-1$ so (8a) reduces to

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

Marchand (p. 173) states "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 $T \to \infty$ $E(R^2) = \rho^2$.

The previous discussion applies to the regression R^2, and while we next argue that the regression R^2 is an incomplete expression of hedging effectiveness, this previous discussion is nonetheless valuable in establishing the properties of hedging effectiveness. First, hedging effectiveness is more explicitly defined as

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

This definition establishes that the variances are for differences between actual and expected outcomes. Accordingly, hedging effectiveness is based not on the variances of simple outcomes but on the variances of unanticipated outcomes. This means that 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 become 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 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 contract price changes. Because the elements of $\mathbf{X_1}$ are systematic, they form anticipations so hedging effectiveness is redefined as

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

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

This expression demonstrates that the regression R^2 and hedging effectiveness are the same only when $\mathbf{X_1}$ consists of a sole column of ones. Otherwise, effectiveness 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 hedging effectiveness follow from analysis of variance definitions as SST = SSR($\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$) + SSE where SST = $\mathbf{Y'Y}$, SSR($\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$) = $\hat{\mathbf{Y}}'\hat{\mathbf{Y}} = \hat{\boldsymbol{\beta}}'\mathbf{X'X}\hat{\boldsymbol{\beta}} = \mathbf{Y'X}(\mathbf{X'X})^{-1}\mathbf{X'Y}$, and SSE is the sum of squared errors (i.e., SSE = $\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} = \mathbf{Y'}[\mathbf{I} - \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}]\mathbf{Y}$). Searle (p. 247) shows (a) that SSR($\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$) = SSR($\boldsymbol{\beta}_2$ | $\boldsymbol{\beta}_1$) + SSR($\boldsymbol{\beta}_1$), where SSR($\boldsymbol{\beta}_1$) = $\mathbf{Y'}\mathbf{X_1}(\mathbf{X_1'X_1})^{-1}\mathbf{X_1'}\mathbf{Y}$, (b) that SSR($\boldsymbol{\beta}_2$ | $\boldsymbol{\beta}_1$) = $\mathbf{Y'}[\mathbf{X}(\mathbf{X'X})^{-1}-\mathbf{X_1}(\mathbf{X_1'X_1})^{-1}\mathbf{X_1'}]\mathbf{Y}$, and (c) that SSR($\boldsymbol{\beta}_2$ | $\boldsymbol{\beta}_1$) / σ^2 has a noncentral χ^2 distribution and is independent of both SSR($\boldsymbol{\beta}_1$) and SSE.

Applying these definitions establishes

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

is distributed as a noncentral F random variable with k_2 numerator and T-K denominator degrees of freedom, and $\lambda = \boldsymbol{\beta}'\mathbf{X'}[\mathbf{I}-\mathbf{X_1}(\mathbf{X_1'X_1})^{-1}\mathbf{X_1}]\,\mathbf{X}\,\boldsymbol{\beta}\,/\,\sigma^2$, and

$$(11b) \quad \hat{e} = \frac{SSR(\boldsymbol{\beta}_2 \,|\, \boldsymbol{\beta}_1)}{SSE + SSR(\boldsymbol{\beta}_2 \,|\, \boldsymbol{\beta}_1)} = \frac{SSR(\boldsymbol{\beta}_2 \,|\, \boldsymbol{\beta}_1)}{SST - SSR(\boldsymbol{\beta}_1)}$$

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

The cumulative distribution of hedging effectiveness is derived from the noncentral F random variable in (11a). First, dividing both the numerator and denominator of (11a) by $\mathbf{Y'}[\mathbf{I}-\mathbf{X_1}(\mathbf{X_1'X_1})^{-1}\mathbf{X_1}]\mathbf{Y}$ expresses (11a) in terms of hedging effectiveness as

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

so that the probability statement

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

defines the cumulative probability distribution for hedging effectiveness.

Methods

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

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

where 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. While this model appears somewhat specific, it encompasses features that are common in hedge ratio regressions as X_1 contains a column of 1s to account for a long term spot price trend, systematic effects represented by dummy variables account for seasonal spot price variation, and the autoregressive term accounts for noninstantaneous spot price equilibration. Another advantage of (13) is that parameter estimation requires the inversion of the 3x3 matrix $X'X$ which is not computationally prohibitive. This consideration is especially important when the autoregressive parameter γ is estimated iteratively.

The model can be subjected to a variety of assumptions by changing its parameters. The parametric assumptions are specified by (a) the structural parameters α, β, γ, and δ (b) the variances of the random variables Δf_t and ε_t as $\Delta f_t \sim \text{N}(0, \sigma_{\Delta f}^2)$, $\varepsilon_t \sim \text{N}(0, \sigma_\varepsilon^2)$, (c) the covariance between Δf_t and ε_t, $\sigma_{\Delta f,\varepsilon}$, and (d) the size of each sample. While the values selected for these parameters are arbitrary, the results are nonetheless illustrative.

Once the parameter values are selected, a sample of size T is drawn and the hedge ratio and hedging effectiveness ($\hat{e}$) are estimated for the sample. Another sample of size T is drawn and the estimated hedge ratio is applied to these data and the out of sample hedging effectiveness ($\hat{\eta}$) is computed by comparing the unhedged outcome with the hedged outcome. This process is repeated 10,000 times for each sample size. The estimates from each sample are used to form the empirical cumulative probability distribution for the regression R^2, hedging effectiveness ($\hat{e}$), and out of sample hedging effectiveness ($\hat{\eta}$). The empirical distributions are compared to the theoretical distributions specified by (7b) and (12b). Also, because the population parameters are known, R^2 can be compared to ρ^2, and effectiveness ($\hat{e}$) and out of sample effectiveness ($\hat{\eta}$) can be compared to η. The sampling distributions for R^2, effectiveness, and of out of sample effectiveness are reported via summary statistics and cumulative probability distribution plots.

Results

Sample sizes (T) of 10, 50, 100, and 500 were selected. A sample size of 10 allows the small sample properties of our three measures to be studied. A sample size of 50 represents an empirical study that uses a year's weekly observations, or four years of monthly observations. Similarly, a sample size 100 represents an empirical study that uses two years of weekly observations, or eight years of monthly observations. Finally, a sample size of 500 represents an empirical study that uses ten years of weekly observations or two years of daily observations.

Parameter values of (0, 1, 0, -2) were selected for (α, β, γ, δ). These values were chosen because $\alpha = 0$ implies no trend in the spot price. $\gamma = 0$ simplifies the model by eliminating serial correlation so that preliminary results can be explored and established. $\beta = 1$ represents a direct hedging application. Finally, $\delta = -2$ is assumed so that the effect of the dummy variable is significant and the distinction between R^2 and hedging effectiveness can be demonstrated and emphasized.

The variances of Δf_t and ε_t are both set to 1 while the covariance between Δf_t and ε_t is set to zero. Applying Marchand's definition of the coefficient of determination to the variables that define effectiveness, we have $\sigma_{yy} = V(\beta \Delta f_t + e_t) = 2$, $\sigma_{yx} = C(\beta \Delta f_t + e_t, \Delta f_t) = 1$, and $\sigma_{xx} = V(\Delta f_t) = 1$, so η (the true value of effectiveness) is $\sigma_{YY}^{-1}\sigma_{YX}\Sigma_{XX}^{-1}\sigma_{XY} = \frac{1}{2}$.

In contrast, the regression R^2 includes the effect of the dummy variable. So $\sigma_{yy} = V(\delta D_t + \beta \Delta f_t + e_t) = (-2)^2 \times 0.25 + 1^2 \times 1 + 1 = 3$, $\sigma_{DD} = V(D_t) = E(D_t^2) - E(D_t)^2 = 0.25$, $\sigma_{\Delta f,\Delta f} = 1$, $\sigma_{yD} = C(\delta D_t + \beta \Delta f_t + e_t, D_t) = -2 \times 0.25 = -0.5$, and $\sigma_{y,\Delta f} = C(\delta D_t + \beta \Delta f_t + e_t, \Delta f_t) = 1$.

Accordingly, $\rho^2 = \sigma_{YY}^{-1}\sigma_{YX}\Sigma_{XX}^{-1}\sigma_{XY} = 3^{-1}\begin{bmatrix} -0.5 & 1 \end{bmatrix}\begin{bmatrix} 0.25 & 0 \\ 0 & 1 \end{bmatrix}^{-1}\begin{bmatrix} -0.5 \\ 1 \end{bmatrix} = \frac{2}{3}$.

Figure 1 summarizes simulations for sample sizes of 10, 50, 100, and 500 with each panel corresponding to a different sample size. Each panel shows the empirical cumulative distribution for the regression R^2, hedging effectiveness, and the out of sample hedging effectiveness. Each panel also shows the theoretical cumulative probability distribution of R^2 and effectiveness as defined by (7b) and (12b), respectively. The theoretical distributions (indicated by dashed lines) and the empirical distributions match so closely that they are indistinguishable.

To interpret the cumulative distributions we note that the farther to the right the cumulative distribution function lies, the higher the mean of the random variable. Accordingly, we observe in all four panels that mean R^2 exceeds the mean in sample hedging effectiveness. We previously established that the mean R^2 should be 0.67 while the mean effectiveness should be 0.50 so this ordering is as expected. Furthermore, the difference between R^2 and in sample hedging effectiveness appears constant across all four panels so the difference appears to be unaffected by sample size.

Figure 1 also shows that the mean in sample effectiveness exceeds the mean out of sample effectiveness. This difference diminishes as the sample size increases and with 500 observations the means in sample and out of sample effectiveness are approximately the same.

Figure 1 also reveals that even though hedging effectiveness is theoretically bounded by zero and one (*i.e.*, hedging in theory can never increase price risk and will potentially reduce price risk), it is possible that the application of estimated hedge ratios might actually increase price risk. This result occurred 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.

The second general principle in interpreting figure 1 is that the steeper the cumulative probability distribution function, the smaller the variance of the random variable. Figure 1 reveals that the cumulative probably distributions get steeper as sample sizes increase. Also, figure 1 indicates that for a given sample size the variances for R^2 and in sample effectiveness are generally the same while variance for out of sample effectiveness is larger. As the sample size increases, this relationship becomes less pronounced.

While figure 1 illustrates these relationships graphically, table 1 illustrates these and other comparisons more precisely. We cited Marchand's statement that generally, the $E(R^2) > \rho^2$. Given the correspondence between R^2 and hedging effectiveness, this also suggests that $E(\hat{e}) > \eta$. While $E(R^2) > \rho^2$ is apparent only for a sample of size 10, $E(\hat{e}) > \eta$ is more readily evident for samples of size 10 and 50. The results in table 1 suggest that the bias diminishes as the sample size increases. The numerical results in table 1 also suggest that while in sample hedging effectiveness is biased upward, out of sample hedging effectiveness is unbiased. This creates a new justification for researchers to forecast out of sample hedging effectiveness in that the estimate obtained is an unbiased estimator of hedging effectiveness.
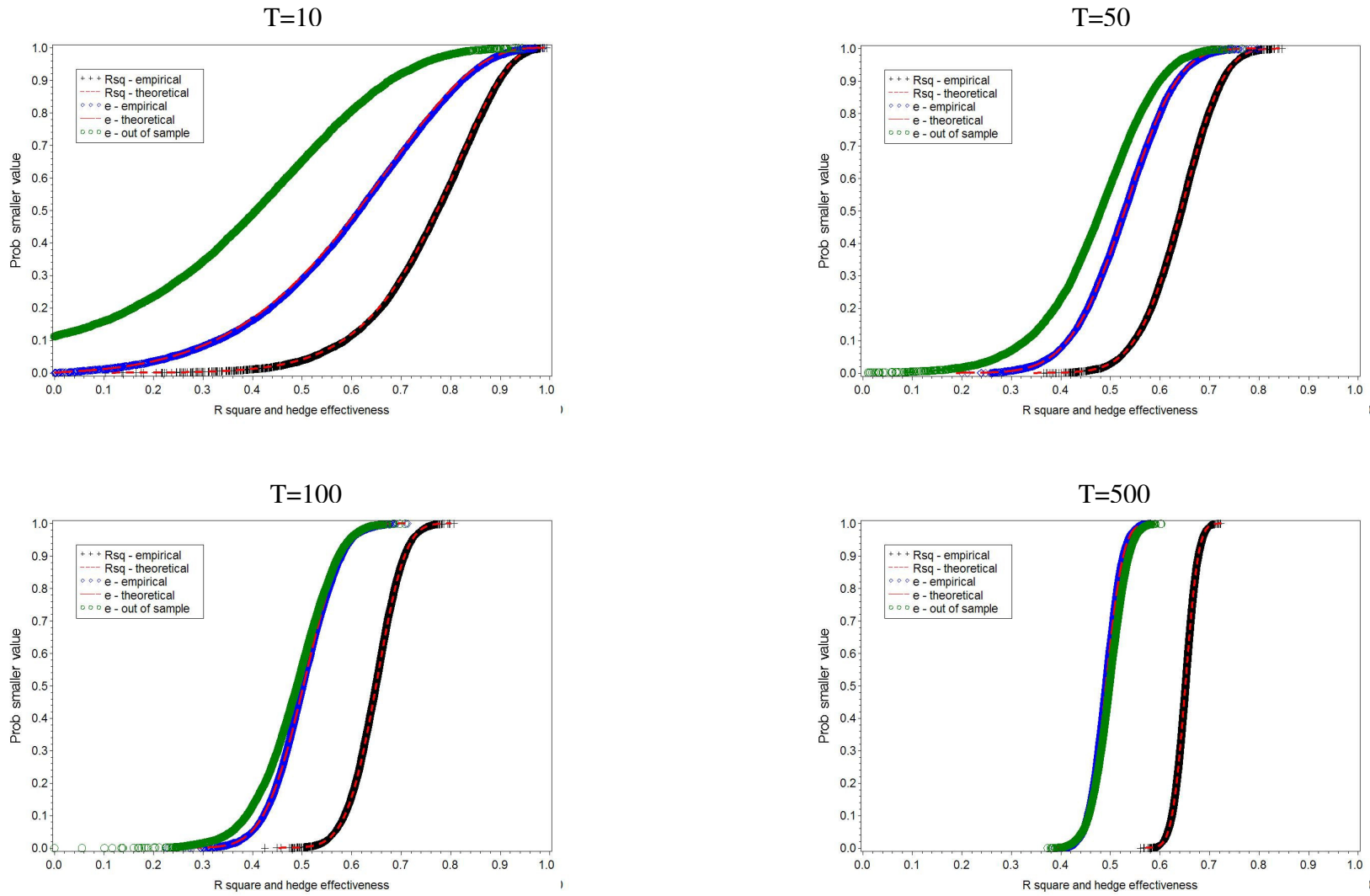
Figure 1. Cumulative probability distributions for R^2, effectiveness and out of sample effectiveness for various sample sizes.
Model: $\Delta p_t = \alpha + \delta D_t + \beta \Delta f_t + \varepsilon_t$, $t = 1, 2, \ldots T$. $\alpha = 0$, $\delta = -2$, $\beta = 1$, $\sigma_{\varepsilon\varepsilon} = 1$, $\sigma_{\Delta f, \Delta f} = 1$, $\sigma_{\Delta f, \varepsilon} = 0$.

Table 1. Aggregate simulation statistics.

T	Descriptive statistic	Regression R square (R^2)	In sample effectiveness $(\hat{e})$	Out of sample effectiveness $(\hat{\eta})$
Population mean		0.67	0.50	0.50

Model: $\Delta p_t = \alpha + \delta\, D_t + \beta\, \Delta f_t + \varepsilon_t$, $\alpha = 0$, $\delta = -2$, $\beta = 1$

T	Descriptive statistic	Regression R square (R^2)	In sample effectiveness $(\hat{e})$	Out of sample effectiveness $(\hat{\eta})$
10	Average	0.755	0.593	0.341
	Std Dev	0.125	0.190	0.360
	Min-max	0.103-0.993	0.000-0.987	-4.460-0.944
	90 %'tile range	0.401	0.631	1.007
50	Average	0.639	0.526	0.470
	Std Dev	0.069	0.086	0.109
	Min-max	0.365-0.847	0.200-0.803	-0.142-0.767
	90 %'tile range	0.227	0.283	0.255
100	Average	0.647	0.501	0.484
	Std Dev	0.047	0.061	0.074
	Min-max	0.425-0.806	0.227-0.713	0.000-0.708
	90 %'tile range	0.155	0.202	0.238
500	Average	0.651	0.490	0.497
	Std Dev	0.021	0.028	0.032
	Min-max	0.561-0.722	0.387-0.581	0.374-0.603
	90 %'tile range	0.069	0.091	0.104

Conclusions

This study is a preliminary investigation of the problem of forecasting how a hedging strategy will perform out of sample. 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, measured hedging effectiveness overstates the expected 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, effectiveness. When the estimated hedge ratios are applied out of sample, this maximization cannot be exercised so the expected in sample effectiveness will exceed the expected out of sample effectiveness.

We have also established that the difference between the expected in sample effectiveness and out of sample effectiveness 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 naturally overstates what will be experienced out of sample. In fact, the lower out of sample estimate is an unbiased estimate of 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 out of sample effectiveness is more variable than the in sample result and that the variance of both the in sample and the out of sample hedging effectiveness fall as the number of observations increase. This gives rise to an interesting question. Suppose the objective is to find an unbiased estimate of hedging effectiveness using the out of sample estimator. Then how should the available observations be allocated so as to obtain the most precise estimate? If more observations are allocated to hedge ratio estimation, then the hedge ratios will be more precise, but there will be fewer observations from which to compute out of sample hedging effectiveness causing its variance to increase. Conversely, using more observations for the computation of out of sample hedging effectiveness will leave fewer observations from which to compute hedge ratios so they become less precise. Is there a rule for the determination of the most efficient estimator given the number of observations?

This study leaves many issues unaddressed which gives rise to other questions. For example, many model specifications have not been examined. In particular, suppose the hedge ratio regression displays serial correlation. How does the distribution of hedging effectiveness and out of sample hedging effectiveness change? Also, the hedge ratio regression is but one equation out of a potentially simultaneous system. Suppose changes in futures prices and errors are correlated. How does this affect the distributions of the in sample and out of sample hedging effectiveness statistics? And finally, only one specification with $\eta = 0.5$ 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. 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.