Tests of the Difference between In-Sample and Post-Sample Hedging Effectiveness

by

Roger A. Dahlgran

Suggested citation format:

Dahlgran, R. A. 2015. "Tests of the Difference between In-Sample and Post-Sample Hedging Effectiveness." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.illinois.edu/nccc134].

Tests of the Difference between In-Sample and Post-Sample Hedging Effectiveness

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, 2015

Roger Dahlgran (dahlgran@u.arizona.edu) is an Associate Professor in the Department of Agricultural and Resource Economics at the University of Arizona. Mailing address: Department of Agricultural and Resource Economics, PO Box 210023, The University of Arizona, Tucson, AZ 85721-0023.

Tests of the Difference between In-Sample and Post-Sample Hedging Effectiveness

Hedging effectiveness is proportional price-risk reduction achieved by hedging. Typically, hedging studies estimate hedging effectiveness for the sample period then use estimated hedge ratios to simulate hedging and estimate hedging effectiveness in a "post-sample" period. This paper derives the statistical properties of the sample-period effectiveness estimator and the statistical properties of the difference between the sample-period and the post-sample period estimators. We find that the bias associated with the sample-period estimator is negligible and that a difference between the sample estimator and the post-sample estimator ties directly to changes in the structural parameters of the hedge-ratio regression. We develop tests for structural change and demonstrate those tests with an empirical example.

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

Introduction

Hedging studies typically specify a price-risk minimization problem; compile spot and futures price data; then estimate hedge ratios by subjecting part of the data to regression analysis. Hedging effectiveness, defined as the proportion of price risk eliminated by hedging, is reported as the regression R-square. These studies then apply the estimated hedge ratios to the remainder of the data to simulate hedged outcomes in a post-sample period as described by Lien(2007).[1] Comparing the variance of simulated unhedged post-sample outcomes with the variance of simulated hedged post-sample outcomes provides an estimate of the price-risk reduction expected from implementing the hedging strategy during a non-sample period. A comparison of the in-sample and the post-sample hedging effectiveness indicates the robustness of the hedging strategy. Given the prevalence of these comparisons, their fundamental assumptions merit scrutiny. That is the purpose of this paper.

While the comparison of in-sample and post-sample hedging effectiveness might be construed as a test for bias in the in-sample effectiveness estimator, this notion is flawed. The in-sample effectiveness estimator is in fact biased (Marchand, 1997; Lien, 2006), but determining the direction and magnitude of the bias requires the parameter's actual value and the estimator's expected value, neither of which are part of the in-sample / post-sample effectiveness comparison. The in-sample and post-sample hedging effectiveness estimates are simply observations of two random variables. Finding that they differ is not especially enlightening. Knowing each estimate's mean and variance, and knowing the statistical significance of their difference is more useful.

Second, the comparison of in-sample and post-sample hedging effectiveness might be construed as an evaluation of the robustness of the hedging strategy. This notion is also flawed. A robust hedging strategy requires an unchanged cash-futures price relationship in moving from the sample period to the post-sample period. Structural change will reduce the hedging strategy's effectiveness in the post-sample period and hence will make it appear less robust. This notion of robustness is better expressed as a test for parameter equality between the sample and post-sample periods. A simple comparison of two hedging effectiveness estimates lacks the statistical foundation provided by such a test.

Finally, detecting a difference between in-sample and post-sample hedging effectiveness is less informative than identifying the causes of the difference. For example, suppose that price risk in the post-sample period increases due to some structural shock and that the prescribed hedging strategy continues to be optimal though less effective. Should we conclude that the hedging strategy is faulty because it removes a smaller portion of a greater amount of price risk? Thoughtful analysis of the hedging strategy's performance during different periods requires a decomposition of the effectiveness statistic and then an inter-period comparison of the components.

This paper addresses these shortcomings. Our objective is to examine the statistical properties of the in-sample and post-sample hedging effectiveness estimators. We specifically focus on their probability distributions, biases, and standard deviations. We utilize mathematical statistics results, extend these results so they apply to hedging effectiveness, and then verify our extensions with thousands of simulated random draws of given-size samples.

We proceed by summarizing hedging objectives, hedge ratio estimation methods, and the hedging effectiveness statistic. Next, we consider the sampling distribution and the bias of the in-sample hedging effectiveness statistic. Then, we turn our attention to the post-sample effectiveness statistic and its relation to the in-sample statistic. This comparison suggests a series of tests that identify statistically significant discrepancies between the in-sample and post-sample effectiveness measures. We then provide an empirical application to demonstrate the workings of the suggested analysis. We end with some general conclusions.

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) $$\pi = x_s (p_1 - p_0) + x_f (f_{M1} - f_{M0}),$$

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_M is the M-maturity futures contract's price, and subscripts 0 and 1 indicate initiating and terminating transaction times. The optimal futures position, x_f^*, is the value of x_f that minimizes the variance of π. This minimum occurs when $x_f^*/x_s = -\sigma_{\Delta p,\Delta s} / \sigma^2_{\Delta f}$.

The risk minimizing hedge ratio (x_f^*/x_s) is estimated by $\hat{\beta}_1$ in the regression

(2) $$\Delta p_t = \beta_0 + \beta_1 \Delta f_{Mt} + \varepsilon_t, \; t = 1, 2, \ldots T$$

where Δ represents differencing over the hedging horizon, ε_t represents stochastic error at time t, and T represents the number of observations used in estimating of β_0 and β_1.[2] The risk minimizing futures position is $x_f^* = -\hat{\beta}_1 x_s$.

Anderson and Danthine (1980, 1981) generalized this approach to accommodate multiple futures positions. In this case, x_f, f_{M1}, and f_{M0} in (1) represent vectors of length k and hedge ratios are the parameters in the multiple regression

(3) $\Delta p_t = \beta_0 + \sum_{j=1}^{k} \beta_j \Delta f_{jt} + \varepsilon_t$, 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.

For commodity processors the profit outcome is $\pi = y\, p_{y,1} - x\, p_{x,0} + x_f\,(f_{M1} - f_{M0})$. In this case, input purchases (x) and output sales (y) are temporally separated by hedge horizon H but connected by product transformation with $y_t = \kappa\, x_{t-H}$. Hedge ratios result from fitting

(4) $p_{y,t} - \kappa\, p_{x,t-H} = \beta_0 + \sum_{j=1}^{k} \beta_j \Delta f_{jt} + \varepsilon_t$, t = 1, 2, 3, … T.

This specification has been applied 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, 2000; Rahman, Turner, and Costa).

Ederington defines hedging effectiveness (e) as the proportionate price-risk reduction available through hedging, or

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

where V is the variance operator, π_u the agent's unhedged outcome ($x_f = 0$) and π_h is the agent's hedged outcome ($x_f = -\hat{\beta}_1 x_s$). Lindahl observes, "The most popular measure of hedging effectiveness is commonly called R^2 … ". If hedge ratios are estimated with regressions (2), (3), or (4), then the regression R^2 is the hedging effectiveness estimator. However, if (2), (3), or (4) is augmented, then we shall see that the effectiveness estimator is related to, but is not the R^2.

R^2 is an estimator of the coefficient of determination, defined by Marchand (p. 168) as follows. Let $[\,\mathbf{Y} : \mathbf{X}\,] = [\,\mathbf{Y} : \mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_k}\,]$ be distributed as a k+1 variate normal with covariance matrix Σ. Let S be the covariance matrix obtained from a sample of size T where T > k > 1. Partition Σ and S as $\Sigma = \begin{bmatrix} \sigma_{YY} & \sigma_{YX} \\ \sigma_{XY} & \Sigma_{XX} \end{bmatrix}$ and $S = \begin{bmatrix} S_{YY} & S_{YX} \\ S_{XY} & S_{XX} \end{bmatrix}$ where σ_{YY} and S_{YY} are scalars. The multiple correlation coefficient between **Y** and $[\,\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_k}\,]$ is $\rho = (\sigma_{YY}^{-1} \sigma_{YX} \Sigma_{XX}^{-1} \sigma_{XY})^{1/2}$ 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 points out (p. 173) that R^2 is a biased estimator of ρ^2 because $E(R^2) > \rho^2$, but as $T \to \infty$, $E(R^2) = \rho^2$.

Determining the magnitude of the bias in R^2 requires its probability distribution, which follows from the probability distribution of the regression F statistic. For (2), (3), or (4)

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

is used to test whether the noncentrality parameter, $\lambda = \boldsymbol{\beta}'(\mathbf{X}-\overline{\text{X}})'(\mathbf{X}-\overline{\text{X}})\boldsymbol{\beta}/\sigma_{\varepsilon\varepsilon} = T\rho^2/(1-\rho^2)$, of the numerator chi-square is zero.[3] The F statistic follows the well-known F distribution so the cumulative distribution for R^2 results from restating the probabilities

$$\text{(6b)} \qquad \Pr\left\{\frac{R^2}{1-R^2}\left(\frac{T-k-1}{k}\right) < f_{T-k-1}^{k,\lambda}(\alpha)\right\} = \Pr\left\{R^2 < \frac{k\ f_{T-k-1}^{k,\lambda}(\alpha)}{(T-k-1)+k\ f_{T-k-1}^{k,\lambda}(\alpha)}\right\} = \alpha,$$

where $f_{T-k-1}^{k,\lambda}(\alpha)$ is the numerical value that has probability α of a smaller value of a noncentral F random variable with k (numerator) and T-k-1 (denominator) degrees of freedom, and noncentrality parameter λ. Thus, given α, T, k and ρ^2, we can compute F-values and construct the cumulative distribution function (CDF) for R^2.

Chattamvelli provides a more general result.

> *"If $\chi^2_{n_1}$ and $\chi^2_{n_2}$ are independent central chi squared random variables with n_1 and n_2 degrees of freedom, then $F = (\chi^2_{n_1}/n_1)/(\chi^2_{n_2}/n_2)$ 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."*

If $\varepsilon_t \sim NID(0, \sigma^2_{\varepsilon\varepsilon})$, then the regression F statistic is composed of the requisite independent chi square random variables, and the regression R^2 has a singly noncentral beta distribution with $n_1 = k$, $n_2 = T - k - 1$, $\lambda_1 = T\rho^2/(1-\rho^2)$, and $\lambda_2 = 0$. (6b) expresses this CDF.

Pe and Drygas derive the probability density, moments, and cumulative distribution of a doubly noncentral beta random variable. They define X_1 and X_2 as independent noncentral χ^2s with n_i degrees of freedom and noncentrality parameters λ_i (i = 1, 2), so $Z = X_1/(X_1+X_2)$ is a doubly noncentral beta distribution with parameters n_1, n_2, and λ_1, λ_2. When $\lambda_2 = 0$ the probability density, moments, and cumulative distribution functions simplify to

$$\text{(7a)} \quad f(z;n_1,n_2,\lambda_1) = \sum_{i=0}^{\infty} p(i;\lambda_1) \frac{z^{i+n_1/2-1}(1-z)^{n_2/2-1}}{B(i+\frac{n_1}{2},\frac{n_2}{2})},$$

$$\text{(7b)} \quad E(Z^r;n_1,n_2,\lambda_1) = \sum_{i=0}^{\infty} p(i;\lambda_1) \frac{B(i+\frac{n_1}{2}+r,\frac{n_2}{2})}{B(i+\frac{n_1}{2},\frac{n_2}{2})}, \text{ and}$$

$$\text{(7c)} \quad F(Z;n_1,n_2,\lambda_1) = \sum_{i=0}^{\infty} p(i;\lambda_1) \frac{B_Z(i+\frac{n_1}{2},\frac{n_2}{2})}{B(i+\frac{n_1}{2},\frac{n_2}{2})},$$

where $p(i;\lambda_1) = e^{-\lambda_1/2} \frac{(\lambda_1/2)^i}{i!}$, i=0, 1, 2, … is a Poisson pdf, B(a,b) represents the complete beta function, and $B_Z(a,b)$ represents the incomplete beta function. The similarities among (7a), (7b) and (7c) include a Poisson weighting applied to corresponding central beta functions. (7b) indicates that the moments of R^2 are expressed by an infinite series, a result that is consistent with Muirhead's claim that "*The exact moments of R^2 … are notoriously complicated.*"

(7a) generates probability density functions (pdfs) for R^2 (represented by z) given n_1, n_2 and $\lambda_1 = \lambda_1(T, \rho^2)$. Figure 1 shows these pdfs for ρ^2 of 0.2, 0.5, and 0.9 and T of 15, 52, and 202. We see in figure 1 that the variance of R^2 declines as sample size increases, that R^2 is an upward-biased estimator for ρ^2, and that the bias in R^2 diminishes as the sample size increases. (7b) was used to compute the bias and standard deviation of R^2 assuming various sample sizes and ρ^2s. These results, shown in table 1, indicate that the bias is positive and diminishes as the sample size increases. For the sample sizes shown, the bias is less than one standard deviation.

The preceding R^2 distributional results provide the foundation for hedging effectiveness distributions. Combining the Ederington's effectiveness definition (5) with the hedging strategies derived from the regression models (2), (3), or (4) gives

$$\text{(8)} \quad e = \frac{E\{[\Delta p_t - E(\Delta p_t)]^2\} - E\{[\Delta p_t - \mathbf{\Delta f_t}\,\hat{\boldsymbol{\beta}} - E(\Delta p_t - \mathbf{\Delta f_t}\,\hat{\boldsymbol{\beta}})]^2\}}{E\{[\Delta p_t - E(\Delta p_t)]^2\}}.$$

where $\mathbf{\Delta f_t}$ represents either a scalar or a vector with $\hat{\boldsymbol{\beta}}$ defined accordingly. The variances are for differences between actual and expected outcomes, or more simply, the variances are for unanticipated outcomes. In this regard, R^2 potentially overstates hedging effectiveness because if the cash-price changes display systematic behavior such as seasonality or serial correlation, then these systematic components are present in the expected outcome whether or not hedging occurs. To isolate systematic behavior, the general hedge ratio regression, $\mathbf{Y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, in (2), (3), or (4) is restated as $\mathbf{Y} = \mathbf{M}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\varphi} + \boldsymbol{\varepsilon}$ where the K columns of **X** are partitioned as k_1 deterministic and/or conditioning variables in **M**, and k_2 futures price changes in **Z**. In addition to a column of ones for the intercept, **M** might also contain dummy variables representing seasonal spot price behavior or lagged spot prices to account for serially correlated errors

(Meyers and Thompson). The elements of **M** form anticipations so the unanticipated outcomes are $\mathbf{Y}-\mathbf{M}\hat{\boldsymbol{\alpha}}$ without hedging and $\mathbf{Y}-\mathbf{M}\hat{\boldsymbol{\alpha}}-\mathbf{Z}\hat{\boldsymbol{\varphi}}$ with hedging.

The hedging effectiveness estimator that accounts for systematic spot price behavior, is

$$(9) \qquad e = \frac{(\mathbf{Y}-\mathbf{M}\hat{\boldsymbol{\alpha}})'(\mathbf{Y}-\mathbf{M}\hat{\boldsymbol{\alpha}}) - (\mathbf{Y}-\mathbf{M}\hat{\boldsymbol{\alpha}}-\mathbf{Z}\hat{\boldsymbol{\varphi}})'(\mathbf{Y}-\mathbf{M}\hat{\boldsymbol{\alpha}}-\mathbf{Z}\hat{\boldsymbol{\varphi}})}{(\mathbf{Y}-\mathbf{M}\hat{\boldsymbol{\alpha}})'(\mathbf{Y}-\mathbf{M}\hat{\boldsymbol{\alpha}})}$$
$$= \frac{\mathbf{Y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{M}(\mathbf{M}'\mathbf{M})^{-1}\mathbf{M}']\mathbf{Y}}{\mathbf{Y}'[\mathbf{I}-\mathbf{M}(\mathbf{M}'\mathbf{M})^{-1}\mathbf{M}']\mathbf{Y}}.$$

Analogous to Marchand's definition of the coefficient of determination, e is an estimator for $\eta = (\sigma_{yy}^{-1}\sigma_{yZ}\Sigma_{ZZ}^{-1}\sigma_{Zy})^2$ where $\mathbf{y} \equiv \mathbf{Y}-\mathbf{M}\boldsymbol{\alpha} = \mathbf{Z}\boldsymbol{\varphi}+\boldsymbol{\varepsilon}$. (9) indicates that R^2 and e differ unless **M** is simply a column of ones. Otherwise, R^2 overstates hedging effectiveness by over-allocating degrees of freedom (K-1 instead of K-k_1) and sums of squares to the numerator, and by under-allocating degrees of freedom (T-1 instead of T-k_1) and sum of squares to the denominator. Thus, in addition to the upward bias that R^2 displays in estimating ρ^2, it also potentially overstates hedging effectiveness by failing to recognize systematic spot price behavior.

The statistical properties of e follow from the statistical properties of ε_t and analysis of variance definitions. Let SSR($\boldsymbol{\alpha}, \boldsymbol{\phi}$) = $\hat{\mathbf{Y}}'\hat{\mathbf{Y}} = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{X}\hat{\boldsymbol{\beta}}$ and SSE= $\mathbf{Y}'[\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}$ so $\mathbf{Y}'\mathbf{Y}$ = SSR($\boldsymbol{\alpha}, \boldsymbol{\phi}$) + SSE. Searle (p. 247) shows that (a) SSR($\boldsymbol{\alpha}, \boldsymbol{\phi}$) = SSR($\boldsymbol{\phi} \mid \boldsymbol{\alpha}$) + SSR($\boldsymbol{\alpha}$), where SSR($\boldsymbol{\alpha}$) = $\mathbf{Y}'\mathbf{M}(\mathbf{M}'\mathbf{M})^{-1}\mathbf{M}'\mathbf{Y}$, and SSR($\boldsymbol{\phi} \mid \boldsymbol{\alpha}$) = $\mathbf{Y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{M}(\mathbf{M}'\mathbf{M})^{-1}\mathbf{M}']\mathbf{Y}$, (b) SSR($\boldsymbol{\phi} \mid \boldsymbol{\alpha}$) is independent of both SSR($\boldsymbol{\alpha}$) and SSE, and (c) SSR($\boldsymbol{\phi} \mid \boldsymbol{\alpha}$) / σ^2 has a non-central χ^2 distribution if $\boldsymbol{\varepsilon} \sim$ MVN($\mathbf{0}, \boldsymbol{\sigma}^2\mathbf{I}$). These definitions establish that

$$(10a) \qquad F = \frac{SSR(\boldsymbol{\varphi} \mid \boldsymbol{\alpha})/k_2}{SSE/(T-K)} = \frac{\mathbf{Y}'[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - \mathbf{M}(\mathbf{M}'\mathbf{M})^{-1}\mathbf{M}']\mathbf{Y}/(K-k_1)}{\mathbf{Y}'[\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}/(T-K)}$$

has a noncentral F distribution with k_2 numerator degrees of freedom, T-K denominator degrees of freedom, and $\lambda = \boldsymbol{\beta}' [\mathbf{X}'\mathbf{X} - \mathbf{X}'\mathbf{M}(\mathbf{M}'\mathbf{M})^{-1}\mathbf{M}'\mathbf{X}]\boldsymbol{\beta} / \sigma^2 = T\eta / (1-\eta)$, and that

$$(10b) \qquad e = \frac{SSR(\boldsymbol{\varphi} \mid \boldsymbol{\alpha})}{SSE + SSR(\boldsymbol{\varphi} \mid \boldsymbol{\alpha})} = \frac{k_2 F}{(T-K)+k_2 F}$$

is a singly noncentral beta random variable with corresponding degrees of freedom and λ.[4] Thus, (7a), (7b), and (7c) express the pdf, moments and CDF for hedging effectiveness where $n_1 = k_2$, $n_2 = T - K$, and $\lambda = T\eta / (1-\eta)$. (10b) provides an expression for computing hedge effectiveness based on the F statistic for the for the hypothesis that the hedge ratios are zero.

The preceding relies on the classical regression assumption of nonstochastic **X**, making $\lambda_1 = \boldsymbol{\varphi}'(\mathbf{Z}-\overline{\mathbf{Z}})'(\mathbf{Z}-\overline{\mathbf{Z}})\boldsymbol{\varphi}/\sigma^2$ constant for all samples. This assumption is untenable in our context because it requires identical observations of futures-prices changes in the sample and the post-sample periods[5]. In reality, the futures price changes in **Z** are beyond the experimenter's control and we observe a different **Z** in each sample. If **Z** is assumed to be generated by a MVN process, and if **Z** and **ε** are jointly independent, and if $\overline{\mathbf{Z}} = (\mathbf{M'M})^{-1}\mathbf{M'Y}$, then $\lambda_1 = \boldsymbol{\varphi}'(\mathbf{Z}-\overline{\mathbf{Z}})'(\mathbf{Z}-\overline{\mathbf{Z}})\boldsymbol{\varphi}/\sigma^2$ is a chi-square random variable with T-k_1 degrees of freedom.

(7a) through (7c) are conditional on λ_1 so the pdf is g(z | n_1, n_2, λ_1). If λ_1 is random, then f(z | n_1, n_2) = $\int_0^\infty g(\,z\,|\,n_1, n_2, \lambda_1)\,h(\,\lambda_1\,|\,n_2, n_2)\,d\lambda_1$ where h(.) is a chi-square density with T-k_1 degrees of freedom. Substitution and integration gives the analogs to (7a) through (7c),

$$\text{(12a)}\quad f(z;n_1,n_2) = \sum_{i=0}^{\infty} g(i;\eta,k_1)\frac{z^{i+n_1/2-1}(1-z)^{n_2/2-1}}{B(i+\frac{n_1}{2},\frac{n_2}{2})},$$

$$\text{(12b)}\quad E(Z^r;n_1,n_2) = \sum_{i=0}^{\infty} g(i;\eta,k_1)\frac{B(i+\frac{n_1}{2}+r,\frac{n_2}{2})}{B(i+\frac{n_1}{2},\frac{n_2}{2})}, \text{ and}$$

$$\text{(12c)}\quad F(Z;n_1,n_2) = \sum_{i=0}^{\infty} g(i;\eta,k_1)\frac{B_Z(i+\frac{n_1}{2},\frac{n_2}{2})}{B(i+\frac{n_1}{2},\frac{n_2}{2})},$$

where $g(i;\eta,k_1) = \dfrac{\Gamma\left(\frac{T-k_1}{2}+i\right)}{i!\,\Gamma\left(\frac{T-k_1}{2}\right)}(1-\eta)^{\left(\frac{T-k_1}{2}\right)}\eta^i$, $\eta = \dfrac{\phi'\Sigma_{ZZ}\phi}{\sigma^2+\phi'\Sigma_{ZZ}\phi}$, and Σ_{ZZ} represents the stochastic regressor's covariance matrix. The probability density, expected values, and cumulative distribution are again weighted sums of beta functions. However, when λ_1 is random the weights applied to the beta functions are more dispersed than when λ_1 is fixed. As a result, z (which generally represents e or R^2) is more dispersed.

Figure 2 shows CDFs generated by (12c) for sample size 15, η = 0.3, and for sample size 100, η = 0.7. To validate (12c), figure 2 also shows the corresponding empirical cumulative distribution of e for 100,000 randomly drawn samples of **Z** and **ε**. The correspondence between the CDF (from 12c) and the empirical cumulative distribution (generated from the random draws) supports (12c) as the hedging effectiveness probability model when **Z** and **ε** are stochastic. Figure 2 also shows the CDF for nonstochastic regressors (i.e., 7c). The fixed-regressor CDF is steeper that the stochastic-regressor CDF indicating that stochastic regressors impart additional variability on hedging effectiveness.

Comparing E(e) from (12b) to η gives the effectiveness estimator's bias when futures prices are stochastic. Table 2 shows the bias and standard deviation of e for various sample size and η combinations. The bias is generally positive when $\eta \le 0.5$ and negative when $\eta > 0.5$.[6] This contrasts with the non-stochastic regressor assumption where the bias is positive for all values of η (table 1). Table 2 also reveals that (a) the absolute value of the bias is smaller with stochastic regressors than with non-stochastic regressors (compare to table 1), and (b) the bias is small relative to the standard deviation of hedging effectiveness.

Having determined the statistical properties of the in-sample hedging effectiveness estimator, we now turn to the post-sample hedging effectiveness estimator. To begin, represent the hedge ratio regression as $\mathbf{Y_i} = \mathbf{X_i}\boldsymbol{\beta}_\mathbf{i} + \boldsymbol{\varepsilon}_\mathbf{i} = \mathbf{M_i}\boldsymbol{\alpha}_\mathbf{i} + \mathbf{Z_i}\boldsymbol{\varphi}_\mathbf{i} + \boldsymbol{\varepsilon}_\mathbf{i}$ where i = 1 indicates the sample period with T_1 observations, and i = 2 indicates the simulation or post-sample period with T_2 observations. Define $\hat{\boldsymbol{\alpha}}_\mathbf{i} = (\mathbf{M_i}'\mathbf{M_i})^{-1}\mathbf{M_i}'\mathbf{Y_i}$ and $\hat{\boldsymbol{\beta}}_\mathbf{i} = (\mathbf{X_i}'\mathbf{X_i})^{-1}\mathbf{X_i}'\mathbf{Y_i}$ for each period (i = 1, 2). Let $e_{2|1}$ represent the post-sample hedging effectiveness estimate computed by applying period 1 hedge ratio estimates to period 2 data. From (8),

$$\text{(13a)} \quad e_{2|1} = [\hat{V}_2(\pi_u) - \hat{V}_2(\pi_h)] / \hat{V}_2(\pi_u)$$

where $\hat{V}_2$ indicates a variance estimated in the post-sample period, $\pi_u = x_s\,\Delta p$, $\pi_h = x_s\,\Delta p + x_f\,\Delta f$ with $x_f = -x_s\,\hat{\boldsymbol{\varphi}}_\mathbf{1}$. Then

$$\text{(13b)} \quad \hat{V}_2(\pi_u) = T_2^{-1}\, x_s^2\, (\mathbf{Y_2} - \mathbf{M_2}\hat{\boldsymbol{\alpha}}_\mathbf{1})'(\mathbf{Y_2} - \mathbf{M_2}\hat{\boldsymbol{\alpha}}_\mathbf{1})$$
$$= T_2^{-1}\, x_s^2\, [\mathbf{Y_2}'\mathbf{Y_2} - \hat{\boldsymbol{\alpha}}_\mathbf{2}'\mathbf{M_2}'\mathbf{M_2}\hat{\boldsymbol{\alpha}}_\mathbf{2} + (\hat{\boldsymbol{\alpha}}_\mathbf{2} - \hat{\boldsymbol{\alpha}}_\mathbf{1})'\mathbf{M_2}'\mathbf{M_2}(\hat{\boldsymbol{\alpha}}_\mathbf{2} - \hat{\boldsymbol{\alpha}}_\mathbf{1})],$$

$$\text{(13c)} \quad \hat{V}_2(\pi_h) = T_2^{-1}\, x_s^2\, (\mathbf{Y_2} - \mathbf{X_2}\hat{\boldsymbol{\beta}}_1)'(\mathbf{Y_2} - \mathbf{X_2}\hat{\boldsymbol{\beta}}_1)$$
$$= T_2^{-1}\, x_s^2\, [\mathbf{Y_2}'\mathbf{Y_2} - \hat{\boldsymbol{\beta}}_\mathbf{2}'\mathbf{X_2}'\mathbf{X_2}\hat{\boldsymbol{\beta}}_\mathbf{2} + (\hat{\boldsymbol{\beta}}_\mathbf{2} - \hat{\boldsymbol{\beta}}_\mathbf{1})'\mathbf{X_2}'\mathbf{X_2}(\hat{\boldsymbol{\beta}}_\mathbf{2} - \hat{\boldsymbol{\beta}}_\mathbf{1})], \text{ and}$$

$$\text{(13d)} \quad e_{2|1} = \left(e_2 - \frac{Q_{\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1} - Q_{\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1}}{SSE_2(\boldsymbol{\alpha})} \right) \Bigg/ \left(1 + \frac{Q_{\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1}}{SSE_2(\boldsymbol{\alpha})} \right)$$

where $e_2 = \dfrac{\hat{\boldsymbol{\beta}}_\mathbf{2}'\mathbf{X_2}'\mathbf{X_2}\hat{\boldsymbol{\beta}}_\mathbf{2} - \hat{\boldsymbol{\alpha}}_\mathbf{2}'\mathbf{M_2}'\mathbf{M_2}\hat{\boldsymbol{\alpha}}_\mathbf{2}}{\mathbf{Y_2}'\mathbf{Y_2} - \hat{\boldsymbol{\alpha}}_\mathbf{2}'\mathbf{M_2}'\mathbf{M_2}\hat{\boldsymbol{\alpha}}_\mathbf{2}}$, $Q_{\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1} = (\hat{\boldsymbol{\beta}}_\mathbf{2} - \hat{\boldsymbol{\beta}}_\mathbf{1})'\mathbf{X_2}'\mathbf{X_2}(\hat{\boldsymbol{\beta}}_\mathbf{2} - \hat{\boldsymbol{\beta}}_\mathbf{1})$, $Q_{\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1} = (\hat{\boldsymbol{\alpha}}_\mathbf{2} - \hat{\boldsymbol{\alpha}}_\mathbf{1})'\mathbf{M_2}'\mathbf{M_2}(\hat{\boldsymbol{\alpha}}_\mathbf{2} - \hat{\boldsymbol{\alpha}}_\mathbf{1})$, and $\mathbf{Y_2}'\mathbf{Y_2} - \hat{\boldsymbol{\alpha}}_\mathbf{2}'\mathbf{M_2}'\mathbf{M_2}\hat{\boldsymbol{\alpha}}_\mathbf{2} = SSE_2(\boldsymbol{\alpha})$. (13d) expresses $e_{2|1}$ in terms of post-sample regression estimates. Specifically, e_2 is the hedging effectiveness estimate from the post sample regression, $\hat{\boldsymbol{\alpha}}_\mathbf{2}$ and $\hat{\boldsymbol{\beta}}_\mathbf{2}$ are defined above, and $SSE_2(\boldsymbol{\alpha})$ is the sum of squares for an unhedged position in the post-sample period. $Q_{\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1}$, $Q_{\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1}$, ($Q_{\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1}$ - $Q_{\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1}$), and $SSE_2(\boldsymbol{\alpha})$ are all positive-define quadratic forms so $e_{2|1} = e_2$ if and only if $Q_{\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1} = 0$ and $Q_{\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1} = 0$. Otherwise, $e_{2|1}$ is less than e_2.

The distribution of $e_{2|1}$ is unknown but it is composed of random variables with known distributions under the $\boldsymbol{\varepsilon} \sim$ MVN assumption. Specifically, e_2 has a non-central beta distribution with a random non-centrality parameter. Further, each of the quadratic forms $Q_{\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1}$, $Q_{\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1}$, ($Q_{\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1}$ - $Q_{\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1}$), and $SSE_2(\boldsymbol{\alpha})$ consists of random variables[7] surrounding the inverse of their covariance matrix, so the realization of these quadratic forms has a chi-square distribution with the appropriate degrees of freedom. Hence, (13d) indicates that the procedure for analyzing differences between the in-sample effectiveness estimate (e_1) and the post-sample effectiveness estimate ($e_{2|1}$) is to test the sources of a difference between e_2 and $e_{2|1}$, and then to test for a significant difference between e_1 and e_2. (13d) specifies hypotheses to test and the statistic to use for each test. We will see that this procedure amounts to testing for structural change in the hedge-ratio regression model between the sample and the post-sample periods.

First, the systematic behavior of spot prices could change between the sample and post-sample periods. (i.e., $\boldsymbol{\alpha}_1 \neq \boldsymbol{\alpha}_2$). $Q_{\hat{\alpha}_2-\hat{\alpha}_1}$ in (13d), suggests testing H_1: $\boldsymbol{\alpha}_2 = \hat{\boldsymbol{\alpha}}_1$.[8] The test statistic, F = $[Q_{\hat{\alpha}_2-\hat{\alpha}_1}/k_1]/[SSE_2(\boldsymbol{\alpha})/(T_2-k_1)]$ is distributed as a central F random variable with k_1 and T_2-k_1 degrees of freedom under H_1.

Second, after allowing for systematic spot-price behavior change between the two periods, the optimal hedge ratios could also change (i.e. $\boldsymbol{\varphi}_2 \neq \boldsymbol{\varphi}_1$). The expression $Q_{\hat{\beta}_2-\hat{\beta}_1} - Q_{\hat{\alpha}_2-\hat{\alpha}_1}$ in (13d) suggests testing H_2: $\boldsymbol{\varphi}_2 = \hat{\boldsymbol{\varphi}}_1$ while allowing $\boldsymbol{\alpha}_1 \neq \boldsymbol{\alpha}_2$. The test statistic, F = $\{[Q_{\hat{\beta}_2-\hat{\beta}_1} - Q_{\hat{\alpha}_2-\hat{\alpha}_1}]/k_2\}/[SSE_2(\boldsymbol{\beta})/(T_2-K)]$ is distributed as a central F random variable with k_2 and T_2-K degrees of freedom under H_2.[9]

If neither H_1 nor H_2 is rejected, then the difference between e_2 and $e_{2|1}$ is not significant. In the extreme case, if $\hat{\boldsymbol{\beta}}_2 = \hat{\boldsymbol{\beta}}_1$, then $e_{2|1} = e_2$.[10] This exposes a third source of a difference between e_1 and $e_{2|1}$, namely the difference between e_1 and e_2. Testing H_3: $\eta_1 = \eta_2$ determines whether this difference is significant. The methodology for this test is well established (Papoulis, Snedechor and Cochran) as it derives from testing the equality of correlations (i.e., H_0: $\rho_1 = \rho_2$). Note that (a) R^2 is the squared correlation between $Y_t - \overline{Y}$ and $\hat{Y}_t - \overline{Y}$, (b) because of least squares estimation, $R \geq 0$, hence, (c) R is the positive root of R^2. Similar arguments apply to hedging effectiveness except that the Y_t deviations are from the conditional mean, $\mathbf{M}_t\ \hat{\boldsymbol{\alpha}}$. Hence, the positive root of hedging effectiveness is a conditional correlation. The test for equality of two correlations applies the Fisher Z-transformation to each correlation so $Z_i = ½ \ln((1+e_i) / (1 - e_i))$ and $z = (Z_1 - Z_2)/[(N_1-3)^{-1} + (N_2-3)^{-1}]^{½}$ is approximately normally distributed.

If we reject H_3, then η_1 and η_2 could differ for two reasons. First, the regression error may be heteroscedastic vis-à-vis the sample and post-sample periods. We test H_{3a}: $\sigma_1^2 = \sigma_2^2$ with an F-test. Second, η_1 and η_2 could differ because, after correcting for systematic behavior, the variability of the hedged outcome (i.e., $\boldsymbol{\varphi}'\boldsymbol{\Sigma}_{ZZ}\boldsymbol{\varphi}$) may be heteroscedastic vis-à-vis the sample and post-sample periods. This variability shift may be due to a change in volatility of one or more of the futures prices or due to a change in co-volatility among the futures prices. This change in the covariance structure of futures prices is more likely to occur in a cross-hedging application than in a direct-hedging application. Regardless of the source, an F-test of H_{3b}: $\boldsymbol{\varphi}_1'\boldsymbol{\Sigma}_{ZZ,1}\boldsymbol{\varphi}_1 = \boldsymbol{\varphi}_2'\boldsymbol{\Sigma}_{ZZ,2}\boldsymbol{\varphi}_2$ is appropriate.

An Empirical Application

We have shown that the comparison of e_1 and $e_{2|1}$ is equivalent to testing for changes in the structural parameters $\boldsymbol{\alpha}_i$, $\boldsymbol{\varphi}_i$, σ_i^2, and $\boldsymbol{\Sigma}_{zz,i}$ between the sample period (i=1) and the post-sample period (i=2). The rigor that is missing in the simple comparison of these two statistics is brought to bear in testing for parametric change for each source of the e_1, $e_{2|1}$ difference. We demonstrate our tests with an application to Brazilian ethanol inventory hedging that uses Chicago Board of Trade ethanol and Chicago Mercantile Exchange Brazilian real (R$) futures contracts.[11] A typical hedge ratio model for this application is

(14a) $\Delta(p_t\, r_t) = \beta_0 + \beta_1\, Q_{1t} + \beta_2\, Q_{2t} + \beta_3\, Q_{3t} + \beta_4\, \Delta r_{Tt} + \beta_5\, \Delta f_{Tt} + \varepsilon_t$

where p_t is the Brazilian fuel ethanol spot price (R$/liter) at time t, r_t is the R$ spot exchange rate ($/R$) at time t, Q_{it} is a dummy variable representing quarter i at time t (1 if quarter i, 0 otherwise), r_{Tt} is the T-maturity R$ futures contract price ($/R$) at time t, f_{Tt} is the T-maturity ethanol futures contract price ($/gal) at time t, Δ represents an eight-week hedge horizon, and $\varepsilon_t = \rho\, \varepsilon_{t-1} + \nu_t$. This model illustrates multi-contract hedging, and the presence of quarterly dummies and serial correlation permits a contrast between R-square and hedging effectiveness.

Table 3 summarizes the data used to estimate (14a). The Brazilian ethanol spot prices are weekly averages. We treat these averages as midweek prices and match them with Wednesday futures prices to avoid weekend-related volatility effects.

The March 24, 2005 start of U.S. ethanol futures trading determines the start of the daily ethanol futures price series. We use the nearby futures contract as the hedge vehicle if its last trading day is at least one week beyond the hedge termination date. Otherwise, we use the next nearby maturity. This one-week maturity buffer avoids potential price volatility increases related to contract maturity. Figure 3 shows our spot and futures prices. The spot price spike in the first half of 2011 is due to a brief inter-harvest sugarcane shortage (Jelmayer, 2011).

We consider the nearby Brazilian real futures contract as a potential hedge vehicle. We again use the nearby contract if last trading day is at least one week beyond the hedge termination date. We do not require ethanol and R$ futures maturities to match. While this pairing is attractive as both contracts have maturities for all months and nearly matching last trading days[12], this correspondence is not universal as the R$ has only four maturities per year through April 2007 and ethanol's last trading day was the business day prior to the 15th of the month through the August 2006 maturity.

Finally, to model the practice of using sample results to estimate post-sample effectiveness, we divide our data into a 2005-2010 sample period and a 2011-2013 post-sample period. The post-sample period contains the 2011 price spike (figure 3). Because of the spacing of our observations and the serial correlation implicit in the model, this spike affects two observations. Rather than deleting these two observations and creating a missing value gap in a serially correlated time series, we account for the errors with dummy variables.
A model that represents parameter differences between the two periods is

(14b) $\Delta(p_t\, r_t) = \beta_0 + \beta_1\, Q_{1t} + \beta_2\, Q_{2t} + \beta_3\, Q_{3t} + \beta_4\, \Delta r_{Tt} + \beta_5\, \Delta f_{Tt} +$
$D_t\, (\delta_0 + \delta_1\, Q_{1t} + \delta_2\, Q_{2t} + \delta_3\, Q_{3t} + \delta_4\, \Delta r_{Tt} + \delta_5\, \Delta f_{tTt}\,) + \delta_6\, D_{2011.1,t} + \delta_7\, D_{2011.2,t} + \varepsilon_t$

$$\varepsilon_t = \rho\, \varepsilon_{t-1} + \psi\, D_t\, \varepsilon_{t-1} + \nu_t$$

where $D_t = 1$ for observations in 2011 through 2013, 0 otherwise, and $D_{2011.1,t}$ and $D_{2011.2,t}$ represent dummy variable for the first and second observations in 2011. δ_i and ψ represent changes in the base period parameters in the post-sample period.

Table 4 summarizes estimation results for (14b). In period 1 we observe the following. The quarterly effects are significant as the first and second quarters differ from the fourth (base) quarter with p-values of 0.016 and 0.038, respectively. Serial correlation is significant as indicated by the p-value of 0.001. The hedge ratios for both the R$ and the ethanol futures are significant and of the expected sign with p-values of 0.0002 and 0.006. The regression R-square is 0.528 while hedging effectiveness is 0.442.

Period 2 begins with the price spike. The t-values for the parameters associated with this price spike are 6.89 and -8.50. After accounting for these two outliers, we observe statistically significant changes in the quarterly effects with an increase in the fourth quarter's effect (p-value of 0.047) and a decrease in the second quarter's effect (p-value of 0.015). The change in the serial correlation is not significant (p-value of 0.692) and likewise, the changes in the R$ and ethanol hedge ratios are not significant. Compared to the sample period, the post-sample period displays a larger R^2 (0.960) and a smaller hedging effectiveness (0.197). The R^2 increases because dummy variables account for the 2011 price spike while effectiveness decreases because hedging does not mitigate the variability created by this unanticipated shock.

Table 4 also shows the post-sample results where period 1 estimates are simulated in period 2 (Simulation 2 | 1). This case shows a reduction in R^2 from 0.528 in period 1 to 0.184 in period 2 as the period 1 model does not anticipate the price spike in period 2. Hedging effectiveness also declines from 0.442 in period 1 to 0.135 in period 2. Our analytical framework permits statistical tests of sources of this difference. We first test whether the expectation of the unhedged outcomes changed (i.e., test H_1: $\boldsymbol{\alpha}_2 = \hat{\boldsymbol{\alpha}}_1$). The F-statistic for this test of 33.11 (= [(0.9514-0.0505)/(20-13)]/0.0505/13) has an associated p-value of less than 0.0001 so H_1 is rejected and we conclude that our expectation of the unhedged outcome in period 2 is significantly different from that in period 1.

We next test whether the optimal hedge ratios changed between the two periods (i.e., H_2: $\boldsymbol{\varphi}_2 = \hat{\boldsymbol{\varphi}}_1$) after allowing for the change in the expected unhedged outcomes. For this test, $(\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1)'\mathbf{M}_2'\mathbf{M}_2(\hat{\boldsymbol{\alpha}}_2 - \hat{\boldsymbol{\alpha}}_1) = 0.9514 - 0.0505$ with 7 degrees of freedom and $(\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1)'\mathbf{X}_2'\mathbf{X}_2(\hat{\boldsymbol{\beta}}_2 - \hat{\boldsymbol{\beta}}_1) = 0.8221 - 0.0406$ with 9 degrees of freedom so the resulting F-statistic of 0.687 (= [(0.9514 - 0.0505 -0.8211 + 0.0406)/2] / [(0.8211 + 0.0406) / 11]) with a p-value of 0.523 does not lead to a rejection of H_2. Hence, there is no evidence that a change in the optimal hedge ratios caused the reduction in post-sample hedging effectiveness.

Perhaps the observed decline of hedging effectiveness from the sample period (0.442) to the post-sample period (0.135) occurs because hedging with period 2 estimates would have been less effective in the post sample period. Hence, we test H_3: $\eta_1 = \eta_2$. The test statistic of 0.3288 is asymptotically standard normal and has a p-value of 0.7423 so we cannot reject H_3.[13]

Although we do not reject H_3: $\eta_1 = \eta_2$ we can nonetheless test for constancy of the individual components of effectiveness ($\boldsymbol{\varphi}'\boldsymbol{\Sigma}_{ZZ}\boldsymbol{\varphi}/(\boldsymbol{\varphi}'\boldsymbol{\Sigma}_{ZZ}\boldsymbol{\varphi}+\sigma^2)$) with H_{3a}: $\sigma_1^2 = \sigma_2^2$ and H_{3b}: $\boldsymbol{\varphi}_1'\boldsymbol{\Sigma}_{ZZ,1}\boldsymbol{\varphi}_1 = \boldsymbol{\varphi}_2'\boldsymbol{\Sigma}_{ZZ,2}\boldsymbol{\varphi}_2$. The F statistic for H_{3a} of 1.131 (=0.00417/0.00369) has an associated p-value of 0.435 so we find no evidence of heteroscedasticity. Likewise, the F statistic for H_{3b} of 9.603 (= (0.216624-0.120933) / (0.050531 - 0.040566)) has a p-value of 0.0943 which does not lead to the rejection of the null hypothesis at the five percent significance level.

Summary and Conclusions

This paper has focused hedging effectiveness and is motivated by a lack of context when comparing a sample estimate of hedging effectiveness with an out-of-sample simulation-based estimate of hedging effectiveness. Such a comparison provides no indication of the source of the difference and provides no basis for determining the statistical significance of the difference.

Many researchers consider the regression R^2 as the measure of hedging effectiveness. We have argued that the regression R-square overstates hedging effectiveness whenever the spot price displays systematic effects such as seasonality, serial correlation, or day of the week effects. Similarly, reporting R-square as the estimate of the more narrowly defined hedging effectiveness assumes that other spot price relationships with conditioning variables such as inventory levels or planted acreage are controlled by hedging, when they obviously are not.

Despite the deficiencies of the regression R^2 as a hedging effectiveness estimator, the statistical distributions for R^2 are useful for deriving similar distributions for hedging effectiveness estimators. We discovered that the distributions derived in the mathematical statistics literature did not match those obtained from simulation analysis because classical regression analysis assumes fixed regressors. This assumption is untenable for describing the generation of futures prices as exemplified by a post-sample period with futures prices that differ from the sample period. Hence, we derived the pdf, CDF and moments function for the effectiveness estimator that assumes stochastic futures prices and we verified our functions against simulated random draws.

The pdf, CDF, and moments function are useful but rarely employed. The CDF, for example, permits the construction of confidence intervals for the underlying hedging effectiveness parameter and the moment function permits the determination of the bias associated with a given effectiveness parameter and sample size. We used the moments function to show that the effectiveness estimator is biased, that the bias is small relative to its standard error (table 2), that the bias is positive or negative depending on the sample size and effectiveness parameter (table 2), and that the bias is smaller with stochastic regressors than with fixed regressors (table 1 vs table 2). The finding of negligible bias for the in-sample hedging effectiveness estimator counters one justification for reliance on the post-sample estimator.

With these statistical results, we sought a method for evaluating the difference between the estimated, sample-period, hedging effectiveness (e_1) and post-sample effectiveness estimator ($e_{2|1}$). While the distribution functions apply to effectiveness estimators for the sample (e_1) and the post-sample (e_2) periods, the corresponding functions for the post-sample estimator ($e_{2|1}$)

remain elusive but are composed of mixtures of beta and F distributions. We found that $e_{2|1}$ is unambiguously less than the corresponding e_2. Our approach identifies sources of a difference between the in-sample and the post-sample hedging effectiveness estimators and permits easily testing each source for statistical significance. These tests relate directly to changes in the regression parameters between the sample and the post sample periods.

Finally, we provide an empirical illustration of how using sample-period hedge ratios as a hedging strategy in the post-sample period results in drastically less effective hedging ($e_1 = 0.442$ vs. $e_{2|1} = 0.135$). Our methodology ties the post-sample effectiveness estimator to post-sample structural change in the underlying hedge-ratio model. In our example we determined that the apparent decline in hedging effectiveness was not due to hedge ratios that were clearly suboptimal in the post-sample period but rather was due the an increase in the variability of unhedged outcomes. This finding is clearly a cautionary note that indicates that when evaluating hedges in terms of effectiveness, we must not treat the variability of the unhedged outcome as given. In fact, the stability of the hedge ratios is more important than stability of hedging effectiveness because hedge ratio stability tells us that we are receiving the greatest amount of risk reduction, regardless of the proportion removed. The results presented in this paper generally devalue the post-sample estimator in favor of the in-sample estimator.

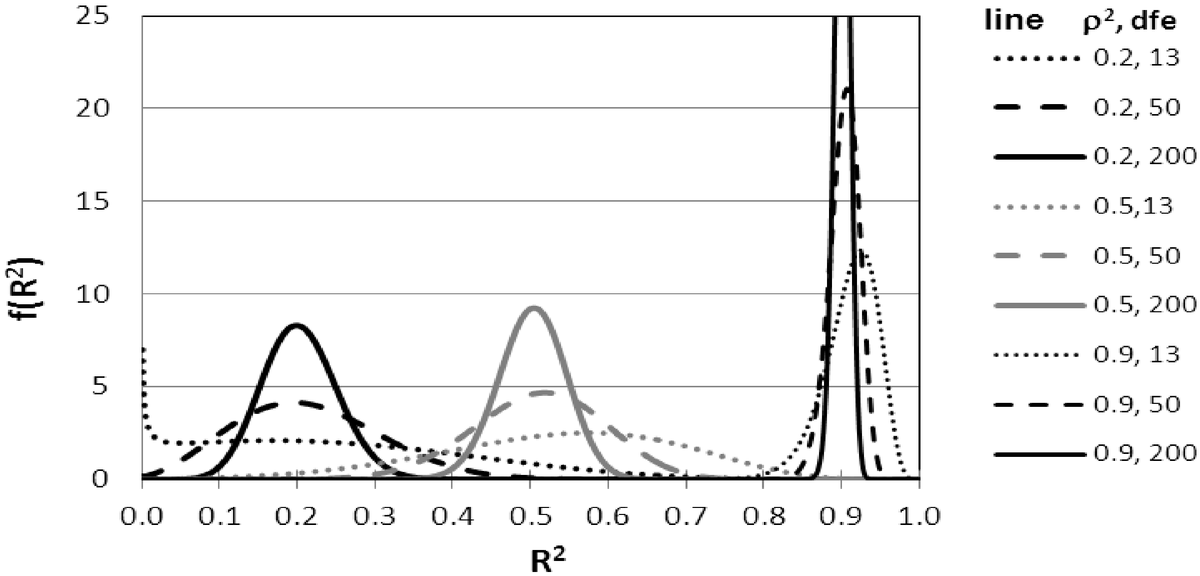

Figure 1. Probability densities for R^2, for various ρ^2 and degrees of freedom.

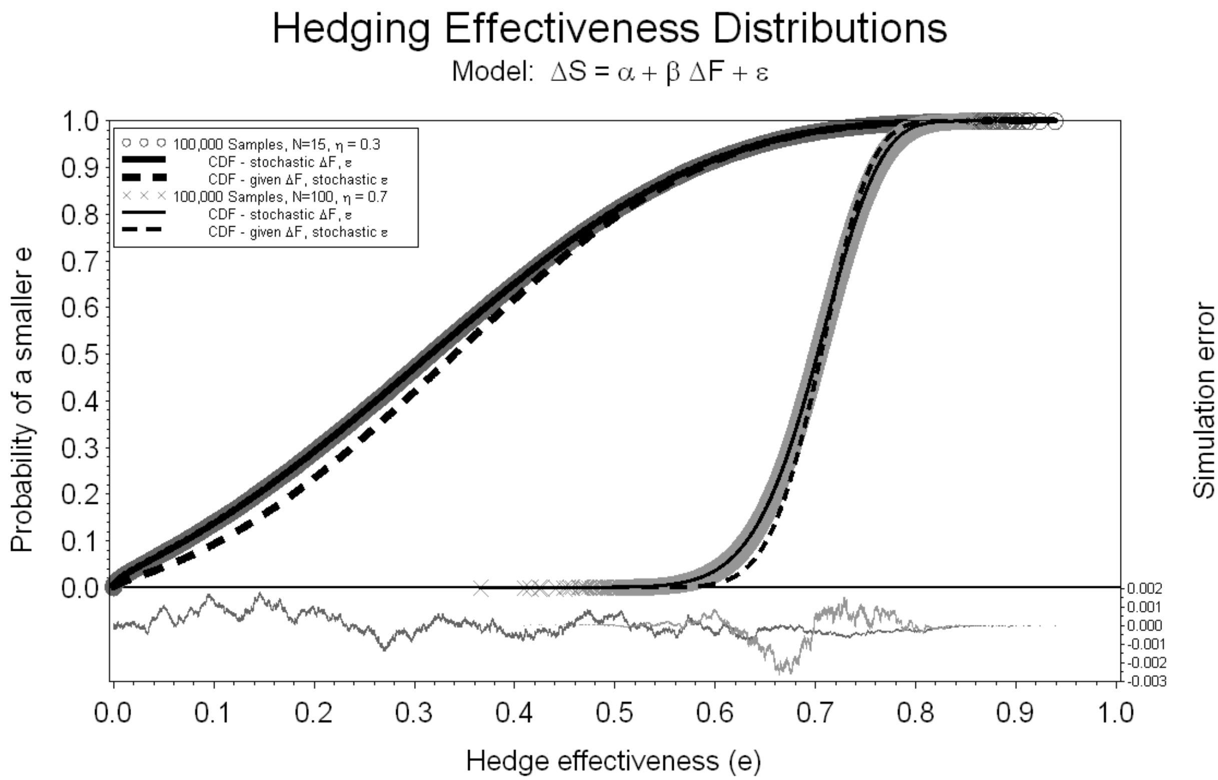

Figure 2. Theoretical versus simulated effectiveness CDFs and error
(N = 15 and η = 0.3 and N = 100, η = 0.7).

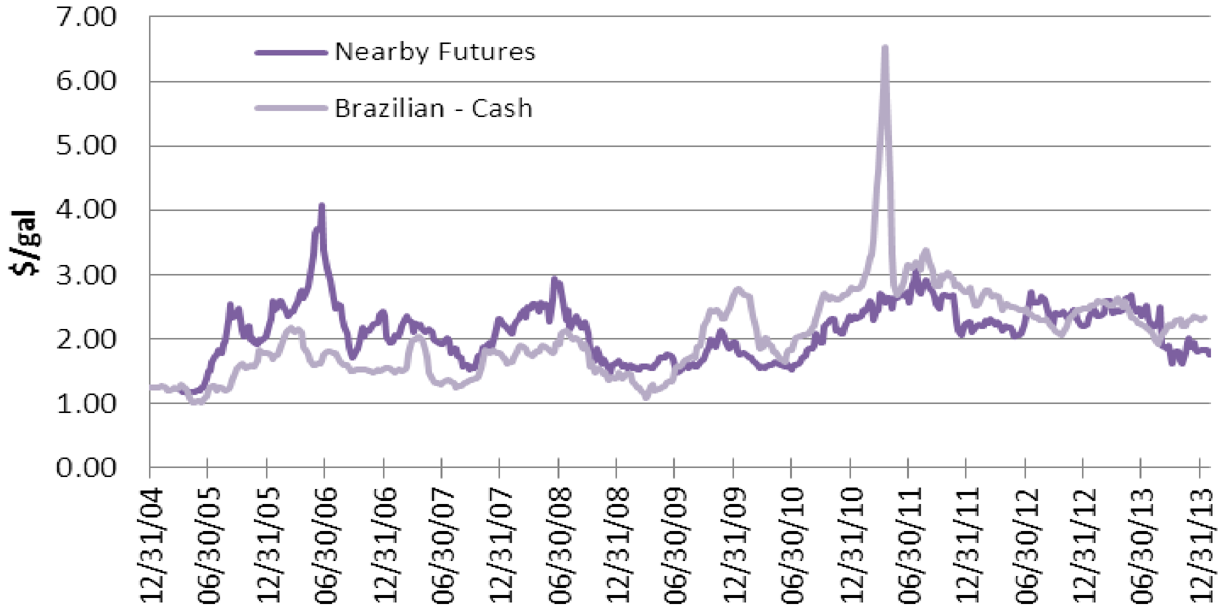

Figure 3. Brazilian anhydrous ethanol cash price and nearby CBOT ethanol futures price.

Table 1. Biases and standard errors for R^2 with non-stochastic regressors.

N	ρ^2									
	0.1		0.3		0.5		0.7		0.9	
	Bias(e)	SE(e)	Bias(e)	SE(e)	Bias(e)	SE(e)	Bias(e)	SE(e)	Bias(e)	SE(e)
15	0.0599	0.1477	0.0443	0.1781	0.0351	0.1558	0.0266	0.1023	0.0117	0.0345
25	0.0348	0.1129	0.0257	0.1394	0.0206	0.1215	0.0157	0.0799	0.0070	0.0273
50	0.0170	0.0791	0.0126	0.0993	0.0102	0.0863	0.0078	0.0569	0.0035	0.0196
100	0.0084	0.0557	0.0062	0.0705	0.0050	0.0611	0.0039	0.0404	0.0017	0.0140
200	0.0042	0.0393	0.0031	0.0499	0.0025	0.0433	0.0019	0.0286	0.0009	0.0099
400	0.0021	0.0278	0.0015	0.0353	0.0013	0.0306	0.0010	0.0202	-0.0004	0.0070

Table 2. Biases and standard errors for hedging effectiveness with stochastic regressors.

N	η									
	0.1		0.3		0.5		0.7		0.9	
	Bias(e)	SE(e)	Bias(e)	SE(e)	Bias(e)	SE(e)	Bias(e)	SE(e)	Bias(e)	SE(e)
15	0.0536	0.1468	0.0238	0.1877	0.0029	0.1826	-0.0078	0.1408	-0.0063	0.0594
25	0.0308	0.1132	0.0130	0.1487	0.0009	0.1420	-0.0048	0.1059	-0.0035	0.0424
50	0.0149	0.0801	0.0060	0.1068	0.0002	0.1004	-0.0024	0.0730	-0.0017	0.0283
100	0.0073	0.0567	0.0029	0.0761	0.0001	0.0709	-0.0012	0.0509	-0.0008	0.0195
200	0.0036	0.0402	0.0014	0.0540	0.0000	0.0501	-0.0006	0.0358	-0.0004	0.0136
400	0.0018	0.0284	0.0007	0.0383	0.0000	0.0354	-0.0003	0.0252	-0.0002	0.0096

Table 3. Data sources and descriptions.

Brazilian Ethanol - Spot	Anhydrous fuel ethanol, \$/liter, weekly, Feb 21, 2000 to present. Source: CEPEA (Center for Advanced Studies on Applied Economics)
Brazilian Real - Spot	Noon buying rates, R\$/\$, daily, Feb 22, 1995 to present. Source: Federal Reserve Bank of New York
Ethanol - Futures	CBOT, 12 maturities/year, \$/gal, daily March 24, 2005 to present. Source: Advanced Commodity Service - barchart.com
Brazilian Real - Futures	CME, 12 maturities/year, \$/R\$, daily, April 2, 2007 to present. Source: Advanced Commodity Service - barchart.com CME, Mar, Jun, Sept, Dec contracts, \$/R\$, daily, Dec 1, 1995 to present. Source: quandl.com

Table 4. Nonlinear OLS estimation results.

Summary of Residual Errors

Period	**Obs**	**SSE(α, φ) (df)**	$\hat{V}(\pi_h)$	**SSE(α) (df)**	$\hat{V}(\pi_u)$	**R-Sq**	**eff**
1 (2005-10)	36	0.120933(29)	0.004170	0.216624(31)	0.006988	0.528	0.442
2 (2011-13)	20	0.040566(11)	0.003687	0.050531(13)	0.003887	0.960	0.197
Total	56	0.161500(40)	0.004038	0.267154(44)	0.006072	0.872	0.395
Simulation 2 | 1	20	0.822090(20)	0.041104	0.951419(20)	0.047571	0.184	0.135

Parameter Estimates

Parameter	**Variable**	**Estimate**	**Approx Std Err**	**Approx t Value**	**Pr > |t|**
Period 1 (2005-2010)					
β_0	Intercept	-0.0373	0.0180	-2.07	0.0445
β_1	1st quarter	0.0706	0.0280	2.53	0.0156
β_2	2nd quarter	0.0523	0.0244	2.14	0.0382
β_3	3rd quarter	-0.0111	0.0273	-0.41	0.6874
β_4	ΔR$ futures	1.1408	0.2826	4.04	0.0002
β_5	Δethanol futures	0.0868	0.0302	2.88	0.0064
ρ	serial corr	-0.5418	0.1542	-3.51	0.0011
Period 2 (2011 - 2013) adjustments					
δ_0	Intercept	0.0615	0.0301	2.05	0.0474
δ_1	1st quarter	-0.0890	0.0507	-1.75	0.0871
δ_2	2nd quarter	-0.1143	0.0450	-2.54	0.0151
δ_3	3rd quarter	-0.0462	0.0614	-0.75	0.4561
δ_4	ΔR$ futures	-0.3571	0.8539	-0.42	0.6781
δ_5	Δethanol futures	0.0227	0.1002	0.23	0.8223
δ_6	price spike part 1	0.6570	0.0953	6.89	<.0001
δ_7	price spike part 2	-0.6591	0.0775	-8.50	<.0001
ϕ	serial corr	0.1569	0.3937	0.40	0.6924

References

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

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

Barchart.com. Brazilian Real Futures Prices. Available at http://acs.barchart.com/. [Accessed March 1, 2014.]

Barchart.com. U.S. Ethanol Futures Prices. Available at http://acs.barchart.com/. [Accessed March 1, 2014.]

Chattamvelli, R. "On the Doubly Noncentral F Distribution." *Computational Statistics and Data Analysis* 20(1995):481-489.

CEPEA (Center for Advanced Studies on Applied Economics). Ethanol Prices. Available at http://cepea.esalq.usp.br/english/ethanol/. [Accessed March 1, 2014.]

Chen, S., C. Lee, and K. Shrestha. "Futures Hedge Ratios: A Review." *The Quarterly Review of Economics and Finance* 43(2003):433-465.

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

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

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

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

Federal Reserve Bank of New York. Brazilian Real Noon Buying Rates. Available at http://www.ny.frb.org/markets/fxrates/historical/home.cfm. [Accessed March 1, 2014].

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

Jelmayer, R. "Brazil Wants Petrobras to Weigh in on Ethanol." *Wall Street Journal* (8 April 2011). Online. Available at http://online.wsj.com/news/articles/SB10001424052748704415104576250742298954426?KEYWORDS=ROGERIO+JELMAYER&mg=reno64-wsj. [Accessed March 2014.]

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

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

Lien, D. "Estimation Bias of Futures Hedging Performance: A Note." *Journal of Futures Markets* 26(2006):835-841.

Lien, D. "Statistical Properties of Post-Sample Hedging Effectiveness." *International Review of Financial Analysis* 16(2007):293-300.

Lindahl, M. "Measuring Hedging Effectiveness with R^2: A Note." *Journal of Futures Markets* 9(1989):469-475.

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.

Muirhead, R. J.. "Estimating a Particular Function of the Multiple Correlation Coefficient." *Journal of the American Statistical Association* 80(1985):923-925.

Myers, R.J., and S.R. Thompson. "Generalized Optimal Hedge Ratio Estimation." *American Journal of Agricultural Economics* 71(1989):858-868.

Papoulis, A. *Probability and Statistics.* Englewood Cliffs, NJ: Prentice-Hall, 1990.

Pe, T., and H. Drygas. "An Alternative Representation of the Noncentral Beta and F Distributions." *Statistical Papers* 47(2006):311-318.

Quandl.com. Brazilian Real Futures Prices, March, June, September, and December Maturities. Available at http://www.quandl.com/futures/brazilian-real-futures. [Accessed March 1, 2014.]

Rahman, S.M., S.C. Turner, and E.F. Costa. "Cross-Hedging Cottonseed Meal." *Journal of Agribusiness* 19(2221):163-171.

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

Snedecor, W.G., and W.G. Cochran. Statistical Methods, 6th ed. Ames, IA: Iowa State University Press. 1967.

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

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

Endnotes

---

1 In considering this methodology, we designate the estimation period as the *in-sample* period and the period where estimation results are used for simulation as the *post-sample* period.

2 Chen, Lee, and Shrestha (2003) review other hedging objectives that lead to other estimation procedures. We restrict our attention to those that that seek to minimize risk, defined as the variance of the hedged portfolio outcomes. Also note that (2) precludes neither heteroscedastic nor serially correlated errors as data transformation can be applied to (2) to impart homoscedasticity and serial independence on the resulting model.

3 λ and ρ^2 are related as follows. $\Sigma_{XY} = T^{-1}(X-\overline{X})'E(Y-\overline{Y}) = T^{-1}(X-\overline{X})'(X-\overline{X})\beta$, $\Sigma_{XX} = T^{-1}(X-\overline{X})'(X-\overline{X})$ and $\sigma_{YY} = T^{-1}\beta'(X-\overline{X})'(X-\overline{X})\beta + \sigma_{\varepsilon\varepsilon}$. By these definitions, $\rho^2 = \lambda\sigma_{\varepsilon\varepsilon}T^{-1}/(\lambda\sigma_{\varepsilon\varepsilon}T^{-1} + \sigma_{\varepsilon\varepsilon})$ and the claimed result follows.

4 The cumulative probability distribution of e also follows from (10b). The second expression can be stated in probability terms as $\Pr\left\{e < \frac{k_2 f_{T-K}^{k_2,\lambda}(\alpha)}{(T-K)+k_2 f_{T-K}^{k_2,\lambda}(\alpha)}\right\} = \alpha$ where α is the probability of a larger F value.

5 The inappropriateness of the fixed-regressor assumption became apparent during the simulation of random samples in an attempt to compare the known CDF of the in-sample effectiveness estimator with unknown post-sample effectiveness estimator. While fixed regressors might be acceptable as conditioning data for the sample period, using these same conditioning data in the post-sample period is clearly unrealistic.

6 As error degrees of freedom approach zero, effectiveness $\rightarrow$ 1.0 but we do not delve into this region.

7 Assumed multivariate normal.

8 Note that the inverse of the covariance matrix ($\mathbf{M_2'M_2}$) in the quadratic form indicates that the proper test is that the period 2 $\boldsymbol{\alpha}$s are equal to the numerical values already obtained in period 1, not the more general test of the hypothesis that $\boldsymbol{\alpha_1} = \boldsymbol{\alpha_2}$. The same cautionary note applies to testing H_2: $\boldsymbol{\varphi_2} = \hat{\boldsymbol{\varphi}}_\mathbf{1}$

9 This statistic compares the period 2 normalized sum of squares attributable to period 1-period 2 differences in all parameters ($\hat{\boldsymbol{\beta}}_i = \begin{bmatrix}\hat{\boldsymbol{\alpha}}_i \\ \hat{\varphi}_i\end{bmatrix}$ i = 1, 2) less the period 2 normalized sum of squares attributable to period 1-period 2 differences in the systematic parameters ($\hat{\varphi}_i$ i = 1, 2).

10 Equivalent to $\begin{bmatrix}\hat{\boldsymbol{\alpha}}_\mathbf{2} \\ \hat{\varphi}_\mathbf{2}\end{bmatrix} = \begin{bmatrix}\hat{\boldsymbol{\alpha}}_\mathbf{1} \\ \hat{\varphi}_\mathbf{1}\end{bmatrix}$.

11 R$ indicates the Brazilian currency, the real, thus avoiding confusion with the word real as in “real price”.

12 The ethanol contract currently matures on the third business day of the month while the R$ currently matures on the last business day of the previous month.

13 We use the respective error degrees of freedom in lieu of N_1-1 and N_2-1 as we apply the test for equality of two correlations (Papoulis).