ARE REGRESSION APPROACH FUTURES HEDGE RATIOS STATIONARY?

ROBERT FERGUSON
DEAN LEISTIKOW*

In contrast to some recent research, this article finds that regression approach futures hedge ratios are stationary. It shows that a previous study's failure to reject the random walk null hypothesis was due to its small sample size and the overlapping hedge ratio calculation approach's bias toward accepting the random walk hypothesis. The impact of overlap on the Dickey-Fuller full model intercept and slope estimates is demonstrated analytically and numerically. Finally, the article shows that out-of-sample hedging performance is not significantly improved by updating the hedge ratios. © 1998 John Wiley & Sons, Inc. Jrl Fut Mark 18:851–866, 1998

INTRODUCTION

The regression approach, introduced by Ederington (1979), is the most widely used approach for calculating futures hedge ratios (HRs).[1]

The authors would like to thank Sharon Gendel, Don Homolka, David Lee, Scott Macdonald, A. Malliaris, Corning Pearson, Ira Seeman, and Jorge Urrutia for their help in the preparation of this article. All errors are the responsibility of the authors.
*Correspondence author, Fordham University–GBA, 113 W. 60th St., New York, NY 10023.
[1]See, for example, Ederington (1979), Hill and Schneeweis (1982), and Grammatikos and Saunders (1983).

- *Robert Ferguson is President of Axiomatic Systems, Inc.*
- *Dean Leistikow is an Associate Professor of Finance at Fordham University.*

The Journal of Futures Markets, Vol. 18, No. 7, 851–866 (1998)

CCC 0270-7314/98/070851-16

Grammatikos and Saunders (1983) were the first to examine HR stability. They studied five currencies using three different tests. Generally, they found that the null hypothesis of HR stability could not be rejected.

Malliaris and Urrutia (1991) (MU) were the first to test whether the HR follows a random walk. They found that futures HRs for direct hedges follow a random walk for currencies and equities.[2] Perfect and Wiles (1994) examined this issue using different data and different tests and state (p. 10) "Overall, the test results support Malliaris and Urrutia's finding."

The conclusion that HRs follow a random walk is disturbing for three reasons. First, it is not generally consistent with the commonly held view that for most assets the spot and futures positions should be opposite in sign and roughly equal in size for direct near-to-maturity hedges.[3] Second, it implies that hedgers need to continuously readjust their hedges (MU). Third, it implies that the HR will wander aimlessly and will reach any specified positive or negative value with probability one given sufficient time.[4] Such aimless wandering is an unrealistic hypothesis for hedge ratios, as MU would likely agree. This suggests possible inadequacies with MU's data or tests.

In marked contrast, McNew and Fackler (1994) tested whether the HR is constant using a nested hypotheses tests approach. Their tests, using live cattle and corn data, support the constant HR hypothesis.

The MU, Perfect and Wiles (1994), and McNew and Fackler (1994) findings could be reconciled in at least two ways. HRs may be stationary for some assets and follow a random walk for others. Or perhaps one of the studies' conclusions is unwarranted.

This article's results strongly reject the random walk hypothesis. It uses the Dickey-Fuller test employed by MU. The reasons MU fails to reject nonstationarity are its small sample size and the overlapping observations used in the HR calculation. Overlap biases the test toward accepting the random walk hypothesis.[5] The impact of overlap on the

[2]In a direct hedge, the spot asset being hedged is also the asset that underlies the futures contract doing the hedging.

[3]For example, see page 14 of the Chicago Board of Trade's (1994) *Commodity Trading Manual*: "Commercial firms note the strong tendency for cash and futures prices to move in the same direction by roughly equal amounts . . . These firms realize that . . . a loss on a cash transaction . . . could be offset . . . in the futures market. This is possible if they initiate a futures position equal and opposite to their cash position."

[4]This is a standard gambler's ruin problem. See Feller (1957), *An Introduction to Probability Theory and its Application*, volume 1, third edition, page 347.

[5]MU also concluded that the in-sample hedging effectiveness (calculated as the R^2 from the regression for which the HR is estimated) follows a random walk using the same methodology that they

Dickey-Fuller full model intercept and slope estimates is demonstrated analytically and numerically. Finally, the paper extends previous research by showing that out-of-sample hedging performance is not significantly improved by updating the hedge ratios.

METHODOLOGY AND RESULTS

To examine the above issues, futures and spot price data are gathered for the deutsche mark, British pound, Japanese yen, and Canadian dollar. The futures trade at the IMM. For each of these currencies, biweekly Friday matched spot and futures closing/settlement price data are gathered.[6] When there is no Friday trading in one or both of the markets (perhaps due to a holiday), the preceding Thursday is used. If there is also no trading on Thursday in one or both of the markets, the following Monday is used. This strategy is employed so that each spot/futures price change pair encompasses about 14 days and each member of a price change pair encompasses the same economic events. Days for which futures prices are nonmarket prices, due to futures price limits, are treated as if they are holidays (see the discussion above).

For each currency, all the futures contracts have approximately between three months plus two weeks and two weeks time-to-maturity.[7] From each asset's price data, 6 sequential biweekly spot/futures price change pairs are calculated for each futures contract and are denoted by the futures contract's time-to-maturity.[8] The sixth biweekly price change of each contract represents the price change for which the futures contract approximately has from 108 to 94 days-to-maturity, the fifth price change represents the price change from 94 to 80 days-to-maturity, and so on, down to the first price change, which represents the price change for which the futures contract has from 38 to 24 days-to-maturity.[9] There is no calendar time overlap between price change pairs; in particular, there is no calendar time overlap between low time-to-maturity price

used to show that the HR follows a random walk. Their hedging effectiveness conclusion is unwarranted for the same reasons that their HR conclusion is unwarranted.

[6]It is unlikely to be important, but MU use Tuesday prices rather than the Friday prices considered in this study.

[7]Futures contracts with between 3 months plus 2 weeks to maturity and 2 weeks to maturity are considered rather than the between 3 months and 0 months to maturity, which were considered by MU, because of the dramatic drop in futures contract liquidity when the futures contract gets within two weeks of its maturity.

[8]Care is taken that the distribution of price change times-to-maturities is the same for each futures contract because of the possibility that the HR is a function of the futures contract times-to-maturities used in the HR calculation (Leistikow (1989, 1993)).

[9]Exceptions to these exact numbers of days-to-maturities are solely due to holidays or price limits as discussed above.

change pairs for one contract and high time-to-maturity price change pairs of the subsequent contract.

March, June, September, and December futures contracts are used for all currencies. The data correspond to the March 1978 through September 1992 futures contracts, a total of 59 contracts per currency.

In each case, the HR is given by β in the following OLS regression:

$$S_t - S_{t-1} = \alpha + \beta(F_t - F_{t-1}) + \varepsilon_t \quad (1)$$

where S_t, F_t, and ε_t are the spot price, futures price, and error terms at time t, respectively. Time $t - 1$ is 2 weeks prior to time t. The HR is estimated from the six time-to-maturity price changes associated with the four previous contracts. In this way, the HRs are estimated using 24 observations.[10]

To test the random walk hypothesis using the Dickey-Fuller test, the unlagged HR is regressed on the lagged HR in the full model regression depicted in eq. (2) below.[11,12]

$$\text{Full Model: } HR_t = b_0 + (b_1 \times HR_{t-1}) + e_t \quad (2)$$

The full model sum of the squared errors is calculated and denoted as SSE_F. According to the random walk hypothesis, the full model regression's intercept should be zero and the lagged HR's coefficient should be one. Imposing these assumptions into the full model yields the reduced model depicted in eq. (3) below.

$$\text{Reduced Model: } HR_t - HR_{t-1} = e_t \quad (3)$$

The reduced model sum of the squared errors (SSE_R) is calculated. The Dickey-Fuller test statistic is given as $F = [(SSE_R - SSE_F)/SSE_F] \times [(n - 2)/2]$, where n is the number of lagged-unlagged HR pair observations in the regressions. The F statistic is the product of two terms. The $[(SSE_R - SSE_F)/SSE_F]$ term is expected to be unrelated to sample size if the HR is stationary, while the $[(n - 2)/2]$ term is clearly an increasing function

[10]This is essentially the same as in MU. MU used 26 observations to calculate the HR.

[11]Unlike MU, a drift term is not included in this regression. The existence of a drift would virtually guarantee the bizarre result that in the long-run the HR would breach any specified positive or negative barrier, depending on the sign of the drift. The drift term was rarely statistically significant in the MU study. The inclusion of a drift term in this paper's regressions should not change the Dickey-Fuller test results.

[12]The HRs are estimates that are treated as ordinary data for regression purposes. Thus the data may not satisfy the regression assumptions. This may change the regression estimates' statistical properties in unknown ways. However, there is no obvious reason why the results are not indicative of the HRs' underlying behavior.

of the sample size. Because each HR calculation uses 4 contracts. HRs for use on contracts 5 through 59 are calculated.[13] The overlapping-window lagged-unlagged HR pairs share 75% of the observations used in their estimations. For example, the HRs for contracts 5 and 6 were estimated sharing the price change data from contracts 2 through 4.

The overlapping-windows approach to estimating HRs provides a larger sample size. This improves the significance test's power, making it more likely to reject the random walk null hypothesis when it is false. On the other hand, it is biased in favor of the random walk null hypothesis because sharing a high percentage of observations causes the sequential HRs to be positively correlated. This makes the Dickey-Fuller full model regression more consistent with the random walk hypothesis than is warranted.

It is shown in the Appendix that the Dickey-Fuller full model b_1 estimate is approximately the proportionate overlap when the true HR is always 1. Thus, the greater the overlap, the more consistent is the Dickey-Fuller full model regression with the random walk hypothesis and the more observations it takes to reject a false random walk null hypothesis.

The bias in favor of the random walk hypothesis caused by the overlapped observations window should be much stronger in 2-week data than in 1-week data because there is greater proportionate overlap (75% vs 50%). Evidence of this is presented below.[14]

The F value is compared to the Dickey-Fuller (1981) critical value (Table IV of their paper) to determine if the HRs are stationary. Essentially this test rejects the null hypothesis of nonstationarity if for the full model (1) the intercept is statistically significant and/or (2) the coefficient of the lagged HR is significantly less than one in absolute value.

First, the Dickey-Fuller test is performed on the restricted futures contract data set that MU used. Their study utilized the June 1980 through December 1988 futures contracts.

As in the MU study, and using the MU futures contracts, the random walk null hypothesis generally is not rejected at the 5% confidence level (Table IA). Similarly, the median (calculated across the four currencies) F value does not reject the random walk null hypothesis at the 10% level. In addition, the median value (calculated across the four currencies) of the b_1 estimate, 0.63, is close to 0.75, the value indicated for a true constant HR of 1 with an overlap of 75%.

[13]Contracts are numbered chronologically. March 1978 is contract number 1 and September 1992 is contract number 59.

[14]With 1 week data, there are twice as many price changes per contract than with 2 week data. Therefore only two contracts are used to estimate a HR. Consider the HRs estimated for contracts 3 and 4. They share the price change data from contract 2. Thus the proportionate overlap is 50%.

TABLE IA

Test That the HR Follows a Random Walk: Using Only the June 1980 to December 1988 Futures Contracts (Those Used in MU)

		Full Model			*Reduced Model*		
Asset	*n*	*b0 Standard Error*	*b1 Standard Error*	*SSEf*	*SSEr*	*(SSEr-SSEf) SSEf*	*F*
		0.2554	0.7275				
D mark	30	0.0336	0.137	0.0249	0.0300	0.2048	2.87
		0.3289	0.6558				
B pound	30	0.0288	0.1577	0.0182	0.0222	0.2198	3.04
		0.3742	0.6121				
J yen	30	0.0304	0.1695	0.0203	0.0253	0.2463	3.47
		0.4898	0.4761				
C dollar	30	0.0418	0.1803	0.0385	0.0535	0.0247	5.47[b]
		0.3516	0.634				
Median	30	0.0320	0.1636	0.0226	0.0277	0.2123	3.26

Notes: The 1%, 5%, and 10% confidence level critical F values for a sample size of 25 are 7.88, 5.18, and 4.12, respectively Superscripts [a], [b], and [c] denote F value significance at the 1%, 5%, and 10% confidence levels, respectively.

In Table IA, the median full model estimates for b_0 and b_1 are 0.35 and 0.63, with standard errors of 0.03 and 0.16, respectively. These are far off the values of 0 and 1 called for by the random walk hypothesis.

The analysis also was performed on 1-week data, where the proportionate overlap is 50%. The median b_1 estimate was 0.46, lower than before and remarkably close to the proportionate overlap.[15] This is consistent with the overlap inducing random walk findings. These findings cannot be explained by MU.

Next, the Dickey-Fuller test is performed on our expanded data set (It begins before and ends after that used by MU.). The random walk null hypothesis is rejected in all four currencies at the 10% confidence level and in two of the currencies at the 5% level (upper panel, Table IB). The median F value is significant at the 5% level. Again, the b_0 and b_1 values are inconsistent with the random walk hypothesis. They appear stable, as they are almost the same for both data sets.

Comparing the HRs generated with overlapping windows for the restricted data set (Table IA) with those for the larger data set (upper panel, Table IB) shows that the small sample size contributes to the failure of the MU study to reject the random walk hypothesis. The median F value is 3.3 and insignificant for the restricted data set (Table IA), whereas it is 4.9 and significant at the 5% level for the full data set (upper panel,

[15]The individual currency results are not presented here, but are available from the authors.

TABLE IB

Test That the HR Follows a Random Walk Using the Full Data Set

		Full Model			Reduced Model		
Asset	n	$b0$ Standard Error	$b1$ Standard Error	$SSEf$	$SSEr$	$(SSEr\text{-}SSEf)$ / $SSEf$	F
Overlapping Approach (Unlagged-Lagged Hr pairs were estimated with a 75% observational overlap)							
		0.2727	0.7187				
D mark	54	0.0293	0.0999	0.0447	0.0515	0.1521	4.00[c]
		0.2903	0.7822				
B pound	54	0.0292	0.0767	0.0442	0.0516	0.1674	4.34[c]
		0.3305	0.6599				
J yen	54	0.0242	0.1035	0.0305	0.0368	0.2066	5.41[b]
		0.5011	0.4745				
C dollar	54	0.0397	0.1126	0.0818	0.1163	0.4218	10.96[a]
		0.3104	0.6893				
Median	54	0.2093	0.1017	0.0445	0.0516	0.187	4.88[b]
Nonoverlapping Approach							
		1.3540	−0.1269				
D mark	13	0.0359	0.2793	0.0142	0.0467	2.2887	12.66[d]
		0.9029	0.0572				
B pound	13	0.0443	0.2324	0.0216	0.0547	1.5324	8.40[d]
		1.0079	−0.0386				
J yen	13	0.0367	0.2977	0.0148	0.0311	1.1014	6.09[d]
		0.9128	0.0378				
C dollar	13	0.0407	0.2556	0.0182	0.0443	1.4341	7.87[d]
		0.9604	−0.0004				
Median	13	0.0387	0.2675	0.0165	0.0455	1.4833	8.14[d]

Notes: The 1%, 5%, and 10% confidence level critical F values for a sample size of 50 are 7.06, 4.86, and 3.94, respectively. The 1%, 5%, and 10% confidence level critical F values for a sample size of 25 are 7.88, 5.18, and 4.12, respectively. Superscripts [a], [b], and [c] denote F value significance at the 1%, 5%, and 10% confidence levels, respectively. The sample sizes upon which the critical values are based are chosen for a lower sample size in order to be conservative.
[d]indicates that no F value significance levels are given for the nonoverlapping approach as the smallest number of observations for which significance levels are given in the Dickey-Fuller tables is 25.

Table IB). The same analysis of 1-week data yielded F values significant at the 5% and 1% levels, respectively.

The median F value for the full data set is higher despite the fact that the $[(SSE_R - SSE_F)/SSE_F]$ term is slightly lower for the full data set than the restricted data set (0.19 vs 0.21). The fact that the $[(SSE_R - SSE_F)/SSE_F]$ term is almost constant when the data set is expanded from 30 to 54 observations means that the increase in F is due to the increase in sample size.[16] It also suggests that further increases in the

[16]The constancy of the $[(SEE_R\text{-}SSE_F)/SSE_F]$ term corresponds to the constancy of the b_0 and b_1 coefficients for the two samples.

sample size can be expected to result in further approximately proportionate increases in the F value.

Dickey-Fuller tests are also performed where the observations used in the estimation of the unlagged-lagged HR pairs are nonoverlapping. The lagged and unlagged HR pairs are chosen in the following way from those HRs that have already been estimated. Because the HRs use the four previous contracts to calculate the HR, the first HR chosen is for use on contract 5 (recall it was estimated using the observations from contracts 1 through 4). The second HR chosen is for use on contract 9 (it was estimated using the observations from contracts 5 through 8). The third HR chosen is for use on contract 13 (it was estimated using the observations from contracts 9 through 12). The rest of the nonoverlapping HRs are calculated accordingly.

The nonoverlapping HR Dickey-Fuller test results are presented in the lower panel of Table IB. No significance levels are available for the nonoverlapping HRs because of the small sample sizes.

Evidence that overlap induces the random walk MU finding is seen by comparing the upper and lower panels in Table IB. The median of the 75% overlap b_1 estimates is 0.69 (upper panel, Table IB). The corresponding nonoverlapping estimate is 0.00 (lower panel, Table IB). (This striking difference cannot be accounted for by MU.). For each currency, the Dickey-Fuller full model b_0, b_1 pair estimates for nonoverlapping data are in line with the strict stationarity specification (i.e. b_0 and b_1 are roughly 1 and 0, respectively). They are inconsistent with the random walk hypothesis (i.e. a b_0 and b_1 of 0 and 1, respectively).

Dickey-Fuller results similar to those in Tables IA and IB were found for the S&P 500, T-bills, and gold. Therefore, the foreign exchange results presented in the article appear to have wide applicability. These noncurrency results are available from the authors upon request.

The upper panel of Table II provides summary statistics on the nonoverlapping approach HRs utilized in the Dickey-Fuller regressions presented in the lower panel of Table IB. The HRs are all very tightly distributed around the mean of about 0.95. This is broadly consistent with the commonly held view that the HR in direct near-to-maturity hedges for most assets should be about 1. The HRs are also subdivided into a first half and second half. The descriptive statistics for the first and second half HRs are presented in the middle and lower panels of Table II, respectively. The greatest absolute value difference between the first half and second half HR means is 0.0413 (C Dollar). The similarity between the HR descriptive statistics for the first and second halves is further evidence that the HRs are stationary.

TABLE II

Nonoverlapping Approach Hedge Ratio Descriptive Statistics (HRs Used in the Regressions Presented in the Lower Panel of Table IB)

Asset	*N*	*Mean*	*Standard Deviation*	*Maximum*	*Minimum*
			Whole Period		
D mark	14	0.9641	0.0347	1.0265	0.8773
B pound	14	0.9481	0.0514	1.0239	0.8291
J yen	14	0.9688	0.0331	1.0278	0.9033
C dollar	14	0.9406	0.0474	1.0154	0.8301
			First Half		
D mark	7	0.9679	0.0243	1.0078	0.9382
B pound	7	0.9470	0.0597	1.0239	0.8291
J yen	7	0.9569	0.0310	1.0040	0.9033
C dollar	7	0.9199	0.0455	0.9646	0.8301
			Second Half		
D mark	7	0.9603	0.0423	1.0265	0.8773
B pound	7	0.9492	0.0414	0.9965	0.8720
J yen	7	0.9808	0.0307	1.0278	0.9303
C dollar	7	0.9612	0.0395	1.0154	0.8950

These stationarity results suggest there may be little benefit from continuously updating the HRs. This was checked by comparing the out-of-sample hedging effectiveness (HE) in two cases. The first HE was obtained from the current HR. The second HE was obtained from a dated HR. Five dated HRs were calculated by ignoring the most recent 2, 4, 6, 8, and 10 years of data.[17] HE is defined as one minus the ratio of the hedge profit variance to the spot price change variance. Hedge profit is defined as the spot price change less the product of the HR and the futures price change. Table III shows the results. The HE is higher for the current HR in only 11 of the 20 cases. The differences are small in all cases. The median differences across currencies are tiny, ranging from 0.24 down to -0.09.

CONCLUSION

In sharp contrast to the MU findings on regression approach futures HR nonstationarity, this article strongly rejects the random walk hypothesis

[17]To illustrate the current and dated HRs, consider the first pair. Each HR is calculated using the 6 different biweekly price change times to maturity for each of 4 quarterly contracts. Thus the dated HR of the first pair is calculated using contracts 1 to 4, while the current HR of the first pair uses contracts 9 to 12 (2 years ahead), 17 to 20 (4 years ahead), 25 to 28 (6 years ahead), 33 to 36 (8 years ahead), and 41 to 44 (10 years ahead). The first out-of-sample hedges are implemented on contracts 13, 21, 29, 37, and 45. The same HR is used for all 6 times to maturity for a contract.

TABLE III

Out-of-Sample Hedging Effectiveness: Current Hedge Ratio vs. Dated Hedge Ratio

Asset	*Minimum Age of Data Used to Calculate Dated Hedge Ratio*	*Number of Hedges*	*% Hedging-Effectiveness (Based on Current Hedge Ratio)*	*% Hedging-Effectiveness (Based on Dated Hedge Ratio)*	*Hedging-Effectiveness Difference (Current Minus Date)*
D mark	2 years	282	97.26	97.43	−0.17
	4 years	234	97.21	97.21	−0.01
	6 years	186	97.23	97.52	−0.29
	8 years	138	97.71	97.70	0.00
	10 years	90	98.27	98.36	−0.09
B pound	2 years	282	97.23	96.80	0.44
	4 years	234	97.08	96.75	0.33
	6 years	186	97.54	96.91	0.62
	8 years	138	97.11	96.74	0.36
	10 years	90	97.39	96.57	0.81
J yen	2 years	282	97.27	97.35	−0.08
	4 years	234	97.43	97.43	0.01
	6 years	186	97.44	97.21	0.23
	8 years	138	97.31	97.20	0.11
	10 years	90	98.62	98.67	−0.05
C dollar	2 years	282	96.27	96.38	−0.11
	4 years	234	97.14	97.29	−0.16
	6 years	186	97.55	97.37	0.18
	8 years	138	97.51	97.07	0.44
	10 years	90	96.74	96.81	−0.06
Median (across currencies)	2 years	282	97.25	97.07	−0.09
	4 years	234	97.17	97.25	0.00
	6 years	186	97.49	97.29	0.20
	8 years	138	97.41	97.14	0.24
	10 years	90	97.83	97.59	−0.06

for a given futures contract time-to-maturity. This article's findings are consistent with the live cattle and corn findings of McNew and Fackler (1994). It shows that the MU study's failure to reject the null hypothesis of nonstationarity was due to its small sample size and the overlapping HR calculation approach's bias towards failing to reject the random walk null hypothesis. The impact of overlap on the Dickey-Fuller full model intercept and slope estimates was demonstrated analytically and numerically. The MU study was ahead of its time. Insufficient data was available for the hypothesis tested.

The article shows that out-of-sample hedging performance is not significantly improved by updating the hedge ratios. To the extent the costs of continuously recalculating HRs for a given time-to-maturity are significant, a hedger should consider scaling back this activity. The costs of recalculating HRs include data gathering, data organizing, and data analysis. Moreover, if a position goes unhedged while the HR is being recalculated, the cost of HR recalculation could be quite high.

These findings should be reassuring to hedgers and researchers. The HR does not wander aimlessly. Furthermore, the HR is not going to approach either positive or negative infinity as allowed by the random walk hypothesis.

The HR stationarity result probably applies to most assets for which futures markets exist for hedging periods of at least 1 week. However, this finding does not suggest that HRs are constant for a given futures contract time-to-maturity. For example, if particular types of important announcements come out on Mondays, it may be the case that the Monday HR would be somewhat different than the HR on other days of the week. Likewise, if the announcement were to come out on the 10th of the month, the HR could be somewhat different for the 10th of the month than it would be for the other days. The HR could be somewhat seasonal or a function of the futures contract's time to maturity, as well. Thus this paper's findings are not inconsistent with either Leistikow (1993 or 1989).

The commonly held view that in direct, near-to-maturity, hedges the HR is approximately 1 for most assets is broadly supported by the data.

MU's conclusion that hedging-effectiveness follows a random walk is unwarranted for the same reasons their HR conclusion is unwarranted.

APPENDIX

This appendix shows how overlap in estimating hedge ratios effects the Dickey-Fuller full model regression estimates.

The context is a series of regressions with overlapping observations. Denote the i'th regression by the subscript i and the j'th observation by the subscript j.

The standard two-variable regression model, in the overlap context, is:

$$Y_{ij} = \alpha_i + \beta_i X_{ij} + u_{ij} \tag{1}$$

Y and X are the changes in the spot and futures prices respectively. The OLS estimate of β_i is:

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n} x_{ij} y_{ij}}{\sum_{j=1}^{n} x_{ij}^2} \tag{2}$$

Lower case letters denote differences from means.

The difference $\hat{\beta}_i - \beta_i$ is:

$$\hat{\beta}_i - \beta_i = \frac{\sum_{j=1}^{n} x_{ij} u_{ij}}{\sum_{j=1}^{n} x_{ij}^2} \equiv \varepsilon_i \tag{3}$$

The series of regressions indicated by eq. (1) are used to estimate a series of hedge ratios. $\hat{\beta}_i$ is the *i'th* hedge ratio estimate. The next step is the following second-stage regression.

$$\hat{\beta}_i = a + b\hat{\beta}_{i-1} + v \tag{4}$$

The null hypothesis is that the true hedge ratios don't change. This is equivalent to a = a, constant, and $b = 0$. Since direct near-to-maturity hedge hedge ratios are about 1, this means that $a \approx 1$.

The OLS estimate of b in eq. (4) is:

$$\hat{b} = \frac{\sum_{j=1}^{n} (\hat{\beta}_{(i-1)j} - \bar{\hat{\beta}}_{(i-1)j})(\hat{\beta}_{ij} - \bar{\hat{\beta}}_{ij})}{\sum_{j=1}^{n} (\hat{\beta}_{(i-1)j} - \bar{\hat{\beta}}_{(i-1)j})^2} \tag{5}$$

This estimator's expected value is approximately (for large enough samples):[18]

$$E(\hat{b}) \approx \frac{\mathrm{cov}(\beta_i, \beta_{i-1})}{\mathrm{var}(\beta_{i-1})} \tag{6}$$

Assuming independence between x_{ij} and u_{kl}, and stationarity:

$$E(\hat{b}) \approx \frac{\mathrm{cov}(\varepsilon_i, \varepsilon_{i-1})}{\mathrm{var}(\varepsilon_{i-1})} = \frac{\mathrm{cov}(\varepsilon_i, \varepsilon_{i-1})}{\sigma_\varepsilon^2} \tag{7}$$

Eq. (3) implies:

[18]Note that this does not imply that $E(\hat{b}) = b$. In general, $\hat{b}$ is biased downward in magnitude because the independent variable in the second stage regression, $\hat{\beta}_{i-1}$, is measured with error.

$$\sigma_\varepsilon^2 = E\left[\frac{\sum_{j=1}^{n}\sum_{k=1}^{n} x_{ij}x_{ik}u_{ij}u_{ik}}{\sum_{l=1}^{n}\sum_{m=1}^{n} x_{il}^2 x_{im}^2}\right] \tag{8}$$

$$\sigma_\varepsilon^2 = E\left[\sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{x_{ij}x_{ik}}{\sum_{l=1}^{n}\sum_{m=1}^{n} x_{il}^2 x_{im}^2}\right)u_{ij}u_{ik}\right]$$

$$= E\left(\sum_{j=1}^{n}\sum_{k=1}^{n} c_{iijk}u_{ij}u_{ik}\right) \tag{9}$$

Denote averages by an upper bar. Then by the independence and stationarity assumptions:

$$\sigma_\varepsilon^2 = n\bar{c}_{iijj}\sigma_u^2 \tag{10}$$

Now evaluate $\text{cov}(\varepsilon_i, \varepsilon_{i-1})$ similarly.

$$\text{cov}(\varepsilon_i, \varepsilon_{i=1}) = E\left[\frac{\sum_{j=1}^{n}\sum_{k=1}^{n} x_{ij}x_{(i-1)k}u_{ij}u_{(i-1)k}}{\sum_{l=1}^{n}\sum_{m=1}^{n} x_{il}^2 x_{(i-1)m}^2}\right] \tag{11}$$

$$\text{cov}(\varepsilon_i, \varepsilon_{i=1}) = E\left[\sum_{j=1}^{n}\sum_{k=1}^{n}\left(\frac{x_{ij}x_{(i-1)k}}{\sum_{l=1}^{n}\sum_{m=1}^{n} x_{il}^2 x_{(i-1)m}^2}\right)u_{ij}u_{(i-1)k}\right]$$

$$= E\left(\sum_{j=1}^{n}\sum_{k=1}^{n} d_{i(i-1)jk}u_{ij}u_{(i-1)k}\right) \tag{12}$$

Assume that p of the observations overlap. Then the observations from $n - p + 1$ to n of regression $i - 1$ overlap with the first p observations of regression i. Thus there will be only p terms that contribute to the expectation.

$$\text{cov}(\varepsilon_i, \varepsilon_{i-1}) = p\bar{d}_{iijj}\sigma_u^2 \tag{13}$$

Now solve eq. (10) for σ_u^2 and substitute in eq. (13).

$$\sigma_u^2 = \frac{\sigma_\varepsilon^2}{n\bar{c}_{iijj}} \tag{14}$$

$$\text{cov}(\varepsilon_i, \varepsilon_{i=1}) = \left(\frac{p}{n}\right)\left(\frac{\bar{d}_{iijj}}{\bar{c}_{iijj}}\right)\sigma_\varepsilon^2 \tag{15}$$

Finally:

$$E(\hat{b}) \approx \left(\frac{p}{n}\right)\left(\frac{\bar{d}_{iijj}}{\bar{c}_{iijj}}\right) \tag{16}$$

Eq. (16) shows that full overlap leads to an expectation of 1 and no overlap leads to an expectation of 0. Moreover, $\bar{d}_{iijj}$ and $\bar{c}_{iijj}$ are drawn from two populations with the same mean. Indeed, $\bar{d}_{iijj}$ is the average of a subset of the numbers used to compute $\bar{c}_{iijj}$. Thus their expected ratio is approximately 1, and $E(\hat{b})$ is roughly the proportional overlap.

The behavior of $E(\hat{b})$ implies, through OLS, that $E(\hat{a})$ is approximately 0 with full overlap and approximately 1 with no overlap. Its expectation is rising between these limits. Roughly speaking, $E(\hat{a})$ is 1 less the proportional overlap.

A monte carlo simulation was performed to illustrate these analytical results.

The simulation begins by specifying the true parameters for the regression of change in spot price against change in futures price. The assumptions are an intercept of 0, a slope (the hedge ratio) of 1.0, a \$2.00 standard deviation for the change in futures price, a \$0.40 residual standard deviation, and normal distributions. The standard deviation assumptions imply a true R squared of about 0.96. To keep things simple, the mean change in futures price was assumed to be zero.

Draws from the standard normal distribution, together with the assumed parameters, were used to generate 100 observations for the change in futures price and the change in spot price for use in the first regression. The nonoverlap observations for the second regression were generated similarly. The assumed overlap period observations for the second regression come from the first regression's data set.

Regressions were run on the two data sets to obtain two sets of intercepts and slopes. The two slopes, estimated hedge ratios, represent a single observation for a second stage regression of the second slope against the first slope. A total of 100 second stage observations were generated by repeating the process. The entire process was repeated for proportional overlaps from 0 to 1 with a step size of 0.05.

The monte carlo simulation does not mimic exactly the Dickey-Fuller full model regression procedure. The data generated for the simulation's second stage regression do not correspond exactly to the second stage

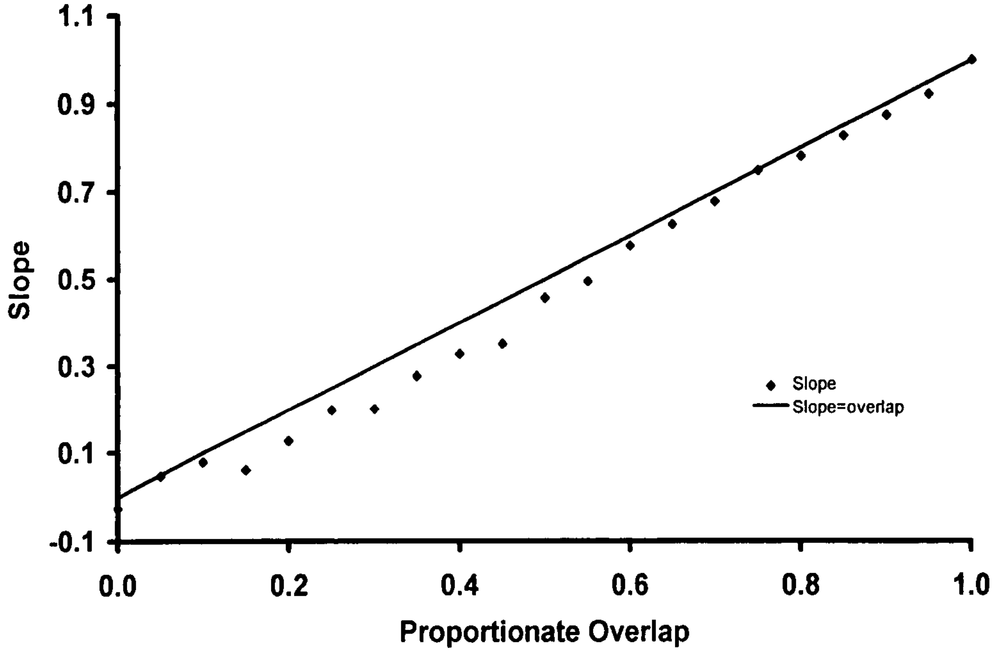

FIGURE 1
Slope for regression of hedge ratio on previous hedge ratio.

regressions run in the paper. In reality, the total number of first-stage observations is fixed and increasing the overlap increases the number of second-stage observations. For simplicity, the number of second stage observations in the simulation is fixed. This introduces second-stage regression sampling variation that is not present in reality. Thus the simulation probably understates the closeness of the relationship in eq. (16).

Figure 1 shows how the simulated second-stage regressions' slopes change with proportionate overlap. The points show the estimated slopes. The line shows the proportionate overlap. The points lie somewhat below the line except as the proportionate overlap approaches 1. Sampling variation, estimation error, and errors in variables vanish at a proportionate overlap of 1. This explains the convergence of the line and the points. Errors in variables tend to reduce the magnitude of an estimated slope. This explains at least part of the tendency of the points to be below the line except near a proportionate overlap of 1. The negative intercept probably reflects sampling variation and estimation error.

All things considered, the line is a fairly good approximation to the points. The estimated slope in the simulated second stage regression is approximately the proportionate overlap.

BIBLIOGRAPHY

Chicago Board of Trade (1994): *Commodity Trading Manual.* Chicago: Chicago Board of Trade.

Dickey, D., and Fuller, W. (1981): "Likelihood Ratio Statistics for Autoregressive Time Series With A Unit Root," *Econometrica,* 49:1057–1072.

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

Feller, W. (1957): *An Introduction to Probability Theory and its Applications* (vol. 1, 3rd edition). New York: Wiley.

Grammatikos, T., and Saunders, A. (1983): "Stability and Hedging Performance of Foreign Currency Futures," *The Journal of Futures Markets,* 3:295–305.

Hill, J., and Schneeweis, T. (1982): "The Hedging Effectiveness of Foreign Currency Futures," *Journal of Financial Research,* 5:95–104.

Leistikow, D. (1989): "Announcements and Futures Price Variability," *The Journal of Futures Markets,* 9:477–486.

Leistikow, D. (1993): "Impacts of Shifts in Uncertainty on Spot and Futures Price Change Serial Correlation and Standardized Covariation Measures," *The Journal of Futures Markets,* 13:873–887.

Malliaris A., and Urrutia J. (1991): "Test of Random Walk of Hedge Ratios and Measure of Hedging Effectiveness for Stock Indexes and Foreign Currencies," *The Journal of Futures Markets,* 11:55–68.

McNew, K., and Fackler, P. (1994): "Nonconstant Optimal Hedge Ratio Estimation and Nested Hypotheses Tests," *The Journal of Futures Markets,* 14:619–635.

Perfect, S., and Wiles, K. (1994): "New Tests of Randomness in Futures Hedge Ratios," working paper presented at the 1994 FMA meetings.