Estimating the Extended Mean-Gini Coefficient for Futures Hedging

Donald Lien
Xiangdong Luo

INTRODUCTION

It is by now well recognized that mean-variance (MV) analysis is only applicable to certain restrictive circumstances. The method, however, remains popular in the literature. For example, the well-known Ederington (1979) hedge ratio is derived within this framework. Any alternative approach to replace the MV analysis must maintain the simplicity while improving its consistency with stochastic dominance criteria. As demonstrated in Yitzhaki (1982, 1983), the extended mean-Gini difference constructs such an approach. Upon adopting a specific member from the class of extended mean-Gini differences, Chung, Kwan, and Yip (CKY; 1990) derive optimal hedge ratios for several futures contracts. A complete analysis is provided in Kolb and Okunev (KO;1992). They find two peculiar results, both pointing to the instability (or sensitivity) of hedge ratios following the extended mean-Gini approach.

First, Kolb and Okunev argue that the optimal hedge ratio minimizing the extended mean-Gini coefficient is a step function of the underlying risk aversion parameter. It is likely that a small change in the risk aversion parameter leads to a large increment in hedge ratios. Thus one needs to characterize a hedger (in terms of risk aversion parameter) very carefully. A small error may create completely misleading hedge ratios. Second, upon applying a 50-day moving window to the whole sample data, they find that the hedge ratio is a step function of consecutive windows. A small change in the sample employed for estimation purposes, therefore, might produce rather diverse hedge ratios.

The above two properties are carefully analyzed in this article. Beginning with a discussion of the methodology employed in CKY (1990) and KO (1992), it is claimed that, in contrast to the KO conclusion, the hedge ratio is more likely a smooth function of the risk aversion parameter. Empirical studies using weekly S&P 500 price data support this claim. Moreover, it is shown that the hedge ratio is a monotonic function of the risk aversion parameter. That is, the hedge ratio either increases or decreases continuously with increasing risk aversion; but it never increases for some ranges while decreasing

We are thankful to De-Min Wu and two anonymous referees for helpful comments and suggestions. A.G. Malliaris at Loyola University of Chicago kindly provided the data for this research. Of course, we are responsible for any remaining errors.

Donald Lien is a Professor in the Department of Economics at the University of Kansas.

Xiangdong Luo is a Doctoral Candidate at the University of Kansas.

The Journal of Futures Markets, Vol. 13, No. 6, 665–676 (1993)

CCC 0270-7314/93/060665-12

for others. This property is demonstrated in the KO study without notice. Using weekly data, this study derives a step function for hedge ratios as the data window moves when the risk aversion parameter is large. An explanation based upon shifts in minimal rank points is provided. More specifically, in this case the hedge ratio is mostly determined by the minimal rank point. Any change in the minimal rank point requires the hedge ratio to change by at least a certain amount. On the other hand, it is also found that there are generally two local minima. When one local minimum overtakes the other, the resulting shift in hedge ratio is sharp. Both factors contribute to the instability of the hedge ratio. In the case of a small risk aversion parameter, it is shown that the hedge ratio is robust to the data windows.

The above conclusions are derived using the KO method of estimating the extended mean-Gini difference. One defect of the approach is that the probability distribution function is estimated by a discrete function, the empirical distribution function. As a consequence, the estimated EMG (extended mean-Gini) coefficient is not differentiable, rendering the conventional first-order condition approach inapplicable. A better alternative relies upon nonparametric density estimation via kernel methods, which produces a smooth estimator for the EMG coefficient. It turns out that for the sample data used in this study, the above results are all retained. The properties discussed in this article are robust to distribution estimation methods.

THE BASIC APPROACH

Adopting the notation from the KO article, the EMG coefficient for a hedger is given by:

$$\Gamma(v) = -v \operatorname{cov}(R_p, (1 - F(R_p))^{v-1}) \tag{1}$$

where v is the risk aversion parameter; $R_p = R_s + xR_f$ is the return on the market portfolio with x being the hedge ratio, and R_s (resp. R_f) being the return on the spot (resp. futures) market.[1] Also, $\operatorname{cov}(\cdot, \cdot)$ denotes the covariance operator and $F(\cdot)$ denotes the probability distribution of R_p. The true distribution function is, of course, unknown. A sample analog is adopted to estimate $F(\cdot)$. More precisely, given observations of R_{st} and R_{ft} ($t = 1, \ldots, T$), and given any hedge ratio x, $F(R_{pt})$ is estimated by $\hat{F}(R_{pt})$, the rank of R_{pt} ($= R_{st} + xR_{ft}$) divided by the number of observations.[2] Similarly, the covariance term in eq. (1) is estimated by the corresponding sample analog. Thus, for empirical implementation purposes, the extended mean-Gini coefficient is rewritten as:

$$S\Gamma(v) = -(v/T)\left\{\sum_{t=1}^{T} R_{pt}(1 - \hat{F}(R_{pt}))^{v-1} - \left(\sum_{t=1}^{T} R_{pt}/T\right)\left(\sum_{t=1}^{T}(1 - \hat{F}(R_{pt}))^{v-1}\right)\right\} \tag{2}$$

An iterative process is applied to find the minimum of $S\Gamma(v)$ over the choice of x. Suppose there is no tie in ranking: $\hat{F}(R_{pt}) \neq \hat{F}(R_{pt'})$ whenever $t' \neq t$, then eq. (2) is differentiable with respect to x and the optimal hedge ratio satisfies the first-order condition $\partial S\Gamma(v)/\partial x = 0$. However, the presumption is invalid. That is, at the optimal

[1]Strictly speaking, R_f (defined as the percentage of change in futures prices) cannot be termed as a return because no initial outlay is required when taking a futures position. The terminology adopted here follows convention.

[2]This method is correct when there is no tie in ranking. Suppose that two data points obtain the same market return (for a given hedge ratio) with n other data points generating smaller returns. Then, the cumulative density at both points should be estimated by $(n + 2)/T$. Empirical estimation rarely incurs ties in ranking; and, therefore, the above modification is generally not applicable.

hedge ratio, there are ties in ranking, at least analytically.[3] In terms of computation, the limitation of convergence criteria prevents a tie in the solution. The implication is discussed later.

Since the hedge ratio is of central interest, replacing EMG by its estimates may not be the most appropriate approach. Alternatively, using eq. (1) (which is continuously differentiable for a continuous differentiable distribution function) to minimize the EMG, the optimal hedge ratio must satisfy the following first-order condition:

$$-\text{cov}(R_f, (1 - F(R_p))^{\nu-1}) + \text{cov}(R_p, (\nu - 1)(1 - F(R_p))^{\nu-2} f(R_p) R_f) = 0 \quad (3)$$

where $f(\cdot)$ denotes the probability density function. An estimate for the LHS of eq. (3) is

$$\begin{aligned} Q(\nu, x) = {} & \sum_{t=1}^{T} R_{ft}(1 - \hat{F}(R_{pt}))^{\nu-1} - \left(\sum_{t=1}^{T} R_{ft}/T\right)\left(\sum_{t=1}^{T}(1 - \hat{F}(R_{pt}))^{\nu-1}\right) \\ & + (\nu - 1) \cdot \left\{\sum_{t=1}^{T} R_{pt} R_{ft} \hat{f}(R_{pt})(1 - \hat{F}(R_{pt}))^{\nu-2} \right. \\ & \left. - \left(\sum_{t=1}^{T} R_{pt}/T\right)\left(\sum_{t=1}^{T} R_{ft} \hat{f}(R_{pt})(1 - \hat{F}(R_{pt}))^{\nu-2}\right)\right\} = 0 \quad (4) \end{aligned}$$

where $\hat{f}(\cdot)$ denotes the estimate of the density function. For a given ν, the optimal hedge ratio is derived by an iterative process to solve the eq. $Q(\nu, x) = 0$. Corresponding to the estimate $\hat{F}(\cdot)$, which may be assumed to arise from histograms [see Silverman (1986)], $\hat{f}(\cdot)$ is chosen to be the number of observations with the same value divided by the total number of observations. T.

To sum up, the hedge ratio is determined by an unknown quantity within the mean-Gini framework. Eqs. (2) and (4) represent two different approaches to resolve the problem upon estimating the objective function and the first-order condition, respectively. While previous research and the current article adopt the first method to derive optimal hedge ratios, the second method is taken to establish the monotonicity of hedge ratios. Empirical studies indicate the two methods yield nearly identical results.

THE EFFECT OF RISK AVERSION PARAMETER

As stated above, the KO approach searches for the minimum of $S\Gamma(\nu)$. In case of no ties in ranking, $S\Gamma(\nu)$ is continuously differentiable with respect to both x and ν. Provided the solution renders convexity of $S\Gamma(\nu)$ in a small neighborhood, the implicit function theorem implies that, contrary to the KO claim, the optimal hedge ratio is a continuous function of the risk aversion parameters, ν. While theoretically the solution must dictate ties in ranking, computational results do not necessarily follow the same pattern due to the existence of the convergence criteria that stops the iterations before reaching the theoretical solution. It is then more likely that the optimal hedge ratio is a continuous function of ν. To help illustrate this conjecture, the S&P 500 weekly price data for

[3] Consider the case that the ranking of data points is fixed. Upon taking partial derivative of $S\Gamma(\nu)$ with respect to x,

$$\partial S\Gamma(\nu)/\partial x = -(\nu/T)\sum_{t=1}^{T}\left(R_{ft} - \left(\sum_{t=1}^{T} R_{ft}/T\right)\right)(1 - \hat{F}(R_{pt}))^{\nu-1}$$

Because the ranking is fixed, so is $\hat{F}(R_{pt})$. The above equation implies that the optimal hedge ratio increases when $\partial S\Gamma(\nu)/\partial x > 0$ and decreases when $\partial S\Gamma(\nu)/\partial x < 0$. The movement continues until the ranking is at the margin of the breakdown, that is, at least two points share the same rank.

the nearby contract are applied. Tuesday closing prices are used. Monday or Wednesday closing prices are substituted when Tuesday closing prices are missing. The sample period extends from January 1, 1984 to December 27, 1988 for a total of 260 observations.

Strictly speaking, continuity is a local property. Consideration of a discrete series of risk aversion parameter values will not help in verifying or rebutting the conjecture that the hedge ratio is a continuous function of the very parameter. Nonetheless, numerical results reported in Table I appear to be exceedingly smooth, in support of the conjecture.[4] When $2 \leq v \leq 10$, an increment of one induces an increase in the optimal hedge ratio by less than 0.015, or less than 1.7%. In cases where $v > 10$, the impacts of increments are even smaller. Table I also shows that the optimal hedge ratio decreases with increasing risk aversion parameter, in terms of absolute value. This result contradicts the conclusion of Kolb and Okunev (1992), in which they consider the daily S&P 500 price data. The increment of hedge ratios decreases as v increases such that the optimal hedge ratio remains constant when v is very large. For comparison, a conventional approach through least squares estimation is applied to determine the Ederington hedge ratio, resulting in a number of -0.8171, which is fairly close to the case of $v = 9$. To repeat, the above hedge ratios are derived by minimizing the estimated mean-Gini coefficient. An alternative approach relies upon the estimated first-order condition, i.e., eq. (4). To compare the two approaches, $Q(v, x)$ is evaluated for a given v where $x^*(v)$ corresponds to the hedge ratios derived by the former approach. As shown in Table I, $Q(v, x^*(v))$ is nearly zero for all v considered. Thus the two approaches are consistent.

Table I
OPTIMAL HEDGE RATIOS FOR VARIOUS RISK AVERSION PARAMETER VALUES

v	$x^*(v)$	$Q(v, x^*(v))$
2	−0.8822	4.7×10^{-7}
3	−0.8789	2.0×10^{-6}
4	−0.8643	6.9×10^{-7}
5	−0.8530	3.1×10^{-7}
6	−0.8427	7.0×10^{-7}
7	−0.8341	1.4×10^{-6}
8	−0.8251	5.2×10^{-7}
9	−0.8182	7.2×10^{-7}
10	−0.8098	-1.3×10^{-6}
20	−0.7725	-3.3×10^{-6}
30	−0.7599	3.1×10^{-6}
40	−0.7524	9.2×10^{-6}
50	−0.7435	2.7×10^{-6}
100	−0.7208	1.8×10^{-5}
150	−0.7185	-5.1×10^{-5}

[4]Only the case when $v \geq 2$ is considered. In fact, when $v < 2$, the hedger may be termed as a risk lover, which is at odds with the purpose of hedging.

Let $SE(\cdot)$ and $S\,\mathrm{cov}(\cdot,\cdot)$ denote the sample expectation and sample covariance operators, respectively. The Appendix provides the following result:

$$Q(\nu + 1, x^*(\nu)) \approx -SE[(1 - \hat{F}(R_{pt}))^{\nu-1}] \cdot S\,\mathrm{cov}(R_{ft}, \Delta(1 - \hat{F}(R_{pt}))) \qquad (5)$$

where "$\approx$" means "approximately equals." Regardless of the value of ν, $SE[(1 - \hat{F}(R_{pt}))^{\nu-1}]$ is always positive. The sign of $Q(\nu + 1, x^*(\nu))$ is therefore opposite to the sign of $S\,\mathrm{cov}(R_{ft}, 1 - \hat{F}(R_{pt}))$. Provided the second order condition of the original minimization problem is satisfied, (i.e., $\partial^2\Gamma(\nu)/\partial x^2 > 0$), $Q(\nu + 1, x)$ is expected to be increasing in the neighborhood of $x^*(\nu + 1)$. Consequently, $x^*(\nu + 1)$ is greater (resp. less) than $x^*(\nu)$ when $S\,\mathrm{cov}(R_{ft}, 1 - \hat{F}(R_{pt}))$ is positive (resp. negative). Note that R_{pt} is defined as $R_{st} + x^*(\nu)R_{ft}$, and hence $\hat{F}(R_{pt})$ is implicitly a function of ν. Nonetheless, as long as the sign of $S\,\mathrm{cov}(R_{ft}, \Delta(1 - \hat{F}(R_{pt})))$ remains unchanged, which it will hold for a sufficiently large interval of x, the relationship between $x^*(\nu + 1)$ and $x^*(\nu)$ is unidirectional. That is, the optimal hedge ratio tends to be a monotonic function of the risk aversion parameter. Empirical evidence from Kolb and Okunev (1992) demonstrates that, as the risk aversion parameter increases, the hedge ratio continuously decreases for copper and German mark markets while it continuously increases for gold, corn, and S&P 500 markets, in terms of absolute value.

Eq. (5) also indicates that $Q(\nu + 1, x^*(\nu))$ is nearly zero when ν is large, implying $x^*(\nu + 1) = x^*(\nu)$. That is, the optimal hedge ratio approaches a constant term when the risk aversion parameter becomes very large. Empirical support of this result is found in Table I and the KO study.

THE EFFECT OF THE DATA WINDOW

Following Kolb and Okunev, moving data windows are applied to examine the stability of EMG hedge ratios. Specifically, a window contains 50 data points. As the window moves across the observations, the most obsolete data point is replaced by a new data point. In total, this approach produces 210 hedge ratios; the stability is examined by comparing the magnitude of change in the consecutive windows. Table II reports the empirical results using the EMG approach with two risk aversion parameter values, $\nu = 2$ and $\nu = 100$. For comparison purposes, the Ederington hedge ratios calculated by the regression models are also reported. Note that the instability of a hedge ratio series may be caused either by structural change in the underlying sample or by the method that generates the series.

The Ederington hedge ratio appears to be stable, with 201 out of 209 possible changes not exceeding 0.02. Since the ratio generally lies between 0.80 and 0.90, a change of 0.02 in hedge ratios implies a relative change of less than 2.5 percentage points. Similar results apply to the EMG ratio when $\nu = 2$. In fact, because each EMG ratio generally exceeds the corresponding Edgerington ratio, the conclusion is strengthened when considering relative changes. The above observations indicate that the sample employed does not contain any structural change; also, both mean-variance and extended mean-Gini difference (with $\nu = 2$) approaches are insensitive to data window selections.

The case for $\nu = 100$ is completely different. While 161 consecutive changes exhibit a small magnitude of less than 0.05, 25 others exceed 0.105. It appears that, in this case, the EMG method tends to resist the impacts of changes in data points. Nonetheless, if the hedge ratio changes, the magnitude is likely to be large. The two factors, inertia and big

Table II
ABSOLUTE CONSECUTIVE CHANGES IN HEDGE RATIOS

Range[a]	Ederington	EMG ($\nu = 2$)	EMG ($\nu = 100$)
(0, 0.005)	151	131	161
(0.005, 0.01)	37	48	3
(0.01, 0.02)	13	20	2
(0.02, 0.03)	1	5	5
(0.03, 0.04)	1	2	4
(0.04, 0.05)	0	1	2
(0.05, 0.06)	2	0	0
(0.06, 0.07)	0	0	4
(0.07, 0.08)	1	0	1
(0.08, 0.09)	1	1	1
(0.09, 0.105)[b]	2	1	1
(0.105, 0.15)	0	0	12
(0.15, 0.20)	0	0	8
(0.20, 0.30)	0	0	3
(0.30, 0.40)	0	0	1
(0.40, 0.495)[c]	0	0	1

[a]The two numbers (a, b) indicate the lower and upper bounds on the changes. For example, from the second row, the number of (absolute) consecutive changes in Ederington hedge ratios, EMG ratios with $\nu = 2$ and EMG ratios with $\nu = 100$ that lie between 0.005 and 0.10 are 37, 48, and 3 respectively.

[b]The maximal (absolute) consecutive changes in the case of Ederington and EMG ($\nu = 2$) are 0.1044 and 0.1041, respectively.

[c]The maximal (absolute) consecutive change in the case of EMG ($\nu = 100$) is 0.4946.

changes (if necessary), contribute to the instability of the EMG hedge ratio. As indicated in the KO article, the instability makes the hedge ratio a widely swung step function of data windows. Note that this empirical study indicates the instability result may apply only to the case of large values of the risk aversion parameters. In fact, when ν is very large, eq. (2) is approximated by:

$$AS\Gamma(\nu) = -(\nu/T)(1 - 1/T)^{\nu-1}\left\{R_{pm} - \left(\sum_{t=1}^{T} R_{pt}/T\right)\right\} \tag{6}$$

where $R_{pm} \leq R_{pt}$ for every $t = 1, \ldots, T$. Suppose both spot and futures markets are unbiased so that the average market return, $\sum_{t=1}^{T} R_{pt}/T$, is nearly zero. Eq. (6) indicates that the expected mean-Gini difference is merely determined by the minimum of the return from market portfolios. Moreover, the optimal hedge ratio is a solution to the following problem:

$$\max_x\{\min_t(R_{st} + xR_{ft})\} \tag{6$'$}$$

As pointed out in the KO study, the EMG ratio is a maximum strategy when ν is large. This property accounts for the instability of EMG ratios.[5] A diagram helps illustrate this conclusion.

[5]A referee correctly points out that the variability of the EMG coefficient for high values of ν is not a difficulty with the EMG approach. Rather, if the trader's preferences scheme is extreme in some sense, small changes in the environment may cause large changes in his behavior.

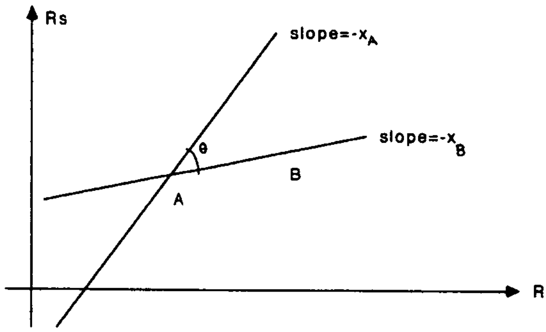

Figure 1
Changes in EMG Hedge Ratios.

In Figure 1, let A denote the data point with minimal return at the EMG hedge ratio, x_A. Assume sampled data are changed, resulting in a new EMG hedge ratio, x_B. The corresponding data point with minimal return is B. To establish this result, $-x_B$ must be smaller than $-x_A - \theta$, where θ is the angle between the original market portfolio line and the line connecting A and B. That is, when a new EMG ratio is created by moving data windows, the change in hedge ratios must exceed θ; otherwise, the hedge ratio should not change at all. As a result, when comparing to the other two methods, more changes fall below 0.005 but fewer changes lie between 0.005 and 0.05. Finally, it is worthwhile pointing out that the big changes (larger than 0.20) are caused by the multiple-root problem. For the two approaches of Ederington and EMG (with $\nu = 2$), the objective function has a unique local minimum. Yet, when adopting $\nu = 100$, the EMG method entails two local minima: one located near 0.4, and the other near 0.8, on average. As the data windows move, it happens on several occasions that one replaces the other as the global minimum, and a very big change results.

When ν is moderate, the approximation of $AS\Gamma(\nu)$ does not perform well. Despite that, R_{pm} remains dominant in the calculation of $S\Gamma(\nu)$, and other data points are also important, as the difference between $(1 - \hat{F}(R_{pt}))^\nu$ and $(1 - (\frac{1}{T}))^\nu$ may not be negligible. In other words, the complete ranking of data points should be considered. A new EMG ratio is more plausible if it generates a different ranking. Moreover, a large difference between the old and new hedge ratios should correspond to more dissimilar rankings. To account for this conjecture, let r_{oi} and r_{ni} denote the ranking of the ith data point from the original sample on the basis of the old and new hedge ratios, respectively. The sum of the squares of changes in rankings across observations is calculated as:

$$\mathrm{SSRC} = \sum_{i=1}^{T} (r_{oi} - r_{ni})^2 \tag{7}$$

This number is employed to measure the dissimilarity of two rankings (with $T = 50$ in our case). Table III shows there is a positive correlation between SSRC and changes in large ratios. When ν is large, the ranking of the sample data is insensitive to small changes in hedge ratios but very sensitive to some moderate changes in the ratios. Again, the instability of the EMG method is well established.

NONPARAMETRIC ESTIMATION OF EMG COEFFICIENTS

To derive the EMG hedge ratio, an estimator for the cumulative density function is necessary. Previous researchers all rely upon the empirical estimator following Lerman

Table III
THE RELATIONSHIP BETWEEN CHANGES IN HEDGE RATIOS AND SSRC

	Sum of Squared Rank Change		
Change in Hedge Ratio[a]	**Minimum**	**Maximum**	**Average**
(0, 0.00016)	2	4	2.8
(0.00016, 0.01)	22	64	31.3
(0.01, 0.027)	188	272	216.5
(0.027, 0.07)	328	710	510.5
(0.07, 0.13)	604	5688	2053
(0.13, 0.20)	368	9584	6965
(0.20, 0.50)	6468	6468	6468

[a]The two number (a, b) indicate the lower and upper bounds on the changes.

and Yitzhaki (1984). The method is less satisfactory if the smoothness of the estimator is of importance. Moreover, Reiss (1981) showed that the empirical estimator is asymptotically deficient with respect to a certain appropriately chosen kernel estimator. That is, a much larger sample is required for the empirical estimator to achieve the same efficiency as the kernel estimator. More importantly, the kernel-based estimates are much smoother than empirically derived estimates. It is interesting to investigate whether the smoothness of the kernel estimator promotes the stability of the EMG hedge ratio.[6]

Let $k(x)$ denote a kernel function such that $k(x) \geq 0$ for every x and

$$\int_{-\infty}^{\infty} k(x)\,dx = 1 \tag{8}$$

Let $K(x)$ denote the integration of $k(\cdot)$ up to x. The kernel estimator for the cumulative density function of a random variable W is:

$$F(z) = \frac{1}{T}\sum_{t=1}^{T} K((z - W_t)/\alpha_T) \tag{9}$$

when evaluated at the point z; where $\{W_1, \ldots, W_T\}$ is a sample of W and $\alpha_T > 0$ is called the bandwidth. Note that the empirical estimator corresponds to the limiting case of $\alpha_T = 0$. The choice of α_T affects the smoothness of the resulting estimator. Although several criteria are proposed for selecting the optimal bandwidth, unanimous agreement is lacking [Silverman (1986)]. In this study, different bandwidth choices are considered when deriving the EMG hedge ratios. The choice of the kernel function is generally of less importance. The following empirical studies adopt a Gaussian kernel, i.e., $k(x)$ is chosen to be the probability density function of a standard normal random variable.

Table IV displays the EMG hedge ratios for various risk aversion parameter values upon applying kernel estimation to the EMG coefficient. From the last column, it is found that the results are nearly identical to those derived from the empirical estimator (see Table I). In general, the EMG ratio appears to be robust to various bandwidth choices. Note that the kernel method provides a smooth estimator of the EMG coefficient. As a result, for each bandwidth choice, the optimal hedge ratio is a smooth function of the risk aversion parameter. Comparison between Table I and Table IV reinforces the previous conclusion that the EMG hedge ratio (calculated with empirical estimator) responds

[6]Besides the kernel estimator, Sarda (1991) applies a polynomial to smooth the empirical distribution function. It is shown that the approach is equivalent to a kernel-type estimator with the kernel linked to the weights.

Table IV
OPTIMAL HEDGE RATIOS USING KERNEL ESTIMATORS

ν	$\alpha_T = 10^{-4}$	$\alpha_T = 10^{-5}$	$\alpha_T = 10^{-6}$
2	−0.88813	−0.88819	−0.88822
3	−0.87888	−0.87904	−0.87894
4	−0.86434	−0.86449	−0.86452
5	−0.85222	−0.85325	−0.85308
6	−0.84266	−0.84268	−0.84251
7	−0.83481	−0.83432	−0.83440
8	−0.82613	−0.82488	−0.82480
9	−0.81800	−0.81778	−0.81800
10	−0.80898	−0.80983	−0.80981
20	−0.77100	−0.77251	−0.77320
30	−0.75977	−0.75636	−0.75969
40	−0.75251	−0.75634	−0.75230
50	−0.74006	−0.74259	−0.74285
100	−0.72081	−0.72081	−0.72081
150	−0.71163	−0.71850	−0.71850

smoothly to the risk aversion parameter. The above method is also applied to examine the effects of data windows.[7] Recall that the empirical estimator generates a robust result for the hedge ratio in response to changes in data windows when $\nu = 2$. Table V shows that a large bandwidth improves the robustness slightly. For example, when $\alpha_T = 10^{-3}$, more consecutive changes fall short of 0.005. As α_T decreases, the number also decreases. The results derived from the empirical estimator are identical to the case of $\alpha_T = 10^{-6}$.

Table V
ABSOLUTE CONSECUTIVE CHANGES USING KERNEL ESTIMATORS ($\nu = 2$)

Range	$\alpha_T = 10^{-3}$	$\alpha_T = 10^{-4}$	$\alpha_T = 10^{-5}$	$\alpha_T = 10^{-6}$
(0, 0.005)	138	136	132	131
(0.05, 0.01)	46	41	46	48
(0.01, 0.02)	17	22	21	20
(0.02, 0.03)	3	5	5	5
(0.03, 0.04)	2	2	2	2
(0.04, 0.05)	1	1	1	1
(0.05, 0.06)	0	0	0	0
(0.06, 0.07)	0	0	0	0
(0.07, 0.08)	0	0	0	0
(0.08, 0.09)	1	1	1	1
(0.09, 0.105)	1	1	1	1

[7]A referee argued that the kernel estimation does not provide superior results and therefore the details should be omitted. Yet, theoretically, the kernel method is much better than an empirically driven estimation. The fact that the two methods generate similar conclusions in this sample implies the latter is acceptable in spite of its theoretical weakness. It is not known whether the two methods provide similar results for other sample data. In case of any divergence, the kernel estimation must be employed. The procedures discussed in this section will be useful.

When $v = 100$, choices of the bandwidth produce more effects. When $\alpha_T = 10^{-3}$, 191 out of 209 changes fall short of 0.005, ten changes lie between 0.10 and 0.30, and three changes exceed 0.495 (see Table VI). As α_T decreases, the three proportions begin to shift. For example, when $\alpha_T = 10^{-5}$, the three numbers become 168, 20, and 0, respectively. Again, the results derived from the empirical estimator are nearly identical to the case $\alpha_T = 10^{-6}$. It is evident that the instability of the EMG hedge ratio in this case is not due to the application of the empirical estimator. In fact, the kernel estimator with a small bandwidth (such as $\alpha_T = 10^{-3}$) may promote instability, at least in these sample data.

CONCLUSIONS

This article analyzes the stability of the hedge ratio following the EMG method. It is shown that, contrary to the previous conclusion by Kolb and Okunev (1992), the EMG hedge ratio is a smooth function of the underlying risk aversion parameter. Moreover, the ratio tends to be a monotonic function of the parameter based on empirical evidence in the KO article and in an alternative S&P 500 weekly data set.

According to Kolb and Okunev (1992), moving data windows also produces a step function for hedge ratios. This article demonstrates that the conclusion applies to a large risk aversion parameter ($v = 100$) but not for the case of $v = 2$. A theoretical analysis is provided to explain the phenomena based on the minimal rank point and changes in the ranking of the whole sample data. Moreover, it is found that multiple local minima prevail when $v = 100$. As the data window moves, occasionally one local minimum overtakes the other to become the global minimum, resulting in a dramatic shift in EMG hedge ratios.

Finally, a nonparametric kernel estimation is applied to reexamine the above two issues. The method, replacing the empirical distribution function with kernel estimation,

Table VI
ABSOLUTE CONSECUTIVE CHANGES USING KERNEL ESTIMATORS ($v = 100$)

Range	$\alpha_T = 10^{-3}$	$\alpha_T = 10^{-4}$	$\alpha_T = 10^{-5}$	$\alpha_T = 10^{-6}$
(0, 0.005)	191	182	168	161
(0.005, 0.01)	0	2	4	2
(0.01, 0.02)	1	0	0	3
(0.02, 0.03)	2	4	5	5
(0.03, 0.04)	0	0	1	4
(0.04, 0.05)	0	0	1	2
(0.05, 0.06)	1	0	0	0
(0.06, 0.07)	0	3	4	4
(0.07, 0.08)	0	1	2	2
(0.08, 0.09)	1	0	1	1
(0.09, 0.10)	0	2	1	0
(0.10, 0.15)	3	4	8	12
(0.15, 0.20)	5	5	6	8
(0.20, 0.30)	2	1	6	3
(0.30, 0.40)	0	4	1	1
(0.40, 0.495)	0	1	1	1
(0.495, 0.60)	3	0	0	0

produces a smooth estimator for the EMG coefficient. It turns out that all previous conclusions are retained even when various choices of bandwidth are considered.

Appendix

Upon substituting $(v + 1)$ for (v) into eq. (4),

$$Q(v + 1), x^*(v)) = \sum_{t=1}^{T} R_{ft}(1 - \hat{F}(R_{pt}))^v - \left(\sum_{t=1}^{T} R_{ft}/T\right)\left(\sum_{t=1}^{T}(1 - \hat{F}(R_{pt}))^v\right)$$
$$+ v\left\{\sum_{t=1}^{T} R_{pt}R_{ft}\hat{f}(R_{pt})(1 - \hat{F}(R_{pt}))^{v-1}\right.$$
$$\left. - \left(\sum_{t=1}^{T} R_{pt}/T\right)\left(\sum_{t=1}^{T} R_{ft}\hat{f}(R_{pt})(1 - \hat{F}(R_{pt}))^{v-1}\right\}. \quad \text{(A1)}$$

Consider the first term, it can be rewritten as:

$$\sum_{t=1}^{T} R_{ft}(1 - \hat{F}(R_{pt}))^v = \sum_{t=1}^{T} R_{ft}(1 - \hat{F}(R_{pt}))^{v-1} \cdot SE(1 - \hat{F}(R_{pt}))$$
$$+ \sum_{t=1}^{T} R_{ft}(1 - \hat{F}(R_{pt}))^{v-1} \cdot \Delta(1 - \hat{F}(R_{pt})) \quad \text{(A2)}$$

where $SE(1 - \hat{F}(R_{pt}) = \sum_{t=1}^{T}(1 - \hat{F}(R_{pt}))/T$ is the sample expectation, and $\Delta(1 - \hat{F}(R_{pt}))$ is defined as $(1 - \hat{F}(R_{pt})) - SE(1 - \hat{F}(R_{pt}))$. Applying similar decomposition to all other terms, $Q(v + 1,\ x^*(v))$ becomes

$$SE(1 - \hat{F}(R_{pt})) \cdot Q(v, X^*(v)) + S\ \mathrm{cov}(R_{ft}, (1 - \hat{F}(R_{pt}))^{v-1} \cdot \Delta(1 - \hat{F}(R_{pt})))$$
$$+ v \cdot S\ \mathrm{cov}(R_{pt}, R_{ft}\hat{f}(R_{pt})(1 - \hat{F}(R_{pt}))^{v-2} \cdot \Delta(1 - \hat{F}(R_{pt})))$$
$$+ SE(1 - \hat{F}(R_{pt})) \cdot S\ \mathrm{cov}(R_{pt}, R_{ft}\hat{f}(R_{pt})(1 - \hat{F}(R_{pt}))^{v-2}) \quad \text{(A3)}$$

Note that $Q(v, x^*(v)) = 0$ and $\hat{f}(R_{pt}) = 1/T$. Provided that R_{pt} is generally small and $v < T$, $Q(v + 1, x^*(v))$ is approximated by the second term. Moreover,

$$S\ \mathrm{cov}(R_{ft}, (1 - \hat{F}(R_{pt}))^{v-1} \cdot \Delta(1 - \hat{F}(R_{pt})))$$
$$= S\ \mathrm{cov}(R_{ft}, \Delta(1 - \hat{F}(R_{pt})) \cdot SE[(1 - \hat{F}(R_{pt}))^{v-1}]$$
$$+ S\ \mathrm{cov}(R_{ft}, \Delta(1 - \hat{F}(R_{pt}))^{v-1} \cdot \Delta(1 - \hat{F}(R_{pt}))) \quad \text{(A4)}$$

where $SE(1 - \hat{F}(R_{pt}))^{v-1} = \sum_{t=1}^{T}(1 - \hat{F}(R_{pt}))^{v-1}/T$ and $\Delta(1 - \hat{F}(R_{pt}))^{v-1} = (1 - \hat{F}(R_{pt}))^{v-1} - SE(1 - \hat{F}(R_{pt}))^{v-1}$. The second term in the above equation is generally very small. Eq. (5) is then obtained.

Bibliography

Cheung, C. S., Kwan, C. C., and Yip, P. C. (1990): "The Hedging Effectiveness of Option and Futures: a Mean-Gini Approach," *The Journal of Futures Markets,* 10:61–73.

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

Kolb, R. W., and Okunev, J. (1992): "An Empirical Evaluation of the Extended Mean-Gini Coefficient for Futures Hedging," *The Journal of Futures Markets,* 12:177–186.

Lerman, R. I., and Yitzhaki, S. (1984): "A Note on the Calculation and Interpretation of the Gini Index," *Economics Letters,* 15:363–368.

Reiss, R. D. (1981): "Nonparametric Estimation of Smooth Distribution Functions," *Canadian Journal of Statistics,* 8:116–119.

Sarda, P. (1991): "Estimating Smooth Distribution Functions," in *Nonparametric Functional Estimation and Related Topics,* Roussas, G. (ed.). New York: Kluwer, pp. 261–270.

Silverman, B. W. (1986): *Density Estimation for Statistics and Data Analysis.* London: Chapman and Hall.

Yitzhaki, S. (1982): "Stochastic Dominance, Mean Variance, and Gini's Mean Difference," *American Economic Review,* 72:178–185.

Yitzhaki, S. (1983): "On an Extension of the Gini Inequality Index," *International Economic Review,* 24:617–628.