Alternative Variance Estimators for Pricing Options

By

Robert Geske, Richard Roll*, Yuzhao Zhang
Anderson School at UCLA

January 2007
This Draft May 2008

JEL Classification: G12, G13, C11
Keywords: Derivatives, volatility, shrinkage estimator

As this research is preliminary please do not quote or distribute without the author's permission. *Contact Roll by mail at The Anderson School at UCLA, 110 Westwood Plaza, Los Angeles, California 90095, USA, by telephone at 310-825-6118, or by e-mail. roll@anderson.ucla.edu.

Abstract

This paper examines a volatility estimation bias that may be commonly exhibited by all option pricing models on all underlying sources of risk. Black-Scholes (1972) were the first to illustrate the bias by showing that their model under priced options on relatively low variance stocks and over priced options on relatively high variance stocks. The bias is always observed in cross section among individual stocks. We think this bias might have nothing to do with Black-Scholes or any option pricing model but instead might be attributable to sampling error. Thus, this bias should be observed with any option pricing model on any underlying, not just equity, but also fixed income securities, foreign exchange, and commodities. To test this idea, we use shrinkage estimators of James-Stein detailed in Efron-Morris (1976) and Ledoit-Wolf (2004a). While both shrinkage estimators utilize the covariance matrix, Ledoit-Wolf (or LW hereafter) is unique because it does not require matrix inversion. We show that the variance bias can be eliminated using these improved estimators.

1. Introduction

The Black-Scholes (1973) option pricing model exhibits systematic mis-pricing of options on individual stocks and options on indexes of stocks. This mis-pricing has been related to moneyness (S/K), time to expiration, and volatility. The mis-pricing has also been related to the Black-Scholes distributional assumption, to their assumption of no dividend payouts, and to the model's European rather than American nature.[1]

This paper's concern is the volatility bias observed in cross-section when pricing options on individual stocks. Black-Scholes (1972) were the first to report that their model under-priced options on low variance stocks and over-priced options on high variance stocks. Black-Scholes used over the counter option (OTC) data when they reported this variance bias because listed options did not commence trading until April, 1973. OTC options are quasi-European because OTC dividend protection eliminates the probability of early exercise.[2] Black (1975) later reported that the model also under-priced out-of-the-money options and near maturity options, while it over-priced in-the-money options on individual stocks. MacBeth-Merville (1979), Rubinstein (1985), Whaley (1982), Sterk (1982), Geske-Roll (1984a), and others discuss these biases but do not focus on the volatility bias.

There have been many theoretical papers concerned with Black-Scholes assumption of constant or deterministic stock return volatility. (Cf. Merton (1976), Cox-Ross (1976),

---

[1] See Black-Scholes (1972), Black (1975, Macbeth-Merville (1979), Rubinstein (1985), Whaley (1982), Sterk (1982), Geske-Roll (1984a).
[2] See Geske, Roll, Shastri (1983).

Geske (1979), Hull-White (1987), Heston (1993), Bakshi, Cao, Chen, (1997), Heston-Nandi (2000).) However, these papers would potentially alter the prices of options on all individual stocks without a particular focus on the observed cross-sectional mis-pricing of options on low and high variance stocks. Thus, in this paper it is our thought to see whether this variance bias observed in individual option cross-sectional prices can be attributed to estimation error in the sample variance.

There is some a priori reason to suspect estimation error in the sample variance rather than the model as the source of this particular mis-pricing. The reason is that this variance related mis-pricing always arises in the context of an *inter-stock* comparison. This is in contrast to other biases (moneyness, time to expiration), which can be detected in an *inter-option* comparison. Unlike the strike price and time until expiration parameters, the true variance is identical for all identical expiration options on the same stock on a given date. Thus, investigation of the variance related mis-pricing cannot rely on either the implied variance or other more sophisticated option pricing models, but must instead be based on historical estimates of actual stock return volatility.

There are many techniques to improve the accuracy of the volatility estimate for individual stocks. (Cf. Boyle-Ananthanarayan (1977), Parkinson (1980), Garman-Klass (1980), and Butler-Schacter (1986), ARCH, GARCH.) However, the essence of the present problem is that a number of variances are estimated simultaneously, one for each stock, and then option mis-pricing is related cross-sectionally to these multitudinous estimates.

The problem of simultaneously estimating multiple parameters has become well-known in statistical theory. The cross-sectional sampling distribution consists of two parts, variability in the true underlying population parameters and variability in the estimation error. In any sample, larger estimates relative to the cross-sectional mean are more likely to contain positive sampling errors and vice versa for relatively smaller estimates. Thus, in a cross-sectional comparison of option mis-pricing, estimation error alone will cause stocks with larger estimated variances to over-price the market and stocks with smaller estimated variances to under-price the market. The Black-Scholes model price, being a positive function of the sample variance, should display a positive cross-sectional mis-pricing. This is exactly the observed mis-pricing phenomenon.

When many variances are being estimated, one for each stock, a James-Stein (1961) estimator is unambiguously superior to the standard univariate estimator. The James-Stein estimator reduces estimation risk on average over all stocks. Such an estimator "shrinks" each individual variance estimate toward a target such as the grand mean of all estimates. Since the variance bias is characterized by over-pricing options on high volatility stocks and under-pricing options on low volatility stocks, adjusting each estimated volatility toward the average volatility for all stocks obviously has the potential to reduce the observed variance bias. In the multiple variance estimation setting, the superior James-Stein estimation technique has the potential to eliminate this problem.

Geske-Roll (1984b) observed the variance bias and were the first to attempt a correction based on the idea that the problem was related to sampling error in volatility estimation rather than model error. They originally chose to use a shrinkage version of Stein's technique described in Efron-Morris (1976) as related to empirical Bayesian estimation. [3] However, this particular "shrinkage" technique involves two difficult questions.. First, how much historical data should be used to estimate individual stock variances? Second, toward what target should individual stock variance estimates be shrunk? Until recently, the first question was usually resolved by constraints on matrix inversion. The sample covariance matrix is non-singular only when the time series sample size, N, exceeds the number of stocks, k.[4] Because of this requirement, smaller groups of stocks are often formed to estimate parameters, and then results from the smaller groups are combined and analyzed.

The second question of shrinkage target is more complex. The target should have minimal free parameters (a lot of structure), should have less estimation error, and should somewhat reflect the characteristics of the quantity to be estimated. In three recent papers Ledoit-Wolf (2003, 2004a, 2004b) have introduced new techniques that provide solutions to these requirements.

Ledoit-Wolf start with the sample covariance matrix because it is unbiased and easy to calculate. They recognize that it is subject to estimation error, especially when there are fewer time series observations than individual stocks, which is often the case in financial

---

[3] Subsequent to Geske and Roll (1984b), several other papers confront the same volatility problem Karolyi (1993) uses a Bayesian approach. He describes the difference (p. 583) as follows: "What distinguishes the Bayesian estimator of volatility from the "shrinkage" estimator … is in the adjustment process." Karolyi considers only call options and he reports that the Bayesian approach eliminates the volatility bias for high volatility stocks but there remains a statistically significant but small bias for the low volatility stocks. Karolyi also reports that the Bayesian estimator creates an under pricing bias in all the call options. Geske and Torous (1990; 1991) use robust techniques to treat outliers when estimating volatility (1990), and they also examine the effects of a non-normal skewness and kurtosis on option prices.

[4] In addition, no two stocks can be perfectly correlated in sample, Although perfect correlation is rarely an issue, very high correlation can cause instabilities in the resulting shrinkage estimates.

applications. They also recognize that an estimator with more structure would have less estimation error, but would likely be mis-specified and biased. Thus, they find a compromise by computing an optimal linear convex combination of the sample covariance matrix and a structured target. They provide results for three targets, the Sharpe single index model, the identity matrix, and a constant correlation model. Herein we now choose to compare a version of the James-Stein estimator to the Ledoit-Wolf technique. For Ledoit-Wolf we shrink toward the simplest target, the identity matrix, which is well conditioned, structured, and parsimonious.

Section 2 describes the data and test calculations. Section 3 describes alternative variance estimators. Section 4 reports the results and shows that the shrinkage techniques of Stein and Ledoit-Wolf both eliminate the variance bias for puts and eliminate or substantially reduce the bias for calls, but that the Ledoit-Wolf technique is superior with respect to prediction error. Section 5 concludes.

2. Data and Test Calculations

The data come from CRSP for daily stock returns and from Option Metrics (OM) for call and put option prices, dividend distributions, and implied volatilities. The OM data span the 100 months from January, 1996 through April, 2004 inclusive.

Stocks are screened one way and options are screened five ways. To assure that stocks are actively traded, we use only the 500 largest stocks by market capitalization on the last trading day of the previous year. Stocks are limited to common shares with share codes 10 or 11. For options, the first screen limits observations to the first trading day of each calendar month. This potentially provides 100 monthly observations of options on 500 individual stocks and allows estimators of volatility to be computed with return observations through the end of each preceding month. The second screen limits options to being near-the-money, which we define as $0.95 < K/S < 1.05$ (with K the strike price and S the stock price on the first day of the month.) Near-the-money options are the most actively traded of all options with different times to expiration, and since these are options on large companies they usually trade many times every day. Also, near-the-money options should exhibit less moneyness bias. The third option screen restricts the sample to options expiring on the third Friday of the next month. Thus, all options have the same short time to expiration, which should control somewhat for any time bias. Short-maturity options are also the most actively traded of all options with different strike prices. Thus, near-the-money, short-maturity options on large stocks should trade many times every day. The fourth screen restricts options to those that actually did trade on each day. The fifth option screen eliminates any detectable arbitrage violations (e.g., $C > S - K e^{-rT}$; $P > S - K$). After these screens, the sample has on average about 494 call options and 488 put options per month.

Historical volatilities are computed for each individual stock using 126 days (approximately 6 months) of previous CRSP daily data preceding each of the 100 first day of month observations for the stock price and option prices. Stock betas are calculated using 504 days (approximately 2 years) of daily data preceding each of the 100 first day of month observations, using the CRSP value weighted return as the market index. (Betas are inputs for the particular Stein estimator that assumes a one-factor structure for the covariance matrix.) We also compute the sample covariance matrix for all stocks in the sample using the preceding 6 months of CRSP daily data; this is an input for the Ledoit-Wolf estimator.

3. Alternative Variance Estimators

The variance estimate should be forward-looking. An obvious choice for the estimate of an expectation is the average from historical data. Stein (1955) showed that when the number of expectations being estimated exceeds two, estimating each of them by its own historical average is an inadmissible procedure. In other words, no matter what the true values, there are estimation methods with smaller total risk, where risk is defined as the expected value of the squared error of the estimator. Stein and James provided an example of such an estimator. The James-Stein estimator (1961) is given by an equation similar to the following:

$$\hat{\sigma}_{Sj}^2 = \hat{\bar{\sigma}}^2 + \gamma_j (\hat{\sigma}_{Hj}^2 - \hat{\bar{\sigma}}^2) \qquad (1)$$

where $\hat{\sigma}_{Sj}^2$ is the Stein estimator for stock j, $\hat{\sigma}_{Hj}^2$ is an historical estimate for the same stock, $\hat{\bar{\sigma}}^2$ is the grand cross-sectional average of all the historical estimates, and γ_j is a shrinking intensity factor bounded between zero and one.

As a simple example, assume the grand average of all variances in the stock return sample is $(0.29)^2$ annually, that the shrinkage factor is 0.4, and a particular company's standard deviation estimated from historical data is 0.54 annually. For this relatively high volatility stock the James-Stein estimate of the stock's true volatility is about 0.41 instead of 0.54. Consider another company with volatility estimated from historical volatility by traditional methods to be 0.19 annually. For this relatively low volatility stock the James-Stein estimate of the stock's true volatility would be 0.25. Recall that the Black-Scholes model under prices options on relatively low volatility stocks and over prices options on relatively high volatility stocks. The Stein estimator which has an essential process of shrinking all individual estimates toward a less disperse target clearly has the potential to remove this bias.

As already mentioned Stein estimators are reminiscent of Bayesian methods. In the limit, as the number of estimates becomes very large, Stein and Bayes' converge. In practice, the James-Stein estimator is often referred to as an "empirical Bayes" rule.[5] In the above example, the shrinkage intensity factor, γ, is treated as a constant. It is potentially a function of many things, including the sample averages, the number of stocks in the sample, the number of observations for each stock, the estimated historical volatility of each stock, and the grand mean of all stock volatilities. For example, the covariance estimator provided by Efron-Morris (1976) is given by

$$\hat{S}_S = \left[(N-k-2)/(N-1)\hat{S}_H^{-1} + (k+1-2/k)/(N-1)(\hat{\bar{\sigma}}^2 I)^{-1}\right]^{-1} \qquad (2)$$

where S denotes covariance, with subscripts $_H$ and $_S$ denoting historical and Stein, respectively, N is the time series sample size, k is the number of securities, I is the (kXk)

[5] See Efron-Morris (1975). They discuss Stein's rule as an empirical Bayes rule, and present applications such as predicting baseball batting averages, estimating toxomosis prevalence rates, and estimating the exact size of Pearson's chi-square test. In equation (2) N>k or grouping is required for Stein, but this is not true for Ledoit-Wolf. Later we group Ledoit-Wolf only to illustrate a similar comparison to Stein.

identity matrix and $\hat{\bar{\sigma}}^2$ is the grand sample mean of historical variances. In this case, shrinkage produces an estimate of the inverse covariance matrix with shrinkage intensity approximately (N-k-2) / (N-1).

A major limitation of generalized Stein techniques for financial applications is that the sample covariance matrix has too little structure. If, for example, it is beneficial to use the sample covariance matrix of stock returns, but the number of historical returns per stock, N, is of the same order of magnitude as the number of stocks, k, then the total number of parameters to be estimated is of the same order as the total size of the data available. When k is larger than N, the sample covariance matrix is always singular, even if the true covariance matrix is known to be non-singular. Muirhead (1987) reviews the literature on shrinkage estimators of the covariance matrix and shows that they all suffer from two major limitations: (i) they break down when k>N and the matrix cannot be inverted; (ii) they do not utilize a priori knowledge about correlations between stock returns. We can circumvent the second limitation by assuming that asset returns follow a factor model, say the single-factor market model akin to the CAPM. Therefore the off-diagonal entry i, j of $\hat{S}_S$ is simply $\hat{\beta}_i \hat{\beta}_j \hat{\sigma}_m^2$. By imposing more structure in this fashion, one can make the sample covariance matrix behave. Ledoit-Wolf techniques circumvent both of these problems.

For a given (NXk) matrix X of de-meaned observations, Ledoit-Wolf derive an "optimal" estimator, S^*, that is a linear combination of the sample covariance matrix, $S = X X' / N$, and a target matrix, whose expected quadratic loss $E [\| S^* - S \|^2]$ is a minimum. When the target is the identity matrix, I, they show that $S^* = \gamma \mu I + (1 - \gamma) S$ depends only on four unobservable scaler functions of the true covariance matrix (μ, ζ, β, α) which can be consistently estimated ($\rightarrow_{qm}$ as $n\rightarrow_{infinity}$) from their sample counterparts.[6]

Define $m \equiv < S, I >$. Then $E(m_n) = \mu_n$ for all n, $m_n - \mu_n \rightarrow_{qm} 0$ as $n\rightarrow_{infinity}$ (3)

Define $d_n^2 \equiv \| S_n - m_n I_n \|_n^2$. Then $d_n^2 - \zeta_n^2 \rightarrow_{qm} 0$, and $\zeta_n^2 = E[\| S_n - \mu_n I_n \|_n^2]$ (4)

Define $b_n^2 = \min(Ъ_n^2, d_n^2)$, $Ъ_n^2 \equiv 1/n^2\, \Sigma^n (\| XX^t - S_n \|_n^2$; b_n^2 & $Ъ_n^2 \rightarrow_{qm} \beta$ (5)

Define $a_n^2 \equiv d_n^2 - b_n^2$. Then $a_n^2 - \alpha_n^2 \rightarrow_{qm} 0$, and $\alpha_n^2 = \zeta^2 - \beta^2$ (6)

and using these scalers (m_n, d_n, b_n, a_n) a linear combination of S and I that minimizes the expected quadratic loss is:

$$S_n^* = b_n^2/d_n^2\, m_n I_n + a_n^2/d_n^2\, S_n \qquad (7)$$

Now, if γ is defined as $\gamma \equiv b_n^2 / d_n^2$, then $S^* = \gamma\, m\, I + (1-\gamma)\, S$.

4. <u>Experimental Results</u>

We shrink the standard historical volatilities estimates in four ways (two Stein and for exact comparisons two for Ledoit-Wolf) and then compare the five estimators (including the historical.) For the Stein estimators, we assume a one-factor structure for the covariance matrix. We form groups of 50 stocks each for Stein because it requires the cross-section of individual stocks, k, to be smaller than the time series of observations, N (herein N=100). The two Stein estimators differ because the first groups stocks randomly while the second estimator groups stocks to maximize the volatility dispersion within

[6] See Ledoit-Wolf (2004a), p.379-380. The squared Frobenius norm $\| \; \|^2$ is a quadratic form whose inner product is $< X X' > = tr(X X') / N$ and the four unobservable scalers are $\mu = <\Sigma, I>$, the expectation of the grand mean of the eigen-values, $\beta^2 = E[\| S-\Sigma \|^2]$, the error of the sample covariance matrix, and $\zeta^2 = E[\| S-\mu I \|^2]$, the cross-sectional dispersion of the sample eigenvalues, $\alpha^2 = \| \Sigma-\mu I \|^2$, and Σ is the true covariance matrix, and $\rightarrow_{qm}$ denotes convergence in quadratic mean as $n\rightarrow$infinity.

each group. To achieve the volatility dispersion, we first sort all the 500 stocks by their historical volatilities and allocate the stocks ranked 1, 11 … 481, 491 to the first group, 2, 12, … to the second group, and similarly for all 10 groups.

At the beginning of months the various volatility estimates are matched near-the-money implied volatilities for options that expire the next month. For example, on January 4, 1996, we choose the 500 largest stocks by market capitalization at the end of 1995 and compute their 6 month historical volatilities using the previous 126 days of daily data. Four shrunk volatilities are then computed for each stock for both calls and puts, and the implied volatility of the near-the money option expiring the next month (February 20, 1996) is computed.

Table 1 provides summary statistics for all the volatility estimators. The statistics presented are time-series means of the cross-sectional summary statistics. On average, all Stein-type estimators have a lower mean than the original historical estimates as well as the Ledoit-Wolf estimators. The reason is all the Stein-type estimators involve matrix inversions which decrease the average due to the Jensen's inequality. Moreover, all shrinkage estimators exhibit lower cross-section dispersion as expected, consistent with the shrinkage process. The LW estimators still preserves much more cross-sectional variation compared with the Stein estimators, which suggests that the LW shrinkage intensity is effectively smaller.

Bias can be measured by examining the log ratio of implied to estimated volatility. Thus, for each stock we first compute five bias ratios, one for the historical volatility and four for the shrinkage estimators and then regress each bias ratio on the volatility estimator that generated the bias, and on the moneyness of the option as a control. Specifically, we define:

$$Error_{estimator} = log(\sigma_{implied} / \sigma_{estimator}) \qquad (8)$$

with estimator = Historical, James-Stein with random groups, James-Stein with large dispersion groups, Ledoit-Wolf with no groups and with groups for similar comparison.

Then the following cross-sectional regressions are computed for each month i:

$$Error_{estimator,,i} = \alpha + \omega\, \sigma_{estimatort,i} + \eta\, S_i/K_i + \varepsilon_{,i} \qquad (9)$$

Following Fama-MacBeth [1973], time series means of the cross-sectional coefficients are compared against time series standard errors computed using a Newey-West autocorrelation correction with 8 lags.

Table 2 presents the main results from these regressions. The historical volatility column reports the coefficients and test statistics for equation 9 when volatility is computed with the standard historical method; it shows clearly the extent of the previously-observed volatility bias. For calls (and puts), the *vol* coefficient is large, negative, -0.437 (-0.474) and very significant, t=-7.596 (t=-7.752.) This is consistent with the finding in Black-Scholes (1972) that in the cross-section options of low (high) volatility stocks are under priced (over priced) by their model.

In the next four columns of Table 2 for the panel using calls, the *vol* coefficients for the four shrinkage estimators are (-0.119, 0.010, 0.001, -0.040) and t statistics are (-0.943, 0.072, 0.019, -1.417) for James-Stein Random, James-Stein High Dispersion, Ledoit-Wolf No Group, and Ledoit-Wolf Group. This shows that the volatility bias has been eliminated by Stein and by Ledoit-Wolf. The control for moneyness reveals that the moneyness bias is significant and is independent of the volatility bias. For puts Stein and Ledoit-Wolf also eliminate the volatility bias.[7] Thus, we conclude that both Stein and Ledoit-Wolf shrinkage techniques are able to eliminate this volatility bias of under pricing options on low volatility stocks and over pricing options on high volatility stocks.[8]

In Table 3 we present further analysis and comparisons of the historical and shrinkage estimators. This table shows the average prediction errors of the uncorrected historical estimator and of the corrected shrinkage estimators. We wanted to see if the process of shrinking the volatility estimators increased the prediction errors even though it eliminated the volatility bias. Row 1 for call options shows that the uncorrected historical volatility estimator has the smallest prediction error of 0.042. The prediction errors for both Stein 1 (random) and Stein 2 (disperse) are larger (0.057 and 0.054) and very significantly different from the historical estimator (t –stats of 3.473 and 3.163).

---

[7]The sample size of nearly 500 calls and puts was obtained by using the mid-point of the bid-ask spread for days when options on specific stocks apparently did not trade. When we eliminated all stocks whose options did not trade every day and the resultant sample of calls (puts) was reduced to 302 (205), both Ledoit-Wolf and Stein eliminated the volatility bias for both calls and puts.

[8] In unreported results, we also examine other variants of the James-Stein estimators with differing assumptions about the covariance matrix target. One target assumes that all covariances are the same and equal to the average sample covariance. The other target assumes that all covariances are zero. These calculations were again carried out with randomly sorted groups and with groups organized to maximize intra-group volatility dispersion. In all cases, the results are essentially the same as those reported in all Tables (1 through 6 inclusive), for the James-Stein estimators. The authors will be happy to provide detailed results to interested readers.

However, the prediction error for both the Ledoit-Wolf estimators, No Group and Group, 0.043, is almost the same size as the uncorrected historical prediction error, 0.042, and Ledoit-Wolf is not significantly different from historical. For put options the results are very similar, with the only difference being that the Ledoit-Wolf estimators (No Group and Group) now have the lowest prediction errors, both 0.037, but it is not statistically different from the uncorrected historical estimator prediction error, 0.039. Thus, we see that while the Stein shrinking does eliminate the volatility bias, it also increases the prediction error, and this increased error is statistically significant. The Ledoit-Wolf estimator does not have this problem.

The prediction errors can be elucidated by using Theil's decomposition, which separates the error into three components: (i) the error attributable to bias in the forecasts (UM); (ii) the error attributable to low correlation between the actual and the forecast (UR); and (iii) the remaining prediction error (UD). This analysis for call options shows that the portion of the prediction error attributable to bias in the forecast (UM) is not significantly different from the historical estimator for any of the shrinkage estimators. The portion of the prediction error attributable to low correlation between the actual and the forecast (UR) is lowest for the Ledoit-Wolf estimators compared to Stein, 0.001 for both calls and puts, and LW is very significantly different from the uncorrected historical estimator. The remaining portion of the larger prediction errors for both calls and puts for the Stein estimators are significantly different from the uncorrected historical estimate while Ledoit-Wolf is not different.

It could be illuminating to examine whether stocks with different percentages of systematic and idiosyncratic components of risk are shrunk differently. Thus, we define the following relative indicator of systematic risk for each stock I at the beginning of month t, based on approximately two prior years of daily returns:

$$\text{Sys}_{i,t} = 1 - \left(\frac{\sigma_{\text{idiasyncratic},i,t}}{\sigma_{\text{historical},i,t}} \right)^2 ; \tag{10}$$

and the following indicator of relative shrinkage:

$$\text{Shrinkage}_{i,t} = \quad Abs \left[Ln \quad \frac{\sigma^2_{shrunk,i,t}}{\sigma^2_{historical,i,t}} \right] \tag{11}$$

for each shrinkage estimator.

Then the following regression is calculated within each monthly cross-section and test statistics are taken from the time series of cross-sectional coefficients:

$$\text{Shrinkage}_{i,t} = \alpha_t + \omega_t \, \text{Sys}_{i,t} + \varepsilon_{i,t}. \tag{12}$$

Table 4 presents the results. The Ledoit-Wolf estimators, No Group or Group, and the Stein Random, imply that as the systematic portion of the risk increases, the positive coefficient indicates the shrinkage increases and the difference between the corrected and uncorrected estimators increases. However, it appears that for Stein High Dispersion, the negative coefficient indicates that the difference between the corrected and uncorrected estimators decreases.

The higher prediction errors of the Stein-type estimators might arise because the Efron-Morris method assumes normality while stock returns distributions are leptokurtic. In order to investigate whether the leptokurtosis increases the prediction errors, we first insert each stock's 6-month kurtosis into equation 12. The regression results are displayed in Table 5. Higher kurtosis induces a upward bias in all the volatility estimates as evidenced by the significant negative *kurt* coefficients for all estimators. The same results hold in both call and put options. However, the coefficients and significance for this kurtosis variable are virtually the same across all the estimators, implying that the impact of kurtosis is about the same across all estimators and not likely to be the reason for the higher prediction errors of the Stein-type estimators.

Table 6 examines the prediction errors in low and high kurtosis stocks. Each month, we sort all the stocks into two halves by their previous 6-month kurtosis and look at the prediction errors of all the estimators in each half. T-statistics for the difference between shrinkage estimators and the historical estimator are computed from the 100-month time series of each monthly difference in errors with a Newey-West correction for autocorrelation using eight lags.

For call options, the prediction errors using the historical and Ledoit-Wolf estimators are significantly lower than Stein in the low kurtosis group. For Stein-type estimators, the gaps between low and high kurtosis groups are much smaller. Therefore, among low kurtosis stocks, the prediction errors for the Stein-type estimators remain significantly higher than for the historical and Ledoit-Wof estimators. In the high kurtosis group, all estimators have similar prediction errors and the differences are not significant statistically. Hence, the generally higher prediction errors exhibited by the Stein estimators can be attributed mostly to low kurtosis stocks.

The results for put options are quite similar except that the Stein-type estimators produce errors that are also significantly greater than the historical estimator even for high kurtosis stocks, though the significance level is higher for low kurtosis stocks as it is for calls.

We also examine whether an optimal shrinkage estimator that minimizes the sum of squared errors is important for the particular application of volatility estimation. To do this, we compare LW's optimal shrinkage to a random average of combining the historical product moment sample matrix and the target matrix. In a similar comparison for estimation of the covariance matrix both Jagannathan and Ma (2003) and Disatnik and Benninga (2004) report that optimal shrinkage is no better than randomly choosing between the sample matrix and the target matrix, and thus optimality is not worth the effort.[9] We find that the LW optimal shrinkage estimator is much better than the random average of the sample matrix and the target.

5. Conclusion

A volatility bias in option prices was first uncovered by Black-Scholes (1972). They demonstrated that their model over-priced options on relatively high volatility stocks and under-priced options on relatively low volatility stocks. We thought that this bias might have nothing to do with the Black-Scholes model but instead could be attributable to sampling error because it is always observed in cross section with inter stock differences. If this is true, this bias would be observed with any option pricing model on any

[9] Jagannathan and Ma (2003), p. 1667, and Disatnik and Benninga (2007), p. 60 report a random average does as well as optimal shrinkage. Disatnik and Benninga state, "Theoretically, the shrinkage estimator should perform better than any other weighted average of the two estimators, as the proportions in the weighted average of the shrinkage estimator are obtained from minimizing the quadratic risk (of error) function of the combined estimator. Yet it seems that, in practice, estimating these specific proportions gives rise to a new type of error, and overall the shrinkage estimator does not perform better than the random average."

underlying, not just equity, but also fixed income securities, mortgages, foreign exchange, and commodities. To investigate this issue, we implemented the alternative variance estimators of James-Stein and Ledoit-Wolf, which correct historical volatility estimates by shrinking them toward a central value, thereby reducing their cross-sectional dispersion. While both shrinkage estimators utilize the covariance matrix, Ledoit-Wolf is unique because it does not require matrix inversion, and thus it does not require grouping the random variables because the number of stocks can exceed the number of observations.

First, we verify that the same bias Black-Scholes originally observed was present and very significant in both put and call option prices for the 100 months during the period January, 1996 through April, 2004. Second, we find that shrinkage variance estimators can eliminate this volatility bias, independent of the presence of the moneyness bias. Third, we uncover a difference between the Ledoit-Wolf and Stein estimators; the former does not increase the prediction error, but the latter significantly increase prediction error, especially for stocks with low kurtosis. Fourth, we demonstrate the Ledoit-Wolf estimators and Stein Random, imply that as the systematic portion of the risk increases, the positive coefficient indicates the shrinkage increases and the difference between the corrected and uncorrected estimators increases. However, it appears that for Stein High Dispersion, the negative coefficient indicates that the difference between the corrected and uncorrected estimators decreases. Finally, we show that the optimal shrinkage estimator of Ledoit-Wolf is superior to a random combination of the sample matrix and the target for this volatility estimation problem.

References

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen, 1997, Empirical performance of alternative option pricing models, *Journal of Finance* 52, 2003–2049.

Black, Fischer, 1975, Fact and fantasy in use of options, *Financial Analyst Journal* 31, 36–41,61–72.

———, and Myron Scholes, 1972, The valuation of option contracts and a test of market efficiency, *Journal of Finance* 27, 399–417.

———, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637–659.

Boyle, Phelim P., and A. L. Ananthanarayanan, 1977, The impact of variance estimation in option valuation models, *Journal of Financial Economics* 5, 375–387.

Butler, J. S., and Barry Schachter, 1986, Unbiased estimation of the black-scholes formula, *Journal of Financial Economics* 15, 341–357.

Cox, John C., and Stephen A. Ross, 1976, The valuation of options for alternative stochastic processes, *Journal of Financial Economics* 3, 145–166.

Disatnik, David J., and Simon Benninga, 2007, Shrinking the covariance matrix, *Journal of Portfolio Management* pp. 55–63.

Efron, Bradely, and Carl Morris, 1975, Data analysis using stein's estimator and its generalizations, *Journal of American Statistical Association* 70, 311–319.

———, 1976, Multivariate empirical bayes and estimation of covariance matrices, *Annals of Statistics* 4, 22–32.

Fama, Eugene, and J. D. MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *Journal of Political Economy* 81, 607–636.

Garman, Mark B., and Michael J. Klass, 1980, On the estimation of security price volatilities from historical data, 53, 67–79.

Geske, Robert, 1979, A note on an analytical valuation formula for unprotected American call options on stocks with known dividends, *Journal of Financial Economics* 7, 375–380.

———, and Richard Roll, 1984a, Isolating the observed biases in American call option pricing: an alternative variance estimator, Anderson UCLA Working Paper.

———, 1984b, On valuing American call options with the Black-Scholes European formula, *Journal of Finance* 39, 443–455.

———, and Kuldeep Shastri, 1983, Over-the-counter option market dividend protection and "biases" in the black-scholes model: A note, *Journal of Finance* 38, 1271–77.

Geske, Robert, and Walter Torous, 1990, Black-Scholes option pricing and robust variance estimation, in Stuart Hodges, ed.: *Options: Recent Advances in Theory and Practice.* Manchester University Press.

———, 1991, Skewness, kurtosis, and Black-Scholes option mis-pricing, *Statistical Papers* 32, 299–309.

Heston, Steven L., 1993, A closed form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6, 327–343.

———, and Saikat Nandi, 2000, A closed-form garch option valuation model, *Review of Financial Studies* 13, 585–625.

Hull, John, and Alan White, 1987, A closed form solution for options with stochastic volatility, *Journal of Finance* 42, 281–300.

Jagannathan, Ravi, and Tongshu Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* 58, 1651–1683.

James, W., and C. Stein, 1961, Estimation with Quadratic Loss, in *Proceedings of the Fourth Berkeley Symposium on Mathetical Statistics and Probability* , vol. 1 pp. 361–379 Berkeley. University of California Press.

Karolyi, G. Andrew, 1993, A bayesian approach to modeling stock return volatility for option valuation, *Journal of Financial and Quantitative Analysis* 28, 579–594.

Ledoit, Olivier, and Michael Wolf, 2003, Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *Journal of Empirical Finance* 10, 603–621.

——— , 2004a, Honey, i shrunk the covariance matrix, *Journal of Portfolio Management* Summer, 110–119.

——— , 2004b, A well-conditioned estimator for large dimensional covariance matices, *Journal of Multivariate Analysis* 88, 365–411.

Macbeth, James D., and Larry J. Merville, 1979, An empirical examination of the black-scholes call option pricing model, *Journal of Finance* 34, 1173–1186.

Merton, Robert C., 1976, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics* 3, 125–144.

Muirhead, R. J., 1987, Developments in Eigenvalue Estimation, in *Advances in Multivariate Statistical Analysis* pp. 277–288 Dodrecht. D. Reidel Publishing Company.

Parkinson, Michael, 1980, The extreme value method for estimating the variance of the rate of return, *Journal of Business* 53, 61–77.

Rubinstein, Mark, 1985, Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active cboe option classes from august 2, *Journal of Finance* 40, 455–480.

Stein, Charles, 1955, Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution, in *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* , vol. 1 pp. 197–206 Berkeley. University of California Press.

Sterk, William, 1982, Tests of two models for valuing call options on stocks with dividends, *Journal of Finance* 37, 1229–1237.

Whaley, Robert E, 1982, Valuation of American call options on dividend-paying stocks: Empirical tests, *Journal of Financial Economics* 10, 29–58.

Table 1: Summary Statistics of Annualized Volatility Estimates

This table shows time series averages of cross-sectional summary statistics of the historical volatility estimates with 6-months of daily returns and corresponding shrinkage estimators. Stein Random is the volatility shrunk by the Efron-Morris formula in random groups. Stein High Dispersion is the volatility shrunk by Efron-Morris formula in groups formed to have larger volatility dispersion. Ledoit-Wolf is the volatility shrunk by the Ledoit-Wolf method, with and without grouping. The mean is the average across all time series and cross-sections. The std is the time series average of the cross-sectional standard deviation for each sample month. Minimum and maximum are the time series averages of, respectively, the cross-sectional minimum and maximum in each sample month.

	Implied	**Historical**	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
Mean	0.384	0.400	0.360	0.361	0.411	0.410
Std	0.146	0.167	0.106	0.102	0.142	0.144
Min	0.122	0.153	0.114	0.121	0.224	0.218
Max	1.036	1.146	0.704	0.645	1.108	1.100

Table 2: Volatility Biases

Fama-MacBeth type tests were conducted for the following cross-sectional specification, $log(\frac{\sigma_{imp,i,t}}{\hat{\sigma}_{i,t}}) = \alpha + \beta\hat{\sigma}_{i,t} + \gamma X_{i,t} + \epsilon_{i,t}$. The upper panel is for call options, the lower panel for put options. All t-statistics are computed from the 100-month time series of cross-sectional coefficients with a Newey/West correction for autocorrelation using eight lags and are reported below the corresponding coefficient means. The five columns correspond to different volatility estimators. Historical is the standard estimator. Stein Random is an estimator shrunk by the Efron-Morris formula in random groups. Stein High Dispersion is an estimator shrunk by the Efron-Morris formula in groups with large volatility dispersion. Ledoit-Wolf is an estimator shrunk by the Ledoit-Wolf method, with and without grouping.

Panel A: Call Options

	Historical	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
Const	-0.347	-0.399	-0.458	-0.556	-0.535
	(-2.751)	(-3.300)	(-3.736)	(-4.746)	(-4.400)
$\widehat{\sigma}$	-0.437	-0.119	0.010	0.001	-0.040
	(-7.596)	(-0.943)	(0.072)	(0.019)	(-1.417)
Moneyness	0.481	0.479	0.487	0.480	0.478
	(4.145)	(3.830)	(3.915)	(4.085)	(4.054)
Ave. R^2	0.154	0.053	0.060	0.045	0.041
Ave. Cross-section	494	494	494	494	494

Panel B: Put Options

	Historical	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
Const	-0.275	-0.328	-0.384	-0.479	-0.462
	(-2.291)	(-2.892)	(-3.341)	(-4.422)	(-4.144)
$\widehat{\sigma}$	-0.474	-0.170	-0.048	-0.044	-0.084
	(-7.752)	(-1.299)	(-0.352)	(-1.150)	(-2.670)
Moneyness	0.443	0.445	0.453	0.441	0.442
	(3.986)	(3.702)	(3.790)	(4.004)	(4.035)
Ave. R^2	0.185	0.056	0.062	0.048	0.045
Ave. Cross-section	488	488	488	488	488

Table 3: Prediction Errors of Volatility Estimators

Average prediction errors are computed for the historical volatility estimator and all four shrinkage estimators, measured by the root mean square of $log(\frac{\sigma_{imp,i,t}}{\hat{\sigma}_{i,t}})$. T-statistics for the difference between shrinkage estimators and the historical estimator are computed from the 100-month time series of each monthly difference in errors with a Newey-West correction for autocorrelation using eight lags, and are reported below the corresponding coefficient means. The five columns correspond to different volatility estimators. Historical is the standard estimator. Stein Random is an estimator shrunk by the Efron-Morris formula in random groups. Stein High Dispersion is an estimator shrunk by the Efron-Morris formula in groups with large volatility dispersion. Ledoit-Wolf is an estimator shrunk by the Ledoit-Wolf method, with and without grouping. In Theil's decomposition, UM is the proportion due to bias in the forecasts. UR is the error due to a low correlation between the actual and the forecast. UD is the remaining part. T-statistics in the parentheses are computed using Newey-West with 8 lags.

Panel A: Call Options

	Historical	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
MSE	0.042	0.057	0.054	0.043	0.043
		(3.473)	(3.163)	(0.238)	(0.276)
UM	0.010	0.012	0.012	0.014	0.014
		(0.608)	(0.376)	(1.003)	(0.918)
UR	0.005	0.003	0.003	0.001	0.001
		(-1.543)	(-2.375)	(-6.994)	(-7.714)
UD	0.027	0.041	0.040	0.027	0.028
		(6.760)	(6.241)	(0.481)	(0.893)

Panel B: Put Options

	Historical	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
MSE	0.039	0.056	0.054	0.037	0.037
		(3.872)	(3.573)	(-0.637)	(-0.553)
UM	0.010	0.015	0.014	0.012	0.012
		(1.648)	(1.367)	(0.497)	(0.464)
UR	0.006	0.004	0.003	0.001	0.001
		(-1.974)	(-2.979)	(-7.789)	(-8.266)
UD	0.024	0.037	0.036	0.024	0.025
		(7.989)	(7.664)	(0.569)	(1.082)

Table 4: Shrinkage and Systematic Risk

For each of 100 month, cross-sectional regressions were computed to explain the shrinkage proportion as a function of the systematic risk estimated over the previous two years (approximately.) T-statistics, in parentheses, are computed from the time series of cross-sectional coefficients using a Newey-West correction for autocorrelation with 8 lags. The columns correspond to four alternative shrinkage estimators. Stein Random is an estimator shrunk by the Efron-Morris formula in random groups. Stein High Dispersion is an estimator shrunk by the Efron-Morris formula in groups with large volatility dispersion. Ledoit-Wolf is an estimator shrunk by the Ledoit-Wolf method, with and without grouping.

	Stein Random	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
Const	0.146	0.027	0.039	0.133
	(5.106)	(1.711)	(4.472)	(13.495)
$Sys_{i,t}$	0.255	-0.417	0.396	0.056
	(2.196)	(-5.880)	(5.670)	(2.418)
Ave. R^2	0.110	0.140	0.046	0.007

Table 5: Control for Kurtosis

Fama-MacBeth type tests were conducted for the following cross-sectional specification, $log(\frac{\sigma_{imp,i,t}}{\hat{\sigma}_{i,t}}) = \alpha + \beta\hat{\sigma}_{i,t} + \gamma X_{i,t} + \epsilon_{i,t}$. The upper panel is for call options, the lower panel for put options. All t-statistics are computed from the 100-month time series of cross-sectional coefficients with a Newey/West correction for autocorrelation using eight lags and are reported below the corresponding coefficient means. The five columns correspond to different volatility estimators. Historical is the standard estimator. Stein Random is an estimator shrunk by the Efron-Morris formula in random groups. Stein High Dispersion is an estimator shrunk by the Efron-Morris formula in groups with large volatility dispersion. Ledoit-Wolf is an estimator shrunk by the Ledoit-Wolf method, with and without grouping.

Panel A: Call Options

	Historical	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
Const	-0.320	-0.384	-0.446	-0.538	-0.516
	(-2.620)	(-3.319)	(-3.812)	(-4.799)	(-4.413)
$\widehat{\sigma}$	-0.352	0.013	0.161	0.101	0.057
	(-8.010)	(0.116)	(1.324)	(2.822)	(2.315)
Moneyness	0.477	0.475	0.484	0.477	0.475
	(4.282)	(3.944)	(4.033)	(4.231)	(4.200)
Kurtosis	-0.009	-0.010	-0.010	-0.009	-0.009
	(-23.032)	(-16.650)	(-16.061)	(-20.776)	(-21.268)
Ave. R^2	0.241	0.121	0.134	0.136	0.131
Ave. Cross-section	494	494	494	494	494

Panel B: Put Options

	Historical	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
Const	-0.251	-0.316	-0.375	-0.464	-0.445
	(-2.176)	(-2.918)	(-3.419)	(-4.490)	(-4.178)
$\widehat{\sigma}$	-0.395	-0.044	0.094	0.050	0.007
	(-8.162)	(-0.374)	(0.765)	(1.542)	(0.325)
Moneyness	0.442	0.445	0.453	0.440	0.441
	(4.170)	(3.862)	(3.953)	(4.200)	(4.238)
Kurtosis	-0.009	-0.009	-0.009	-0.008	-0.008
	(-28.256)	(-16.760)	(-16.355)	(-23.332)	(-24.646)
Ave. R^2	0.267	0.124	0.134	0.139	0.135
Ave. Cross-section	488	488	488	488	488

Table 6: Prediction Errors for Low and High Kurtosis Stocks

Average prediction errors are computed for the historical volatility estimator and all four shrinkage estimators, measured by the root mean square of $log(\frac{\sigma_{imp,i,t}}{\hat{\sigma}_{i,t}})$ for stocks grouped by kurtosis over the previous six months. T-statistics for the difference between shrinkage estimators and the historical estimator are computed from the 100-month time series of each monthly difference in errors with a Newey-West correction for autocorrelation using eight lags. These t-statistics are given in parentheses below each mean prediction error. The five columns correspond to different volatility estimators. Historical is the standard estimator. Stein Random is an estimator shrunk by the Efron-Morris formula in random groups. Stein High Dispersion is an estimator shrunk by the Efron-Morris formula in groups with large volatility dispersion. Ledoit-Wolf is an estimator shrunk by the Ledoit-Wolf method, with and without grouping. In Theil's decomposition, UM is the proportion due to bias in the forecasts. UR is the error due to a low correlation between the actual and the forecast. UD is the remaining part. T-statistic in the parentheses are computed using Newey-West with 8 lags.

Panel A: Call Options

	Historical	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
Low-Kurt	0.032	0.055	0.051	0.033	0.033
		(4.258)	(4.088)	(0.317)	(0.306)
High-Kurt	0.052	0.058	0.058	0.053	0.053
		(1.349)	(1.306)	(0.198)	(0.255)

Panel B: Put Options

	Historical	**Stein Random**	**Stein High Disp**	**Ledoit No Group**	**Ledoit Group**
Low-Kurt	0.031	0.056	0.052	0.028	0.037
		(4.329)	(4.077)	(-1.206)	(2.268)
High-Kurt	0.047	0.056	0.055	0.046	0.047
		(1.914)	(1.861)	(-0.323)	(-0.085)