MS&E 444 FINAL PROJECT REPORT

Optimal Trading of a Mean-Reverting Process

Instructor: Prof. Kay Giesecke

TA: Benjamin Armbruster

Group Members:

Chen Tze Wee

Shih-Arng Pan

Wei Wang

Ren Fung Yu

Introduction

In *Arbitrage Under Power*, Boguslavsky and Boguslavskaya solve for the optimal strategy to maximize power utility at the end of a period for an Ornstein-Uhlenbeck process. Our implementation extends their result by addressing practical constraints, particularly margin requirements and transaction costs, and uses a moving window to obtain dynamic model parameters. As our results show, these modifications represent an improvement over the original strategy, which is too aggressive when implemented with realistic Ornstein-Uhlenbeck parameters and incurs high transaction costs from continuous trading. However, many of our selected stock pairs did not converge in price out-of-sample, thus lowering our returns. We conclude that to take advantage of our strategy, a more reliable way of identifying mean-reverting processes in the market must be found.

Summary of the original strategy

Like the original strategy, our implementation uses the Ornstein-Uhlenbeck model to describe the mean-reverting process. It is defined by the following equation:

$$dX_t = -kX_t + \sigma dz$$

Essentially, k describes the rate of mean-reversion, σ is the volatility, and dz is a standard Brownian motion. Boguslavsky and Boguslavskaya's solution is expressed by α_t, the optimal position to maximize end-of-period power utility of the following form:

$$U(W_T) = \frac{1}{\gamma} W_T^{\gamma}$$

$W_T$is the terminal wealth and γ is a risk aversion parameter such that $-\infty < \gamma < 1$.

The optimal position, α_t is a function of current wealth W_t, current price X_t, time remaining $\tau = \mathrm{T} - \mathrm{t}$, and k and σ:

$$\alpha_t = -W_t X_t D(\tau) . \frac{k}{\sigma^2}$$

where $D(\tau)$ is defined as follows:

$$v = \frac{1}{\sqrt{1-\gamma}}$$

$$C(\tau) = \cosh v\tau + v \sinh v\tau$$

$$C^{'}(\tau) = \frac{dC(\tau)}{d\tau} = v \sinh v\tau + v^2 \cosh v\tau$$

$$D(\tau) = \frac{C^{'}(\tau)}{C(\tau)}$$

Modifications to strategy

<u>Margins</u>

Margins are implemented as a simple add-on condition:

$$\text{if } \left|\alpha_{opt-original}\right| > m, \text{ then redefine } \alpha_{opt} = \pm m$$

where $\alpha_{opt-original}$ is the theoretically optimal position described before, and α_{opt} is the best position our trader's allowed to hold. Here, m acts as a maximal position the trader is allowed to long/short.

We tested two different margin conditions:

(1) Wall Margin: m is kept constant throughout all time.

(2) Scaled Margin: m is a function of wealth, defined implicitly as follows:

$$m\, \sigma_{severe}\sqrt{dt} = W_t\, R_{loss} = \text{maximal loss allowed per unit time}$$

where W_t is the trader's total wealth at time t. σ_{severe} is chosen to be the number of standard deviations the market has to move per unit time to constitute a rare and severe downturn. R_{loss} is chosen to be the trader's maximal allowed percentage loss of current wealth per unit time. These parameters can be chosen to match the trader's risk appetite.

Statistically, the scaled-margin condition generates superior returns on simulated mean-reverting processes

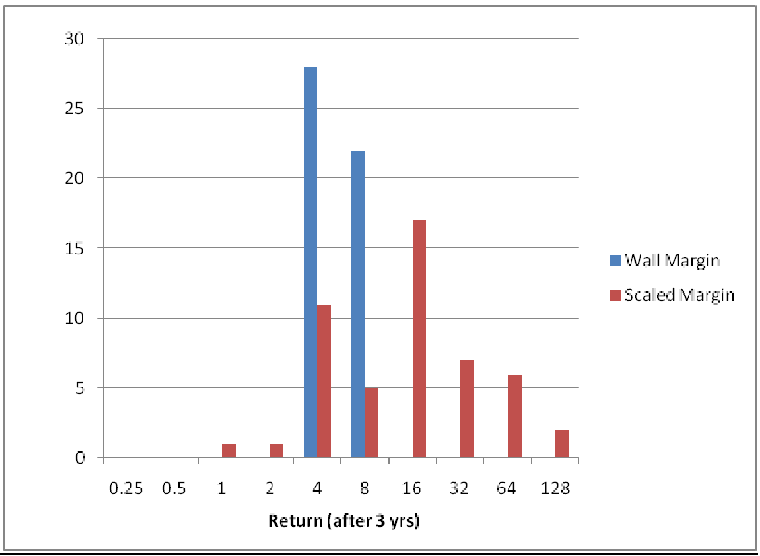

Figure 1: 3-yr return of Wall Margin vs. Scaled Margin conditions

Transaction costs

Assume there is a fixed transaction cost of c. Then the wealth dynamics follow one of the following:

$$dW_{cur} = \alpha_{cur}\, dX \qquad = -\alpha_{cur}\, k\, X_t\, dt + \alpha_{cur}\, \sigma\, dz$$

$$dW_{opt} = \alpha_{opt}\, dX - c \qquad = -\alpha_{opt}\, k\, X_t\, dt + \alpha_{opt}\, \sigma\, dz - c$$

α_{cur} is the current position the trader holds, and dW_{cur}, dW_{opt} are the changes in wealth in the upcoming time step from holding α_{cur} or α_{opt}, respectively.

We optimize the expected utility for the upcoming time step[1]. We change positions from α_{cur} to α_{opt} (we trade) when:

$$E\big(U(W_t + dW_{opt})\big) > E(U(W_t + dW_{cur}))$$

The only variable for the expectancy is dz, which is what we integrate over. Hence, we only trade when the expected utility gain outweighs the assumed transaction cost.

Assume for the upcoming time step, utility is approximately linear. Then the Brownian motion term dz does not contribute to the integral, and our trading condition simplifies to:

$$\big(\alpha_{opt} - \alpha_{cur}\big)dX > c$$

However, our simulations show that the linear-utility approximation is too aggressive, frequently leading to large losses in real market conditions. Hence, we assume power-utility for the next time step, and the trading condition is:

$$\int dz\, \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\gamma} (W_t - \alpha_{opt}\, k\, X_t\, dt + \alpha_{opt}\, \sigma \sqrt{dt}\, z - c)^\gamma$$

$$> \int dz\, \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\gamma} (W_t - \alpha_{cur}\, k\, X_t\, dt + \alpha_{cur}\, \sigma \sqrt{dt}\, z)^\gamma$$

Note that we changed the Brownian motion notation from dz -> $\sqrt{dt}\,z$, and then we integrate over a normal distribution of z. The theoretical range of integration should be from $-\infty \rightarrow \infty$, however, we cannot take arbitrary power of a negative number. In practice, we numerically integrate z from $(-\infty, z_0)$ or (z_0, ∞), for $\alpha < 0$ or $\alpha > 0$, respectively, where the term representing $W_t + dW$ approaches zero when $z = z_0$. This z_0 is different for the two integrals.

[1] Note that for a log-utility trader, this is equivalent to optimizing utility for the end of the time horizon, since the trader does not hedge intertemporally.

We can also optimize the expected utility for *s* time steps into the future. For a Ornstein-Uhlenbeck process, $E(X_{t+s}|X_t) = X_t e^{-ks}$, and $\mathrm{Var}(X_{t+s}|X_t)=(\frac{1-e^{-2ks}}{2k})\sigma^2$, so the trading condition becomes:

$$\int dz\, \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\gamma}(W_t - \alpha_{opt}\, X_t(1-e^{-ks})\, dt + \alpha_{opt}\, \sigma \sqrt{\frac{1-e^{-2ks}}{2k}}\, z - c)^{\gamma}$$

$$> \int dz\, \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\gamma}(W_t - \alpha_{cur}\, X_t(1-e^{-ks})\, dt + \alpha_{cur}\, \sigma \sqrt{\frac{1-e^{-2ks}}{2k}}\, z)^{\gamma}$$

Our simulations show that the trading condition is insensitive to *s*, so we use *s*=1 from now on.

Data and parameter selection

We implemented our strategy on the 18 stock pairs that had the highest correlation of daily returns (>0.75) between January 2003 and December 2004. The list of stocks with their correlations are presented in the Appendix. To adjust for existing trends in the price spreads, we plotted a least-squared regression line through the price spread series and used this as the "adjusted" x-axis for the trading period, January 2004 to December 2005. This adjustment is illustrated in Figures 1a and 1b.

The values for k and σ that we used were maximum likelihood estimators, calculated from the January 2003 to December 2004 data, with dt equal to $\frac{1}{250}$. Note that the actual value of dt is only important to the extent that it is used to calculate k and σ; the latter are renormalizable for a different dt. In our implementation, $\gamma = -0.1$ and transaction costs, c, were constant at 0.15% of our initial wealth.

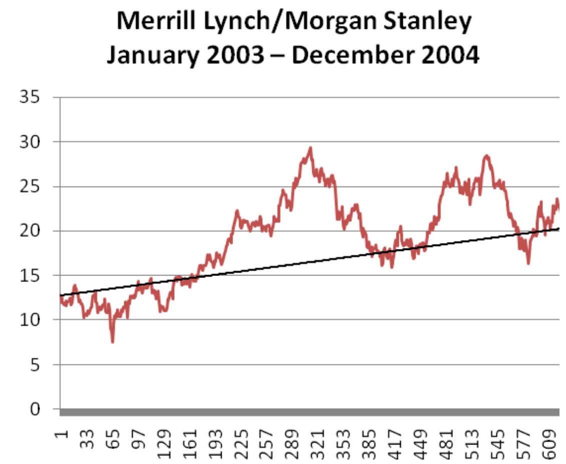

Figure 2a: Unadjusted price spread time series

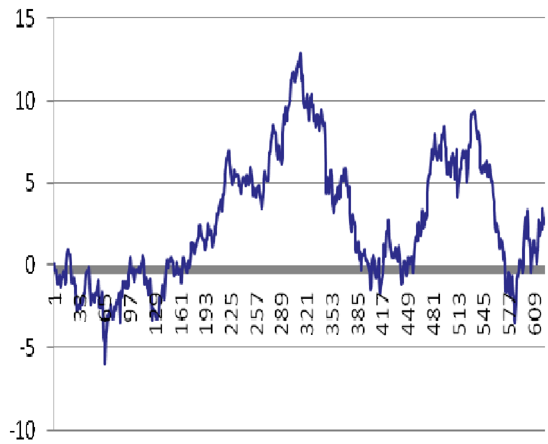

Figure 2b: Adjusted price spread time series

Results

The following examples are chosen to be representative of our strategies' performances under different market conditions.

<u>Chevron - Exxon</u>

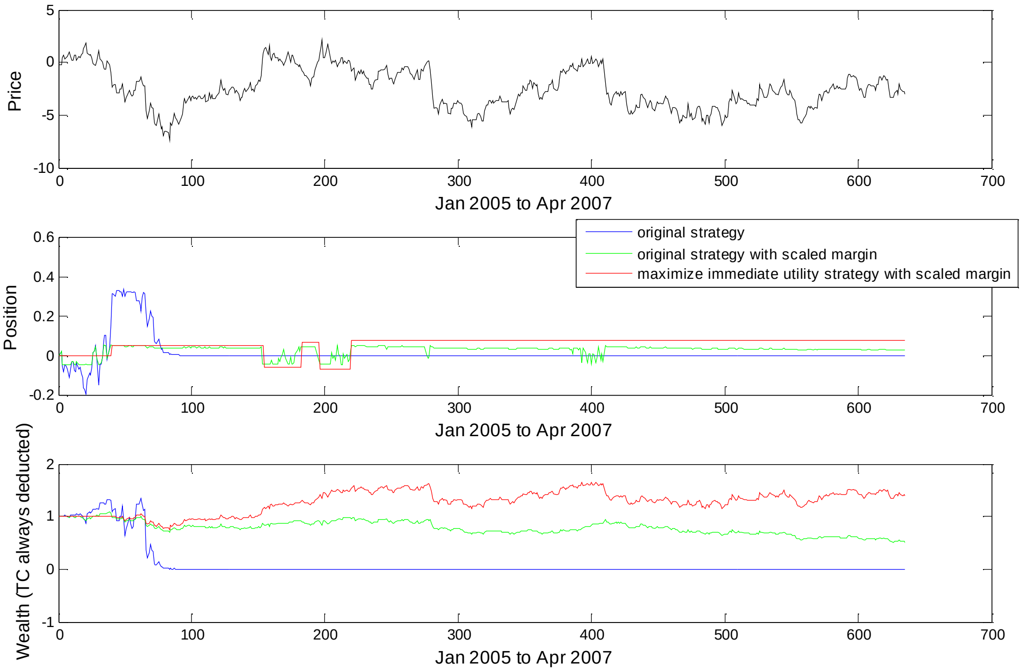

Figure 3: Chevron-Exxon, with scaled margins

Using the market data we selected in the previous section, we analyzed the performance of several strategies, and compared our modifications with the original strategy developed by Boguslavsky and Boguslavskaya.

In Figure 2, we plotted the results of our analysis of the stock pair created with Chevron and Exxon (Chevron-Exxon). The price spread from January 2005 to April 2007 is plotted on the top of the figure. The k, σ parameters of the model used were estimated using training data before the trading window (January 2004 to December 2005), and stays constant throughout the trading period. The second plot in the figure shows the positions taken up by three different strategies studied, and the final plot shows the financial performance of these strategies. The three strategies studied in this part of the work are the original strategy by Boguslavsky and Boguslavskaya (blue line), original strategy with a scaled margin (green), and a new strategy that maximizes immediate utility with a scaled margin (red). A fixed transactions cost equivalent to 0.15% of the starting wealth was deducted for each trade in this study.

The original strategy is too risky, and quickly gets bankrupted during extreme market movements. Therefore, the original strategy is not feasible in the real world. Using the original strategy with scaled margins, we limit the loss of the total wealth per day to 5% in the event of the 3-sigma move against your position ($\sigma_{severe} = 3$, $R_{loss} = 5\%$). This modification has better performance, and does not reach bankruptcy. This is evidence that the risk control measures are working. However, the strategy still trades continuously, and transaction costs are diminishing profits significantly. The third approach, maximizing immediate utility with scaled margins, reduces the impact of transaction costs by limiting the number of times the strategy trades. The positions taken up by this strategy are clearly discrete, as shown in the second plot. By also controlling the risk this strategy takes using scaled margins, the final performance of this approach is the best of the three studied in this case, resulting in the highest final wealth.

Baker Hughes – Schlumberger with Moving Window

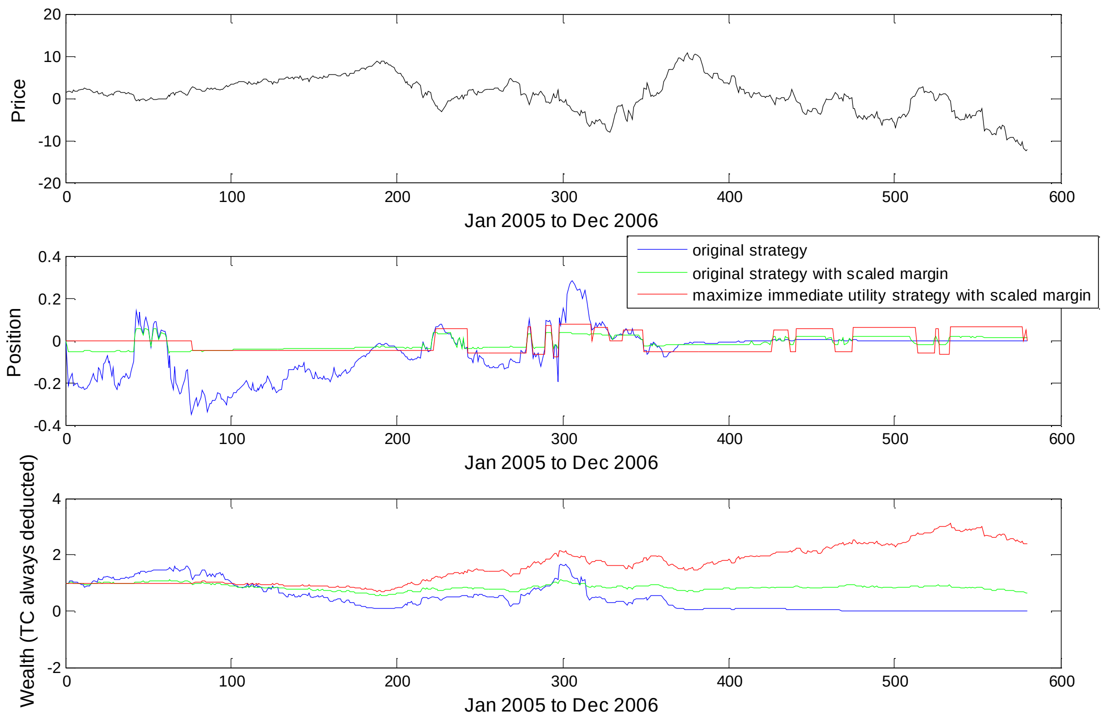

Figure 4: Baker Hughes – Schlumberger, with scaled margins and a 1.5 year moving window

One improvement to the previous study is to consider a dynamic model of mean-reversion, where a moving window of data is used to re-estimate the input parameters constantly. Figure 3 shows the results of this study. A moving window of 1.5 years is used because it gives the best results of several window sizes we analyzed, although the returns are not very sensitive to the window size.

Once again, we see that the original strategy goes bankrupt fairly quickly due to a lack of risk control and transaction costs. The original strategy with scaled margins works better, but its performance is significantly decreased by transactions costs incurred by continuously trading. The new strategy of maximizing immediate utility results in the highest final wealth, which shows that the risk control measures and utility maximization considerations are effective. These conclusions are similar to that of the previous study as shown in Figure 3.

<u>CitiGroup – Lehman Brothers, poor mean reversion</u>

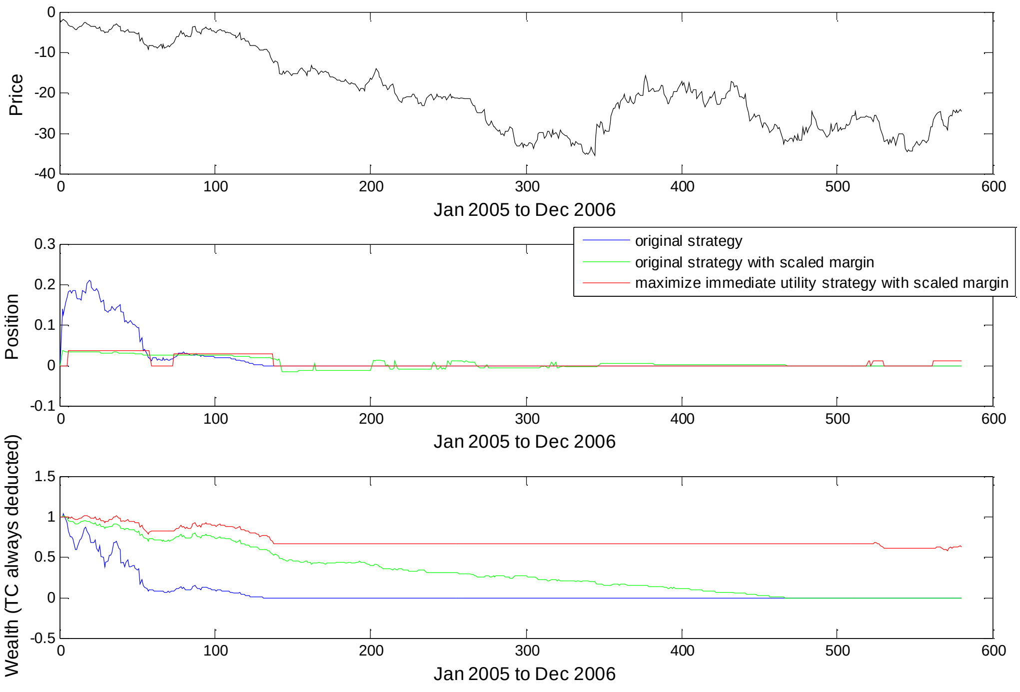

Figure 5: CitiGroup – Lehman Brothers, for a failed mean-reverting process

The fundamental assumption of all our trading strategies implemented is that the stock pairs being traded exhibit a mean-reverting behavior. Unfortunately, even though we already screened out stock pairs and only analyzed pairs that have highly correlated daily returns, the spread between correlated stocks may no longer be mean-reverting out-of-sample. Figure 5 illustrates the consequences of using our trading strategies on stocks with poor mean reversion.

Again, the original trading strategy built by Boguslavsky and Boguslavskaya goes to bankruptcy. This result is expected, as the original strategy tends to bankruptcy even if the stock pair mean-reverts. With scaled margins, it is clear that the risk control measures are effective to a certain extent, as the strategy only bankrupts after about 450 days of trading, which is much longer than the original strategy which bankrupts in about 120 days. The final strategy which maximizes utility while using scaled margins is the most robust of the three strategies, as it never goes bankrupt. This strategy figures out that current market behavior is not beneficial for trading given one's low wealth, thus heavily decreasing trading frequency. At the same time, the moving window calibration of k and σ captures the decline in mean-reversion, which forces the strategy to cut its position. The final wealth is lower than the starting wealth, but it is far from bankruptcy, and perhaps, this is the best behavior one can hope for when fundamental assumptions such as mean-reversion of the stock pairs are violated.

Annual return statistics

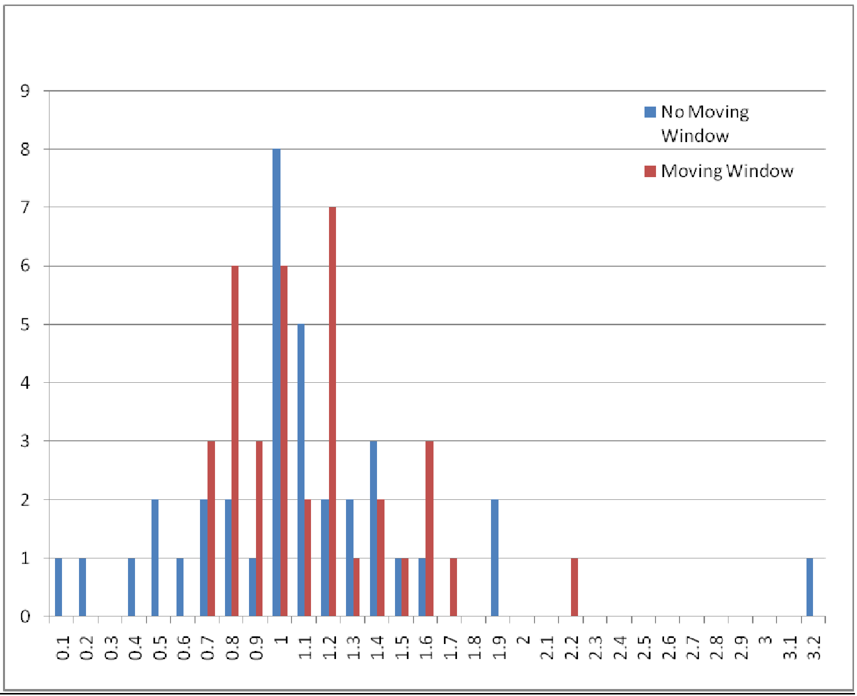

Figure 6: Annual Return Histogram, Moving Window vs. No Moving Window

We implement our best strategy, maximizing immediate utility with scaled margins, on the out-of-sample spreads of 18 pairs of correlated stocks chosen earlier. Using a moving window to actively recalibrate k and σ is clearly superior, increasing returns while decreasing volatility of returns. The following table compares the annual returns of our strategy, with and without the moving window.

	Moving Window	**No Moving Window**
Return	1.0764	0.3428
Volatility	1.0418	0.5511

Conclusions

Boguslavksy and Bogusavskaya's strategy exhibits a strong tendency towards bankruptcy due to high transaction costs from continuous trading. Furthermore, the original strategy tends to take overly-aggressive positions when using k and σ estimates from historical stock pairs. This behavior typically results in massive, often irrecoverable, short-term losses.

We found that simple wall margins on the position reduce risk significantly, and that the strategy's performance can be further improved by scaling the margins in proportion to current wealth. We also managed to reduce transaction costs significantly by reducing the frequency of position changes – our implementation only trades when the expected utility gain outweighs the assumed transaction cost. In general, our strategy performed better when we used a moving window to generate k and σ, as opposed to using static estimators.

Overall, our strategy's returns were positive, but were not high enough to meet an assumed stock market average of around 10% per year. We were hampered by frequent failures of our chosen stock pairs to mean-revert. Our strategy did, however, suffer limited losses compared to the original strategy in the cases where returns were negative. This was mostly due to the scaled margins, which reduced our strategy's position as losses occurred.

To improve performance, it is likely that we will need to look elsewhere to find more consistent mean-reverting financial securities. Otherwise, further study can be done on models that more accurately describe correlated stock price spreads compared to Ornstein-Uhlenbeck processes.

References

1. M. Boguslavsky, E. Boguslavskaya. Arbitrage under Power. RISK magazine, June 2004, pp. 69-73.

Appendix

Chosen stock pairs with highest correlation of daily returns

Stock Pair	Correlation
AMERICAN EXPRESS, CITIGROUP	0.703
BAKER HUGHES, SCHLUMBERGER	0.746
BANK OF AMERICA, CITIGROUP	0.705
BANK OF AMERICA, JP MORGAN CHASE & CO.	0.678
BANK OF AMERICA, REGIONS FINL.NEW	0.666
BANK OF AMERICA, WACHOVIA	0.701
CHEVRON, EXXON MOBIL	0.723
CISCO SYSTEMS, EMC	0.664
CITIGROUP, JP MORGAN CHASE & CO.	0.753
CITIGROUP, LEHMAN BROS.HDG.	0.679
CITIGROUP, MERRILL LYNCH & CO.	0.711
CITIGROUP, MORGAN STANLEY	0.723
HALLIBURTON, SCHLUMBERGER	0.683
INTL.PAPER, WEYERHAEUSER	0.705
JP MORGAN CHASE & CO., MORGAN STANLEY	0.683
LEHMAN BROS.HDG., MERRILL LYNCH & CO.	0.791
LEHMAN BROS.HDG., MORGAN STANLEY	0.790
MERRILL LYNCH & CO., MORGAN STANLEY	0.816