Universität St.Gallen

Relative Implied Volatility Arbitrage with Index Options

Another Look at Market Efficiency

Manuel Ammann and Silvan Herriger

May 2001 Discussion paper no. 2001-06

Department of Economics University of St. Gallen

Editor:	Prof. Jörg Baumberger University of St. Gallen Department of Economics Bodanstr. 1 CH-9000 St. Gallen Phone ++41 71 224 22 41 Fax ++41 71 224 28 85 Email joerg.baumberger@unisg.ch
Publisher:	Forschungsgemeinschaft für Nationalökonomie an der Universität St. Gallen Dufourstrasse 48 CH-9000 St. Gallen Phone ++41 71 224 23 00 Fax ++41 71 224 26 46
Electronic Publication:	www.fgn.unisg.ch/public/public.htm

Relative Implied Volatility Arbitrage with Index Options

Another Look at Market Efficiency[1]

Manuel Ammann and Silvan Herriger[2]

Authors' addresses: Dr. Manuel Ammann
Silvan Herriger, lic. oec. HSG

University of St. Gallen
Swiss Institute of Banking and Finance
Rosenbergstrasse 52
CH-9000 St. Gallen
Tel. ++41 71 224 70 90
Fax ++41 71 224 70 88
Email manuel.ammann@unisg.ch
Website www.sbf.unisg.ch

[1] All data used in this study was provided by the Bloomberg© Professional information service through Wegelin & Co., St.Gallen. We thank David Rey and Maud Capelle for helpful comments.

[2] Manuel Ammann is a lecturer at University of St.Gallen. Silvan Herriger is currently employed as an index derivatives trader at J.P. Morgan.

Abstract

We investigate statistical arbitrage strategies for index options. To test the efficiency of markets in pricing relative implied volatilities in highly correlated markets, U.S. stock indices for which listed options are available are matched into pairs according to their degree of correlation. The interrelationship over time of the three most highly correlated and liquid index pairs is then analyzed. Based on this analysis, the relative implied volatility relationships are calculated. If such a relationship is violated, a relative mispricing is identified. We find that, although many theoretical mispricings can be observed, only a fraction of them are large enough to be used profitably in the presence of bid-ask spreads and transaction costs. A simple no-arbitrage barrier is thus used to identify significant mispricings and a statistical arbitrage trade is implemented every time such a mispricing was recorded, the trades being on average profitable after deduction of transaction costs.

Keywords

Statistical arbitrage, index options, relative implied volatility, market efficiency

JEL Classification

G13, G15

Introduction

Arbitrage relationships on derivatives markets have been studied extensively. Option boundary conditions, as derived by STOLL (1969) and MERTON (1973), for example, have been subjected to numerous empirical studies, such as GOULD and GALAI (1974), KLEMKOSKY and RESNICK (1979), ACKERT and TIAN (1998), ACKERT and TIAN (1999), among others. Index arbitrage has been thoroughly investigated as well. Empirical studies include FIGLEWSKI (1984), CHUNG (1991), SOFIANOS (1993) and NEAL (1996). An example of options arbitrage in imperfect markets is FIGLEWSKI (1989). Clearly, the testing of market efficiency on derivatives markets using arbitrage relationships has drawn a great deal of interest.

Statistical arbitrage, however, be it on derivative or other markets, has received surprisingly little attention in the literature despite its high practical relevance. A possible cause may be the nature of the mispricings underlying statistical arbitrage. Statistical arbitrage is not based on theoretical, exact pricing relationships but rather on empirical, statistically established relationships. Consequently, it involves risk. Omitting the study of such forms of pricing relationships from research agendas altogether, however, may lead to an incomplete understanding of market mechanisms, and thus of market efficiency as well.

One study that has proceeded to test market efficiency on equity markets with a statistical arbitrage approach is GATEV et al. (1999), who investigate the relative pricing mechanism of securities that are close economic substitutes. Motivated by the widespread intermarket hedging activities in commodities markets, various pricing relationships for commodity spreads have been analyzed. This explains the presence of several papers relating to statistical arbitrage in such markets, such as JOHNSON et al. (1991) or POITRAS (1997).

A statistical arbitrage approach to test the efficiency of options markets has not been attempted yet. The aim of this study is thus to devise and implement a statistical arbitrage strategy that tests an aspect of market efficiency that the classical boundary conditions for options fail to reveal: the efficiency of markets in pricing relative risk in highly correlated markets.

There are eleven stock indices in the United States for which listed options are available. Several of these indices are very closely related. The close relationship between indices is often caused by a securities overlap, i.e., the same stocks are included in several indices (for example, every stock in the S&P 100 Index is also included in the S&P 500 Index). If two indices are highly correlated (because of a securities overlap or other reasons), it should be possible to calculate the relationship between their respective volatility levels. Consequently, a similar relationship must also be valid for the implied volatility levels of the respective index options. If the relation between the respective implied volatilities is significantly different from the relation observed between the two index volatilities, the option prices are misaligned, which should not occur in efficient markets. In this case, a statistical arbitrage strategy can be implemented to take advantage of this relative implied volatility mispricing.

Our statistical arbitrage methodology consists of several consecutive steps, briefly outlined below:

(i) First, after ensuring that the return time series are stationary, we calculate the correlation of the different indices. We then select the pairs with the highest correlation coefficients, and do not consider the other indices further.

(ii) In a second step, we study the relationship of the daily returns of index pairs by running an OLS regression in order to establish the past relationship between them. We also test the robustness of the relationships. Because the linear relationship between two indices is time-varying, we estimate statistical boundaries for the OLS coefficients.

(iii) We then establish a conditional forecast of future variance based on the past relationship between the index returns. This is to test (out-of-sample) the predictive powers of the boundaries estimated in the preceding section.

(iv) Once the predictive capacity of the mentioned boundaries is confirmed, we apply the estimated relationship to implied volatility, where a similar relationship should prevail.

(v) Based on the implied volatilities and on the riskless rate recorded every trading day, we calculate the corresponding option prices, incorporating bid-ask spreads in the process. This is to ensure that, should a mispricing of a certain significant magnitude be identified, an option strategy could be implemented and tested taking advantage of the mispricing.

(vi) Finally, we implement a simple arbitrage[1] trading strategy.

Data

In the United States, exchange traded options are available for eleven stock indices. They are the Standard & Poors's 500 index (SPX), the Standard & Poor's 100 Index (OEX), the NASDAQ-100 Index (NDX), the New York Stock Exchange Composite Index (NYA), the

Philadelphia U.S. TOP 100 Sector Index (PTPX), the Philadelphia Stock Exchange Utility Index (UTY), the Standard & Poor's Smallcap 600 Index (SML), the Standard & Poor's Midcap 400 (MID), the American Stock Exchange (AMEX) Major Market Index (XMI), the Russel 2000 Index (RTY), and the Dow Jones Industrial Average (INDU).

Following HARVEY and WHALEY (1991), the implied volatility of at-the-money options with the shortest maturity have been used for this study. At-the-money options are used because they contain the most information about volatility. The front month options are used because they are the most liquid. If less than 20 calendar days are left to expiry, the next available series is taken. This means that the implied volatilities used are calculated from options ranging from 20 to 50 calendar days to expiry, or an average of 35 calendar days. As 35 calendar days represent exactly 5 weeks, and an average week has 5 trading days, an average of 25 trading days shall be used in the calculations. The implied volatilities are calculated based on the closing prices of the options and the underlying. The volatility is derived from an average of the implied volatilities of at-the-money options.

The riskless interest rate is used to calculate the daily option prices. As outlined above, the options used have 35 calendar days left to expiry, on average. Therefore, the one month Eurodollar LIBOR is used as the riskless rate.

Analysis of the index time series

Selection of indices

For our study, we select the SPX, OEX, and NYA indices from the sample of available indices in the United States. The criteria for this selection are stationarity of the index return, high correlation, and liquidity of the market for index options.

To avoid spurious correlation and regression, we first tested the index time series for stationarity. The standard stationarity tests revealed that all return time series (based on continuously-compounded returns) can be considered stationary with the exception of the Philadelphia Stock Exchange Utility Index (UTY). The UTY index is therefore not considered any further.

We then calculate the correlation coefficients for the remaining ten indices. The correlation matrix is given in *Table 1*. We set a minimum of 0.95 for the correlation coefficient as a

---

[1] Unless indicated otherwise, the term arbitrage refers to statistical arbitrage in this paper.

criterion for inclusion in our index sample[2]. The correlation criterion is motivated by the conjecture that index pairs with high correlation exhibit a strong linear relationship between each other. This criterion leaves five pairs of indices to be considered for further calculations. Of these five pairs, the three indices with the most liquid option markets are chosen, namely SPX, OEX, NYA. The fact of using only liquid contracts ensures that potential arbitrage trading strategies can be executed.

That the S&P 500 Index (SPX) is highly correlated with the S&P 100 Index (OEX) is not surprising, as the latter is an integral part of the former. We would expect the OEX to be more volatile than the SPX, due to the comparative heaviness (500 stocks vs 100 stocks) of the SPX. Similarly, the overlap of the SPX with the NYSE Composite Index (NYA) is large, as a large number of stocks in the SPX are listed on the NYSE. The NYA is even heavier than the SPX, so we would expect the latter to be slightly more volatile. The triangle is closed by the pair OEX - NYA. This relationship is, however, more surprising. Apparently, the OEX tracks the heavier SPX so well that it even manages to track the NYA, which is heavier yet. Clearly, we expect the OEX to be more volatile than the NYA because the NYA is more diversified.

Relationship between index pairs

For every pair of indices, the daily returns of one index were regressed onto the daily returns of the other using OLS.

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i \qquad (1)$$

The sample used in every case was half a year of daily returns. One year was approximated as having 250 trading days, half a year thus as having 125 trading days. Consequently, the first regression was made 125 days into the data, using the past 125 daily returns, and for every day after this one, the regression was run again using the last 125 days. This creates a rolling 125-day regression, where the oldest data point is dropped every time a new one is added[3].

---

[2] Although our choice of the minimum limit for the correlation coefficients is somewhat arbitrary, it is guided by two factors: (i) the coefficient should be as high as possible and (ii) several indices should be left for further examination.

[3] As the regressions were used to establish the relationships of the index pairs to enable relative volatility forecasts (see following section), an optimal sample size had to be found. Sample sizes of three months and twelve months were also tested and found to be unsatisfactory; Sample sizes of three months resulted in comparatively unstable relationships between index pairs. This could be explained by the presence of „seasonal“

The fact that the regressions are rolled day by day for a long period for every index pair results in a large number of regressions to be calculated. Figures 1-3 present the resulting linear index relationship over time (in terms of the regression coefficient $\hat{\beta}_2$).

The significance of the regression coefficients is tested with t-tests. Every single estimated slope coefficient $\hat{\beta}_2$ was found to be significantly different from zero at a 99 percent level of confidence. Because the results regarding the significance of the estimated slope coefficients are very similar across the several thousands of regressions, we do not present these results in detail.

The situation is different with the estimated intercepts $\hat{\beta}_1$. For the large majority, the hypothesis that they were not different from zero could not be rejected at a 95 percent level of confidence. In the rare cases where the estimated intercepts were found to be statistically significant, the recorded value of the intercept was very low (zero for practical purposes).

Boundaries for regression coefficients

Table 2 shows minimum and maximum values for the coefficients of determination for all regressions. They confirm a strong, although time-varying, linear relationship between the selected indices. Finding such a strong relationship was the motivation for selecting only index pairs with very high correlation.

However, the relationships of the examined index pairs are not constant over time. This simple observation has a profound impact on this study, as the relative volatility forecasts that will be attempted must be based on these relationships. Time-invariant relationships between index pairs with consistently high R^2 values would be ideal: it would then suffice to establish their interrelationship once in order to predict the respective relative volatilities for any future time interval. With both varying interrelationships and varying goodness of fit of the models, however, an alternative method had to be found to account for these instabilities.

---

relationships, which would prove to be unreliable in forecasting future relative developments in another „season". To illustrate this risk, imagine basing the relationship of an index-pair on the three past summer months, when comparatively little market activity takes place, to forecast the relationship during the historically more turbulent autumn months. In such a situation, a longer sampling period (bridging different potential „seasons") would ensure a smoother development of the relationship across time. On the other hand, relationships between index pairs do change structurally over time, and the longer the sampling period, the slower the reaction of the forecast to this change. The twelve-month sampling period proved to be too slow in adapting to such changes, and thus a six-month sample was used, long enough not to be seasonally biased, but short enough to remain somewhat flexible to structural changes.

The slope coefficients have already been estimated using the respective previous 125 trading days as a sample. These point estimators are subject to estimation error because coefficients have been found to vary across time. An upper and lower boundary should thus be considered to render the estimated slope coefficients more robust as a prediction tool. Applying the method of interval estimation would be inappropriate because ordinary interval estimation makes a statement about the confidence level with which the calculated interval will contain the true slope coefficient - the true slope coefficient is thus assumed to be constant. In our case, however, as we are dealing with rolling regressions, the degree of variance of the slope coefficients across time must also be considered when establishing boundaries for the estimators. Consequently, an empirical boundary calculation based on a simple minimum-maximum approach was chosen instead.

The boundaries are constructed by recording the largest variations of the relevant parameters during a time span that matches the forecasting horizon. This method reflects the market situation: in a volatile situation, when the index relationships vary strongly, the boundaries are wider. On the other hand, when the relationships are relatively constant, the boundaries are narrower.

As 25 trading days is the horizon of the different forecasting calculations, the largest percentage change of the estimated slope coefficients measured in any preceding 25-day interval during the preceding 250[4] trading days was recorded. These extreme changes of the beta coefficient were then used for establishing minimum and maximum boundaries at each point in time to ensure robust beta forecasts. The boundary estimation methodology is illustrated in *Figure 4*.

The following equations are thus used to calculate the boundaries of the estimated slope coefficients at time *t*:

lower boundary: $$\hat{\beta}_{2\,low\,(t)} = \hat{\beta}_{2(t)} - \max(\Delta\,\hat{\beta}_2)\,\hat{\beta}_{2(t)} \qquad (2)$$

higher boundary: $$\hat{\beta}_{2\,high\,(t)} = \hat{\beta}_{2(t)} + \max(\Delta\,\hat{\beta}_2)\,\hat{\beta}_{2(t)} \qquad (3)$$

[4] The choice of using 250 trading days as a sub-sample to establish such maximum variations is a rather delicate matter. The longer the sampling period, the more conservative the result, as a large variation will widen the boundaries for a long period of time. The boundaries are, however, designed to reflect current variation, and not the variation long past that is no longer reflected in the market situation. 250 trading days, covering a period of approximately one trading year, is chosen as an informal compromise between conservatism and recency (and relevance) of information.

where $\hat{\beta}_{2(t)}$ is the estimated slope coefficient at time *t* based on the sample t_{-124} to *t* and $\max(\Delta\hat{\beta}_2)$ is the largest percentage change of the estimated slope coefficients measured across any 25 trading-day period during the sub-samples (250 trading days) preceding the time of estimation.

For example, assume that, during the 250 trading days preceding time *t* of the ABC-XYZ index pair sample, the largest change of the estimated slope coefficient over any 25 trading-day period was *x*%. At time *t*, the estimated slope coefficient is $\beta_{2\tau}$. Thus, at that time, the lower boundary around this estimated slope coefficient would be β_{2t} - (β_{2t} *x*%) and the higher boundary would be β_{2t} + (β_{2t} *x*%).

The beta boundaries are included in *Figures 1-3*.

Volatility relationships between index pairs

For the purpose of this study, the relationships that have been established between the daily returns of index pairs need to be transformed into relationships between the respective volatilities. For the relationship between random variables X and Y such that $Y = a + bX + u$, we obtain for the variance $Var(Y) = Var(a + bX + u)$. By basic statistics, if *a* and *b* are constants and *X* and *u* are independent random variables, we have

$$Var(Y) = b^2 Var(X) + Var(u) \tag{4}$$

Applied to the regressions outputs, this expression can be written as

$$Var(R_{Yi}) = \hat{\beta}_2^2 Var(R_{Xi}) + Var(\hat{u}_i) \tag{5}$$

Based on this relationship, a forecast of the future interrelationship of the index pairs is attempted. From *Equation* (5) it can be seen that two inputs need to be estimated to forecast the variance of R_{Yi}: the slope coefficients $\hat{\beta}_2$ and the variance of the sample disturbance term $\hat{u}_i$.

The second parameter that must be estimated in *Equation* (5) is the variance of the residuals $\hat{u}_i$. An approach similar to the estimation of $\hat{\beta}_2$ was chosen for this estimation: both the smallest and the largest 25-day variance of the residuals during the 250 trading days preceding the respective points of estimation were recorded. Formally:

lower boundary: $$Var(\hat{u}_t)_{low} = \min(Var_{25\,days}(\hat{u}_i)) \tag{6}$$

higher boundary: $$Var(\hat{u}_t)_{high} = \max(Var_{25\,days}(\hat{u}_i)) \tag{7}$$

Now that all necessary parameters have been estimated, their forecasting power will be tested.

Out-of-sample test of volatility boundaries

To test the boundaries around $\hat{\beta}_2$ calculated previously, daily rolling forecasts of the variable to be explained are performed using these boundaries. In other words, we use the established relationships between the indices to forecast their relative future volatilities across a 25 trading-day period, and then compare this forecast to the actual recorded volatility levels. The statistical properties introduced in *Equations* (4) and (5) are applied. For example, to test the established SPX-OEX relationship at time *t*, the following equation is applied:

$$(\hat{\beta}_{2\,low(t)})^2 Var(OEX) + Var(\hat{u}_t)_{low} \leq Var(SPX) \leq (\hat{\beta}_{2\,high(t)})^2 Var(OEX) + Var(\hat{u}_t)_{high} \tag{8}$$

where *Var(OEX)* and *Var(SPX)* are the realized, future 25-day. If the realized SPX variance is within the forecasted boundaries, the forecasting is declared successful. This test is rolled daily, similarly to the regressions. The percentage of successful forecasts can be taken as an indication of the forecasting power of the established boundaries. The results of this out-of-sample test for the boundaries are shown in *Table 3*. Tolerance levels are used to widen the boundaries and thus make the forecast more robust. For a given tolerance Ψ, the forecast is successful if

$$(\hat{\beta}_{2\,low(t)})^2 Var(OEX) + Var(\hat{u}_t)_{low} - \Psi \leq Var(SPX) \leq (\hat{\beta}_{2\,high(t)})^2 Var(OEX) + Var(\hat{u}_t)_{high} + \Psi \tag{8a}$$

Clearly, the wider the boundaries, the higher the probability that the future volatility will fall between these boundaries. This fact is important when evaluating the power of such a forecasting test: the ideal would be a narrow boundary that includes all future volatility readings. However, these are antagonistic goals: the narrower the boundary, the smaller the probability, ceteris paribus, of including the future volatility. This is why tolerance levels have been included in the results shown in *Table 3*; as we are dealing with statistical arbitrage, the forecasting need not be perfect, but sufficiently close to allow for arbitrage trades to be triggered when a mispricing has been identified.

It is important to recall the interpretation of these figures. A result of 100% with no tolerance and a narrow boundary band would effectually indicate the existence of a quasi-deterministic relationship between the considered indices; knowing the volatility of one would then be enough to determine precisely and repeatedly the volatility of the other. That this is not the case is not surprising. However, with a 1% tolerance level, the percentage of successful forecasts is above 90% for all three index pairs, which, considering the relatively narrow boundaries (*Figures 1-3*) used, is considered a satisfactory degree of precision.

In the two previous sections we have developed a method that, based on the past relationship of daily returns, provides a robust forecast of the future relative volatility of three index pairs. These relationships, based on their lower and higher boundaries, are used below to identify relative mispricings in the options markets.

Arbitrage boundaries for relative implied volatilities

The next step is to apply the inferences made in the previous sections regarding relative future volatilities of index pairs to their relative implied volatility levels. It is important to stress again the fact that no relationship between the *magnitude* of the historical volatility and the implied volatility is postulated[5]. Rather, the relationship between the *relative* volatilities is addressed here, i.e., the relative difference in volatility levels, whether historical or implied. In other words, we do not establish a relationship between implied and realized volatility but between the volatility of options on two distinct indexes.

Therefore, once a precise relationship between the relative future volatility of index pairs has been established, then this relationship must also hold for the relative implied volatility of options on the two indexes. Thus, the boundaries calculated for the relative future volatility must also hold for the relative implied volatility, and *Equation* (8) must also apply to implied volatilities. Using the previous example of OEX and SPS options, this gives

$$(\hat{\beta}_{2\,low\,(t)})^2 Var_{impl}(OEX) + Var(\hat{u}_t)_{low} \leq Var_{impl}(SPX) \leq (\hat{\beta}_{2\,high\,(t)})^2 Var_{impl}(OEX) + Var(\hat{u}_t)_{high} \tag{9}$$

where $Var_{impl}(OEX)$ and $Var_{impl}(SPX)$ are the observed implied variances (squared implied volatilities) of at-the-money options with 25 trading days left to expiration. If such a boundary is violated, a theoretical mispricing is identified. The nature of this mispricing is not absolute; it is a relative mispricing, i.e., the relative difference of the implied volatilities of a particular index pair at a specific time differs significantly from the calculated relative future volatility difference between the two indices. The size of the theoretical mispricing is defined as the difference between the violated boundary and the observed implied volatility. The mispricings for all option combinations and for every trading day are presented in Figures 5-10.

However, due to the existence of transaction costs, a security margin similar to the tolerance level in *Table 3* is introduced to identify significant mispricings before an arbitrage trade is initiated. This security margin can be interpreted as a form of no-arbitrage barrier. Its magnitude is fixed at two times the respective bid-ask spreads of at-the-money options. Whenever such a significant mispricing is observed, a statistical arbitrage trade is simulated. In other words, it is not sufficient for the relative implied volatility to fall outside the bounds implied by the expression above. Additionally, the mispricing has to be of a certain minimum size before it is considered a statistical arbitrage opportunity. This additional rule underscores the conservative approach we take towards identifying statistical arbitrage opportunities. The no-arbitrage security margins are also shown in Figures 5-10.

Somewhat surprisingly, the number of theoretical mispricings is rather large. However, most mispricings fail to surpass the no-arbitrage barriers and are thus not identified as substantial enough to represent arbitrage opportunities.

It is also interesting to observe that the mispricings seem to occur in clusters: in certain periods, one index is persistently over/undervalued relative to the other regarding its implied volatility (see, for example, the relationship OEX-NYA). In such periods, the market seems to be making a persistent mistake by failing to recognize the correct relationship between the future volatilities of the respective indices. This persistent mispricing cannot, however, be eliminated by arbitrage as long as it stays inside the no-arbitrage barrier formed by the security margins.

We have not been able to identify factors that could cause the observed concentration of mispricings in phases. The possible explanation linking an increase of mispricings simply to an increase in market volatility, although intuitively appealing, proves to be unsatisfactory. First, mispricings (significant or not) can be observed both in periods of low and high volatility although they appear to be slightly more frequent in volatile markets. Second, linking the concentration of mispricings to volatility levels fails to account for the fact that the observed mispricings are persistent overvaluations of one index over the other in some phases and undervaluations in others. An unknown factor seems to influence the subjective relative

---

[5] See, for example, NCUBE (1994) for an analysis of such a relationship.

risk perception of market participants in such phases, causing a persistent misvaluation of the relative risk of the two indices.

Arbitrage Trading Strategy

To test whether the observed mispricings can be used for profitable arbitrage trading, an arbitrage strategy using at-the-money options is implemented. Option bid-ask prices are calculated from implied volatility data using a bid-ask spread (in terms of implied volatility percentage) around the closing-price implied volatility of at-the-money options. The bid-ask spread used is 1%[6], which for near-month, at-the-money options on liquid option markets is a conservative assumption. 35 calendar days to expiration and the respective one-month Eurodollar LIBOR are used. Because they are European-style and thus not exercisable prior to maturity, we use the Black-Scholes option pricing model for the SPX options. For the American-style OEX and NYA options, we use a binomial-tree model that incorporates the early-exercise feature.

When a mispricing is identified as being significant, one at-the-money option of the overvalued index is sold and the point-equivalent, $\hat{\beta}_2$ adjusted amount of at-the-money options of the undervalued index is bought. Point equivalency refers to the index points of the underlying indices.

Analytically (SPX-OEX example), if either

$$\left[(\hat{\beta}_{2\,low\,(t)})^2 \, Var_{impl}(OEX)_t + Var(\hat{u}_t)_{low} \right] - Var_{impl}(SPX)_t \geq \psi\% \qquad (10)$$

or

$$Var_{impl}(SPX)_t - \left[(\hat{\beta}_{2\,high\,(t)})^2 \, Var_{impl}(OEX)_t + Var(\hat{u}_t)_{high} \right] \geq \psi\% \qquad (11)$$

---

[6] Bid-ask spreads were recorded on a sample of days when realized volatility was relatively high and on days when realized volatility was relatively low so as to give a realistic reflection of such spreads in different market conditions. In relatively volatile markets, bid-ask spreads of approximately 1% were recorded for near-month, at-the-money options for all three indices, with slightly lower readings for the S&P 100 index options (OEX). In less volatile markets, these spreads tend to narrow, so that the choice to use 1% as an overall spread in the calculations can be regarded as a conservative assumption.

we open an arbitrage trade. If the first (second) condition is satisfied, a short (long) position of $\hat{\beta}_2 \cdot \frac{SPX_t}{OEX_t}$ at-the-money options on the OEX and a long (short) position of 1 at-the-money option on the SPX are entered into[7]. The notation is used as previously. SPX_t and OEX_t are the respective index levels. ψ is the security margin in volatility percentage.

Hence, whenever a significant mispricing is observed, one at-the-money option (put or call) on the relatively overvalued index is sold at the bid price (short position) and the point-equivalent, $\hat{\beta}_2$ -adjusted amount of at-the-money options on the undervalued index are bought at the ask price[8] (long position). The two positions are held without rebalancing for delta

---

[7] As the idea behind this offsetting hedge strategy is an approximate immunization of the entire position against all parameter changes except for the normalization of the observed implied volatility mispricing, it is crucial to take positions with equivalent "weight" and adjust them such that they reflect the respective relative volatility differential.

This procedure can be demonstrated by a simple example. Assume that

- index A has 100 points,
- index B has 1000 points,
- one at-the-money call on index A is worth \$8,
- one at-the-money call on index B is worth \$170,
- the relative volatility differential (the estimated slope coefficient $\hat{\beta}_2$) of index A to index B is 2 (for every percent A moves, B will move by two percent),
- index A and B are assumed to be perfectly correlated.

If the implied volatility on index B is overvalued relative to the implied volatility on index A (meaning that the relative implied volatility differential is higher than 2), the following position would be taken: To render the position point-equivalent, ten (1000/100) calls on index A need to be bought (as A was seen to be relatively undervalued; result: 10 x -\$8 = -\$80) to offset the sale of one index B call (+\$170). This point-equivalent position then has to be adjusted by the relative volatility differential ($\hat{\beta}_2$ - adjustment) to keep the offsetting hedge delta and gamma neutral. This is achieved by multiplying the long position by the relative volatility differential (by 2 in this example). A net long position of -\$160 and a net short position of +\$170 would be the result of this trade.

[8] The terms „undervalued" and „overvalued" are used here in reference to the relative implied volatility valuation of the indices, and obviously not to their valuation measured in index points.

neutrality until either (i) the mispricing disappears, i.e. the observed implied volatility of the indices are back inside the calculated boundaries, or (ii) the options expire before the mispricing disappears.

If the mispricing disappears before expiry, the position is unwound by buying back the previously overvalued option at the ask price and selling the previously undervalued amount of options at the bid price. If the mispricing persists until expiry of the options, the intrinsic values of the options are calculated and summed up. The total cashflows upon opening and closing the positions can then be added (interest on cashflows is neglected).

If the method used to define the interrelationship of indices and to forecast their relative volatility relationship is correct, the sum of the cash flows should be positive on average. There are two reasons why this statistical arbitrage strategy may not be profitable for every single trade: (i) the option positions are constructed to be delta and gamma neutral *under the assumption* that the underlying indices vary according to the statistical inferences made about their relationships. Whenever the index pairs fail to behave as predicted, the option position gains in delta exposure, which can result in a losing trade even though the observed implied volatility mispricing may have disappeared. (ii) there is the inherent risk in statistical arbitrage that an observed mispricing, however large, persists or even increases. This is a risk the statistical arbitrageur must always be aware of when deciding on the size of the arbitrage trade.

Results from arbitrage trading strategy

Table 4 shows the results of the statistical arbitrage strategy described in the previous section. Whereas only 10 SPX-OEX statistical arbitrage trades were triggered over the observed time period, trades involving NYA options occur more frequently. This could be due to the fact that the interrelationship between the SPX and the OEX indices is more obvious and thus more widely monitored by market participants, thus reducing the frequency of relative implied volatility mispricings. For every index pair, the average profit/loss per trade was positive. Although some trades generate losses, most are profitable.

One trade in particular needs to be mentioned because of its magnitude: as can be observed in Figures 5-6 regarding the magnitude of mispricings of the SPX-OEX index pair, a very significant mispricing was identified on October 27, 1997, when the SPX and the OEX lost 7.11% and 7.09%, respectively, followed by a bounce on the following day of 4.99% and 5.61%, respectively. Clearly, in such times the probability of a relative mispricing increases.

Bid-ask spreads would obviously be wider in such an occasion, but even if one were to double the spread on these days, a significant mispricing would still have been identified and the the ensuing trade would have remained profitable. Under the very restrictive assumption that no trading in options was possible at all during these two days, the average profit/loss for the SPX-OEX index pair would be greatly reduced with an average of \$0.61 achieved with eight instead of ten trades (on this particular day both call and put positions would be entered into). However, this value is still significantly positive and of a magnitude comparable to the other two index pairs.

Of the two potential risks involved with this relative arbitrage strategy mentioned earlier (persistence of mispricing and varying relationship of indices), the former surprisingly never materialized: the theoretical mispricings seem to be strongly mean reverting, where the mean is defined as the relationship between the relative historical volatility levels of the respective index pairs. In other words, there was no significant relative mispricing that lasted longer than 35 days. The other potential risk, however, illustrates the difficulty of making exact statistical inferences about the future relationships of stock indices. All the losing positions were caused by indices varying in a different manner than predicted with the linear relationship. These positions ended up with a loss even though the implied volatility mispricing disappeared. This is due to the fact that the option positions are taken in a manner that ensures they remain delta neutral as long as the indices vary relatively to one another according to their past relationship. If they cease to do so during the time the option position is open, a delta exposure arises which introduces a new risk parameter. However, because the chosen indices are strongly correlated, and because of the robust statistical inferences (rolling regressions with min-max boundaries), the impact of this risk is kept at a minimum.

Although amount and magnitude of mispricings tend to increase when the underlying stock indices are volatile, the majority of significant mispricings recorded were not linked to such extraordinary market conditions.

The results of *Table 4* do not account for transaction cost. However, even retail transaction costs are low compared to the arbitrage profits displayed in *Table 4*. Commissions are usually charged per traded contract, where one contract typically represents 100 options. Assuming rates of \$1.25 per contract[9] (\$0.0125 per option), this results in transaction costs of

[9] Rates charged by the U.S. discount broker AMERITRADE© as of March 2000.

approximately \$0.05 per round-trip arbitrage trade (4 x \$0.0125)[10]. It is interesting to note that the size of these transaction costs would have enabled even the retail investor to take advantage of the observed relative implied volatility mispricings. It is obvious that professional market participants, who benefit from lower transaction costs still, could profit to an even larger extent from such trades.

It might be argued that our simulated arbitrage strategy might not fully replicate real market conditions because the option prices calculated using the implied volatility data are not exact. For example, one could argue that, because the option prices used for simulated arbitrage trades were theoretical prices and not real market prices, simulated arbitrage opportunities might in fact not be real arbitrage opportunities. This is not the case because our analysis is based on implied volatilities computed from real market option prices. Therefore, by computing option prices from implied volatilities again, a simple reverse transformation is performed without a distorting effect on actual arbitrage opportunities.

However, there are other reasons why the statistical arbitrage opportunities detected might not, or not in all of the detected cases, be implementable. For example:

(i) The option prices were calculated using the closing prices of the indices, which, as HARVEY and WHALEY (1991) pointed out, is imprecise because option markets close 15 minutes after stock markets. Therefore, calculated option prices may be different from real option prices.

(ii) The implied volatilities used to calculate the bid and ask option prices were corrected by a bid-ask implied volatility spread. Although this spread was chosen so as to be conservatively large compared to normal market circumstances, it would be more exact to use the actual intra-day option bid and ask prices for the simulated trades. For example, it is conceivable that bid-ask spreads are occasionally much larger than usual. In this case, although an arbitrage opportunity would be detected, the arbitrage could not be implemented, or would at least be less profitable.

(iii) Whenever an arbitrage opportunity is detected, an arbitrage trade taking advantage of this opportunity has to go through at the next price. That price may be different from the initial price and may no longer allow arbitrage. Furthermore, an arbitrage trade consists of two legs. Often, these legs cannot be executed at the same time.

---

[10] The opening and closing of the offsetting trades involves the buying and selling of approximately 2 options (point-equivalency and $\hat{\beta}_2$ adjustments may somewhat increase or decrease this number.)

Conclusion

We illustrated and tested an aspect of option market efficiency that cannot be tested using conventional, exact arbitrage pricing relationships. Options on several pairs of highly integrated stock indices were used to establish the efficiency with which market participants value the relative risk (implied volatility) of these indices. Boundaries were calculated based on the historical covariance of the index-pair volatilities. Whenever the relative implied volatilities were found to violate such a boundary, a relative implied volatility mispricing was identified. Each one of the three investigated index pairs was found to have a large number of such mispricings, which seem to be concentrated in phases. However, after including bid-ask spreads in the calculations, only a small fraction of these mispricings was identified as an arbitrage opportunity. There is some indication that, occasionally, mispricings occur that can be used profitably by a statistical arbitrage trading rule. However, their number is small and there is the possibility that not all of the arbitrage opportunities detected could actually have been executed because of various arbitrage barriers in special market situations.

Further research might analyze why relative mispricings tend to appear in clusters. Possibly, common factors causing such mispricings could be identified. Additionally, other highly correlated markets, such as certain currencies or commodities, could be investigated with the methodology presented in this study.

Figures and Tables

Table 1. Correlation matrix. *Correlation coefficients based on weekly returns, March 31st, 1995 - December 3rd, 1999. Index pairs with a correlation coefficient higher than 0.95 (bold) are further analyzed in this study*

	SPX	***OEX***	***NDX***	***NYA***	***PTPX***	***INDU***	***SML***	***MID***	***XMI***	***RTY***
SPX	*1*	***0.988***	*0.780*	***0.988***	*0.798*	*0.922*	*0.709*	*0.804*	*0.91*	*0.729*
OEX		*1*	*0.770*	***0.968***	*0.799*	*0.933*	*0.662*	*0.759*	*0.923*	*0.684*
NDX			*1*	*0.721*	*0.563*	*0.633*	*0.688*	*0.745*	*0.593*	*0.708*
NYA				*1*	*0.785*	*0.933*	*0.757*	*0.846*	*0.919*	*0.772*
PTPX					*1*	*0.896*	*0.653*	*0.748*	*0.892*	*0.676*
INDU						*1*	*0.637*	*0.75*	***0.974***	*0.651*
SML							*1*	*0.922*	*0.583*	***0.983***
MID								*1*	*0.700*	*0.922*
XMI									*1*	*0.595*
RTY										*1*

Table 2. Minimum and maximum R^2. *For every series of index pair regressions, the lowest and highest recorded R^2 values are presented here.*

Index pairs	**minimum R^2**	**maximum R^2**
SPX-OEX	0.891	0.988
SPX-NYA	0.935	0.996
OEX-NYA	0.831	0.976

Table 3. Percentage of successful relative-volatility, out-of-sample forecasts using the established relationship boundaries.

Index pairs	**% of successful forecasts (with tolerance level)**			
	no tolerance	**.25% tolerance**	**.50% tolerance**	**1.0% tolerance**
SPX-OEX	68.1%	83.6%	90.8%	97.7%
SPX-NYA	60.6%	82.6%	92.4%	97.3%
OEX-NYA	68.0%	76.7%	84.5%	91.8%

Table 4. Profitability of Arbitrage Trades. *Results of the relative implied volatility arbitrage trades effectuated every time a mispricing was observed that was larger than 2%. Shown are the total number of positions opened per index pair, the number of which were profitable, the average profit in $ per profitable position (normalized to a position of one option of the first-named index in column 1), the number of losing positions, the average loss in $ per losing position, the average time a position was held open and finally, taking all positions into account, the average profit/loss in $ per arbitrage trade (again normalized to one option of the first-named index in column 1).*

	Total number of trades	**Numer of winning trades**	**Average profit ($)**	**Number of losing trades**	**Average loss ($)**	**Average time held**	**Total average profit/loss ($)**
SPX-OEX	10	9	3.62	1	-0.12	2.9 days	3.24
SPX-NYA	37	31	2.00	6	-1.57	9.6 days	1.42
OEX-NYA	28	22	1.07	6	-0.84	3.6 days	0.66

Figure 1. SPX-OEX beta boundaries. *Graph of obtained* $\hat{\beta}_2$ *values from rolling, 125 day regressions starting May 7*th*, 1991 and ending February 10*th*, 2000. The calculated lower and higher boundaries are also shown.*

$$R_{SPX\ i} = \hat{\beta}_1 + \hat{\beta}_2\ R_{OEX\ i} + \hat{u}_i$$

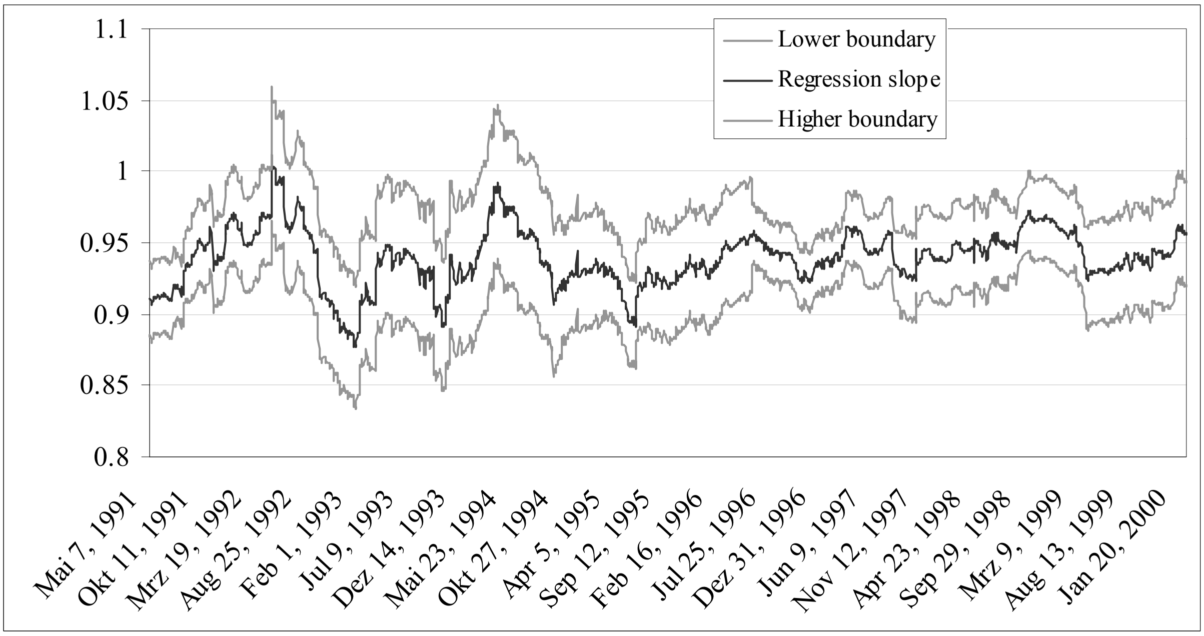


Figure 2. SPX-NYA beta boundaries. *Graph of obtained* $\hat{\beta}_2$ *values from rolling, 125 day regressions starting May 7th, 1991 and ending February 10th, 2000. The calculated lower and higher boundaries are also shown.*

$$R_{SPX\ i} = \hat{\beta}_1 + \hat{\beta}_2\, R_{NYA\ i} + \hat{u}_i$$

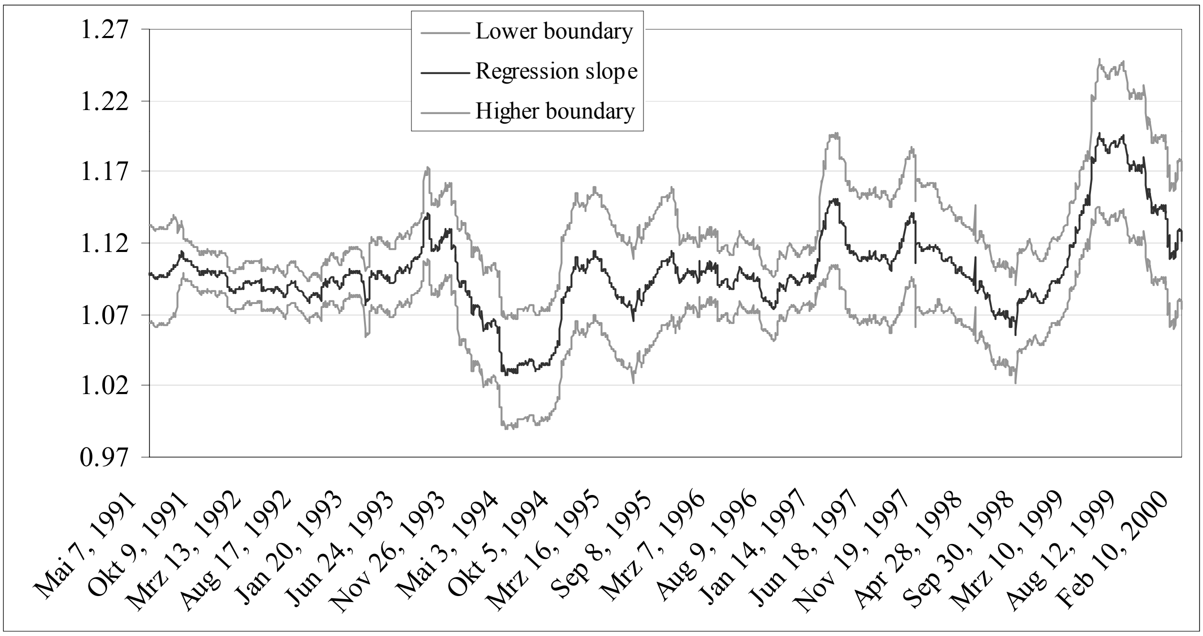


Figure 3. OEX-NYA beta boundaries. *Graph of obtained* $\hat{\beta}_2$ *values from rolling, 125 day regressions starting May 7*th*, 1991 and ending February 10*th*, 2000. The calculated lower and higher boundaries are also shown.*

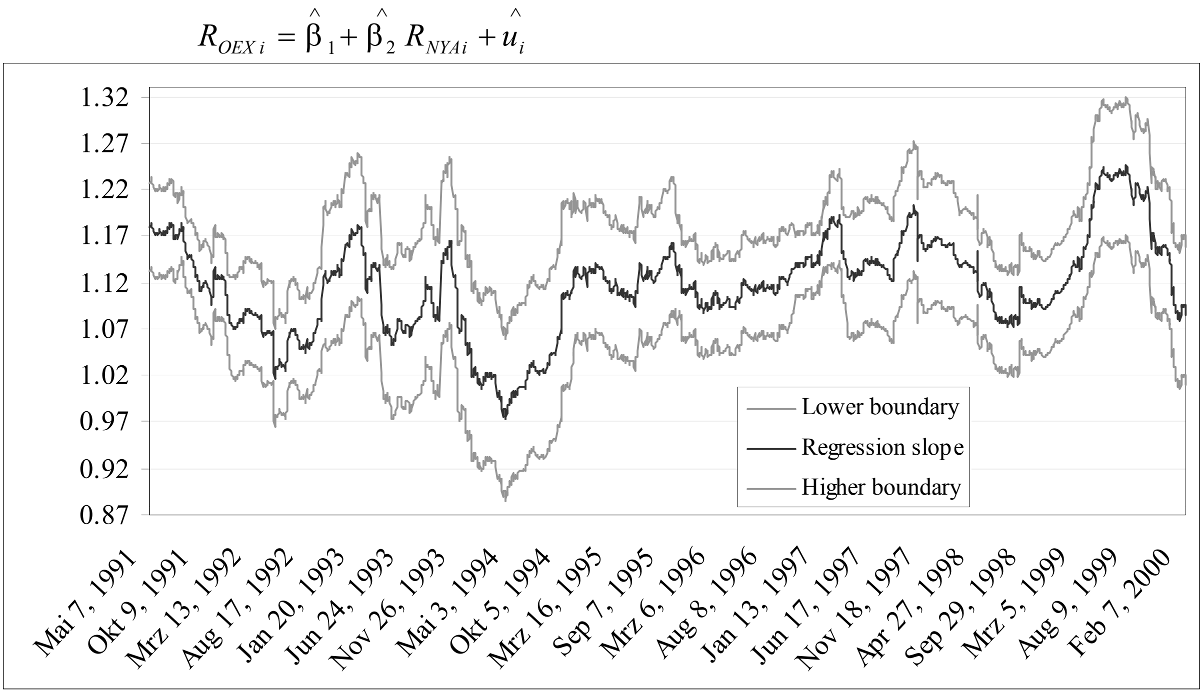


Figure 4. Beta boundary estimation methodology. *The largest change of beta in any preceding 25-day interval during the preceding 250 trading days is used to calculate the beta bounds. Because betas are estimated using 125 trading days of data (rolling 125-day windows), the first boundary estimation is available 375 days into the data.*

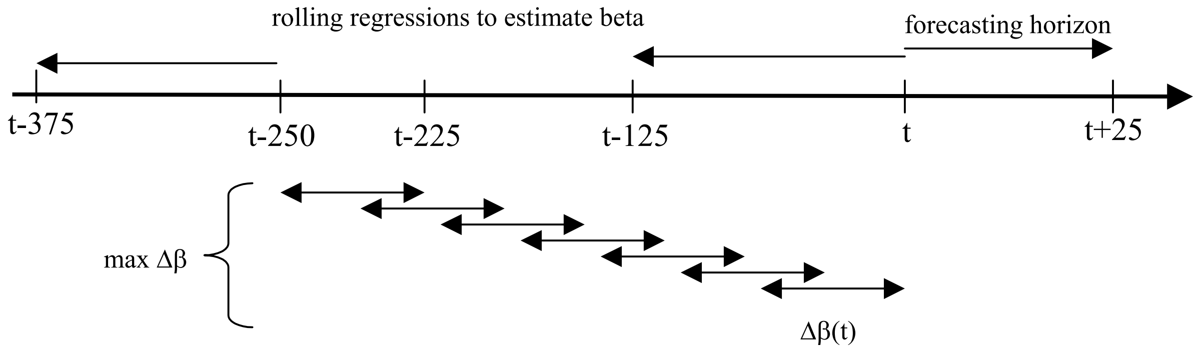


Figure 5. SPX-OEX call mispricings. *Theoretical mispricings of the relative implied volatilities of SPX-OEX calls for the period starting January 3rd, 1995 and finishing February 10th, 2000. The security margins have been included to determine the significant mispricings (gray lines).*

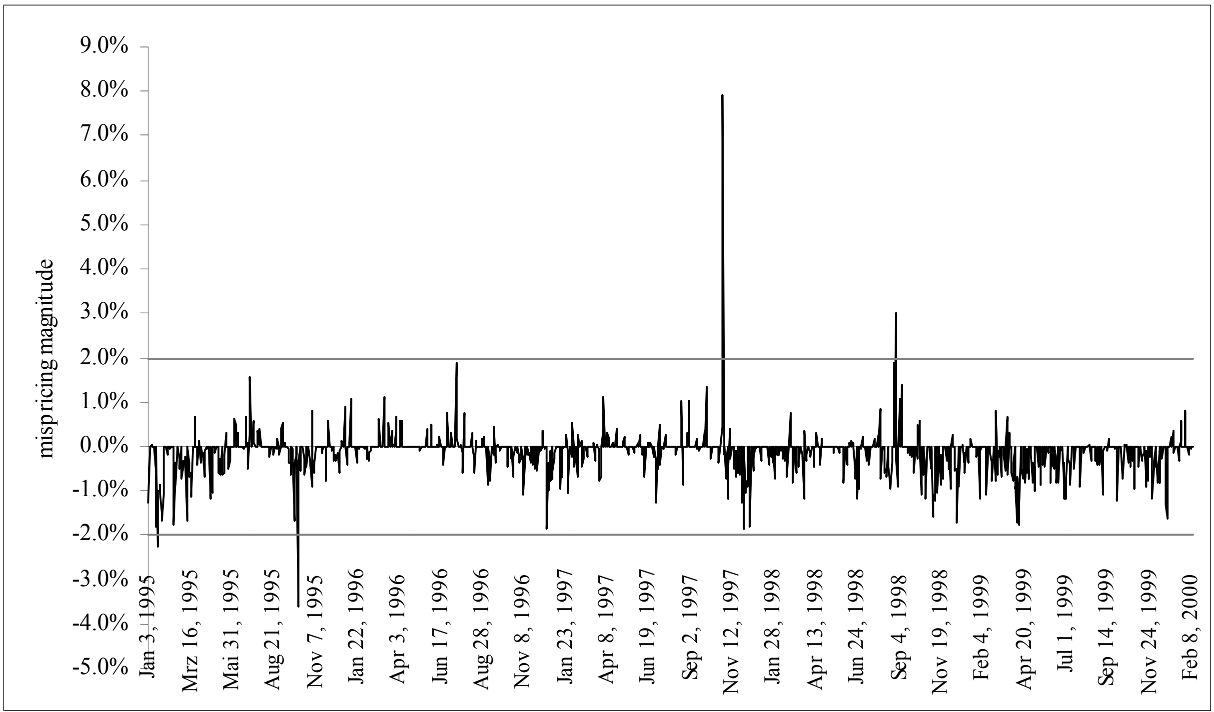


Figure 6. SPX-OEX put mispricings. *Theoretical mispricings of the relative implied volatilities of SPX-OEX puts for the period starting January 3rd, 1995 and finishing February 10th, 2000. The security margins have been included to determine the significant mispricings (gray lines).*

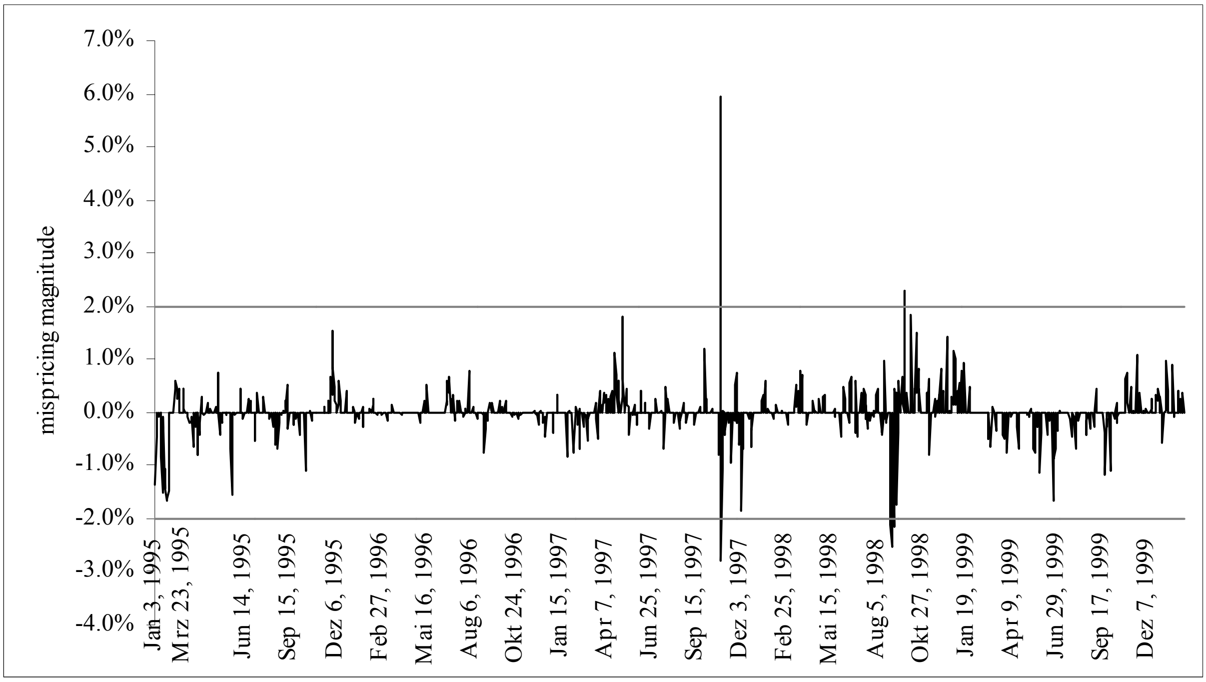


Figure 7. SPX-NYA call mispricings. *Theoretical mispricings of the relative implied volatilities of SPX-NYA calls for the period starting January 3rd, 1995 and finishing February 10th, 2000. The security margins have been included to determine the significant mispricings (gray lines).*

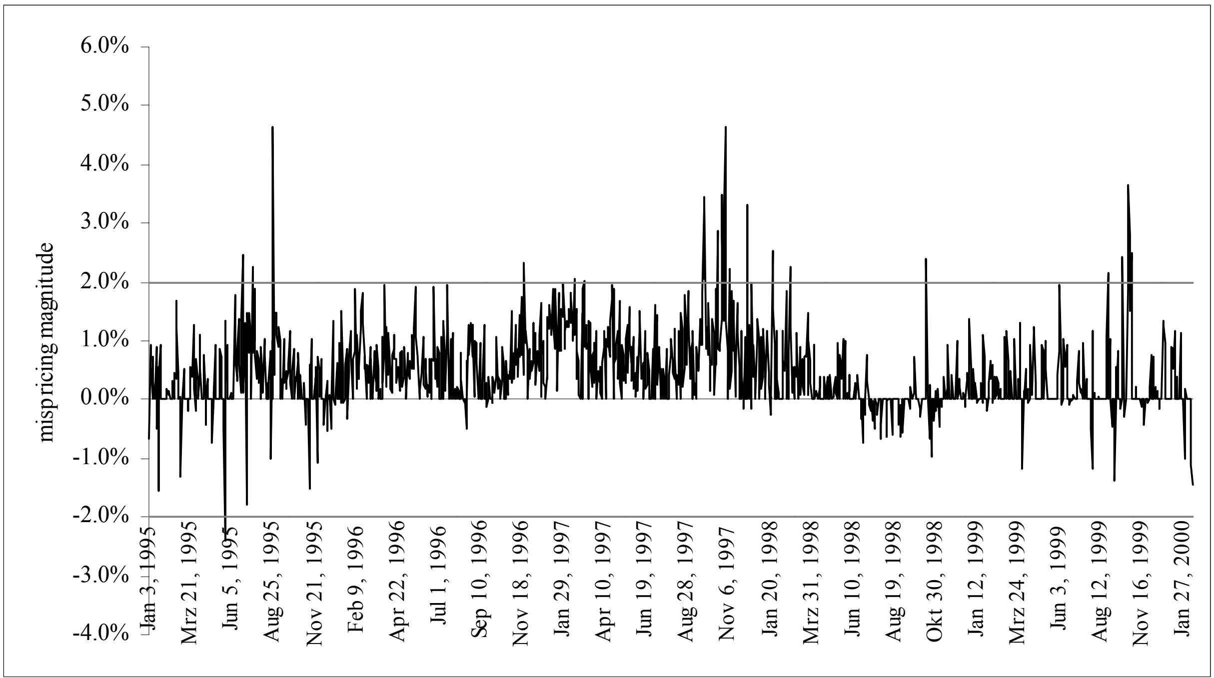


Figure 8. SPX-NYA put mispricings. *Theoretical mispricings of the relative implied volatilities of SPX-NYA puts for the period starting January 3rd, 1995 and finishing February 10th, 2000. The security margins have been included to determine the significant mispricings (gray lines).*

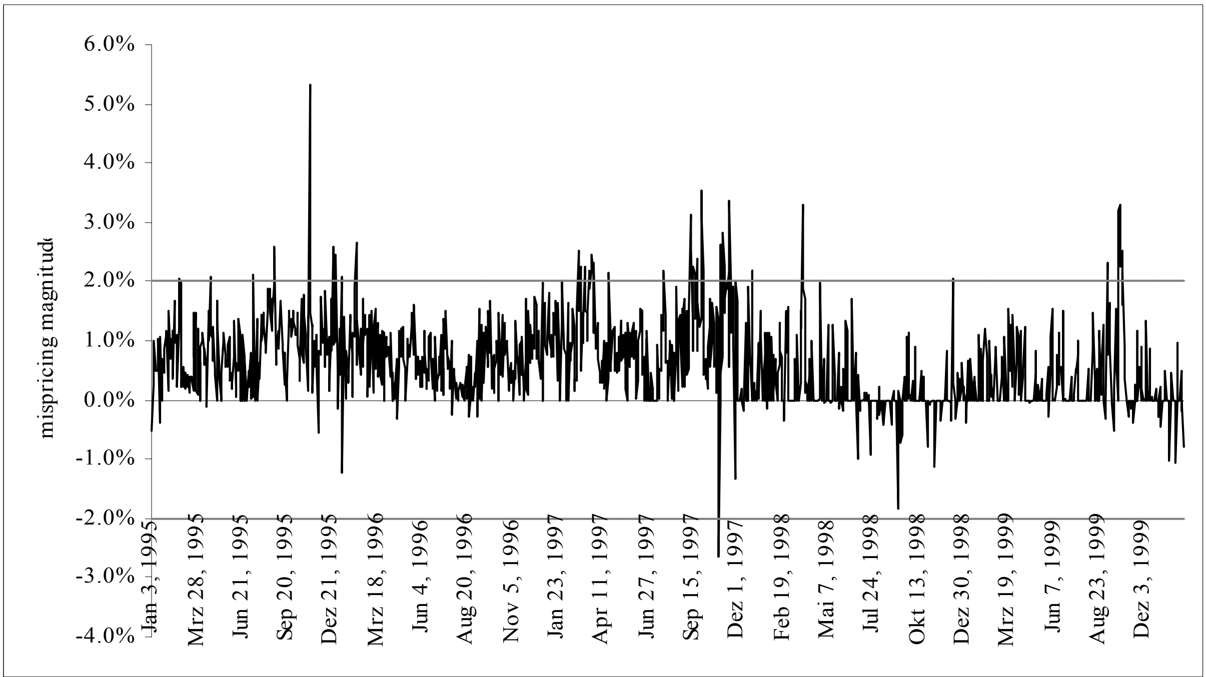


Figure 9. OEX-NYA call mispricings. *Theoretical mispricings of the relative implied volatilities of OEX-NYA calls for the period starting January 3rd, 1995 and finishing February 10th, 2000. The security margins have been included to determine the significant mispricings (gray lines).*

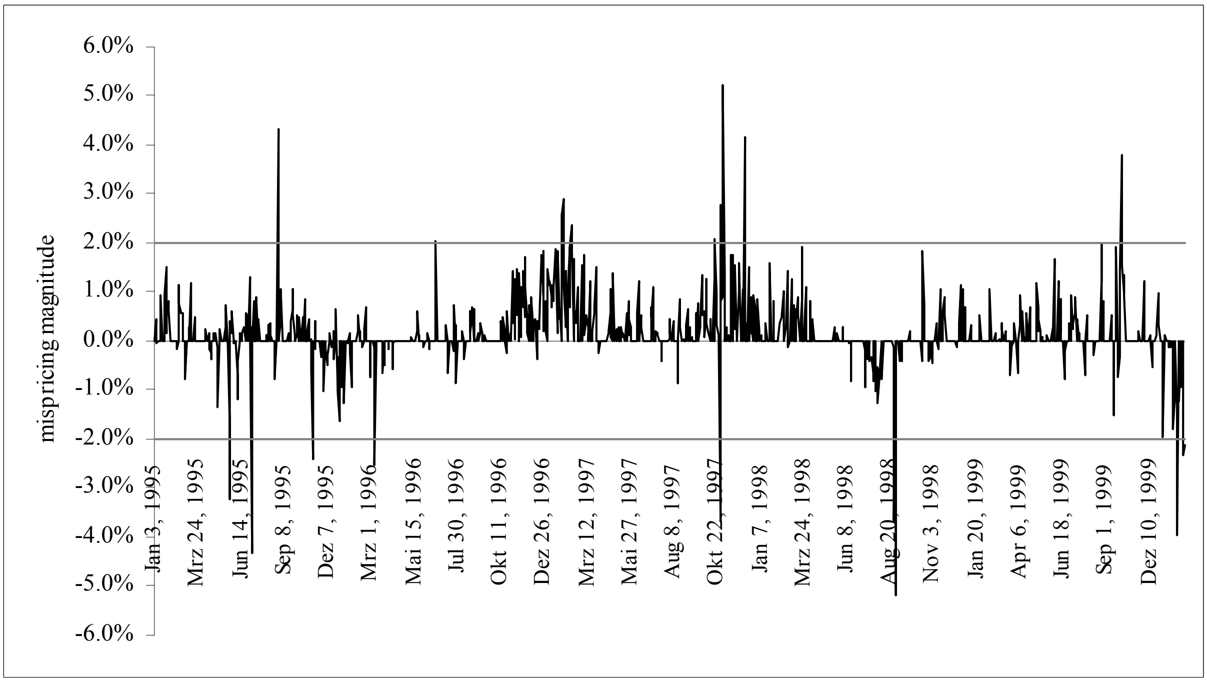
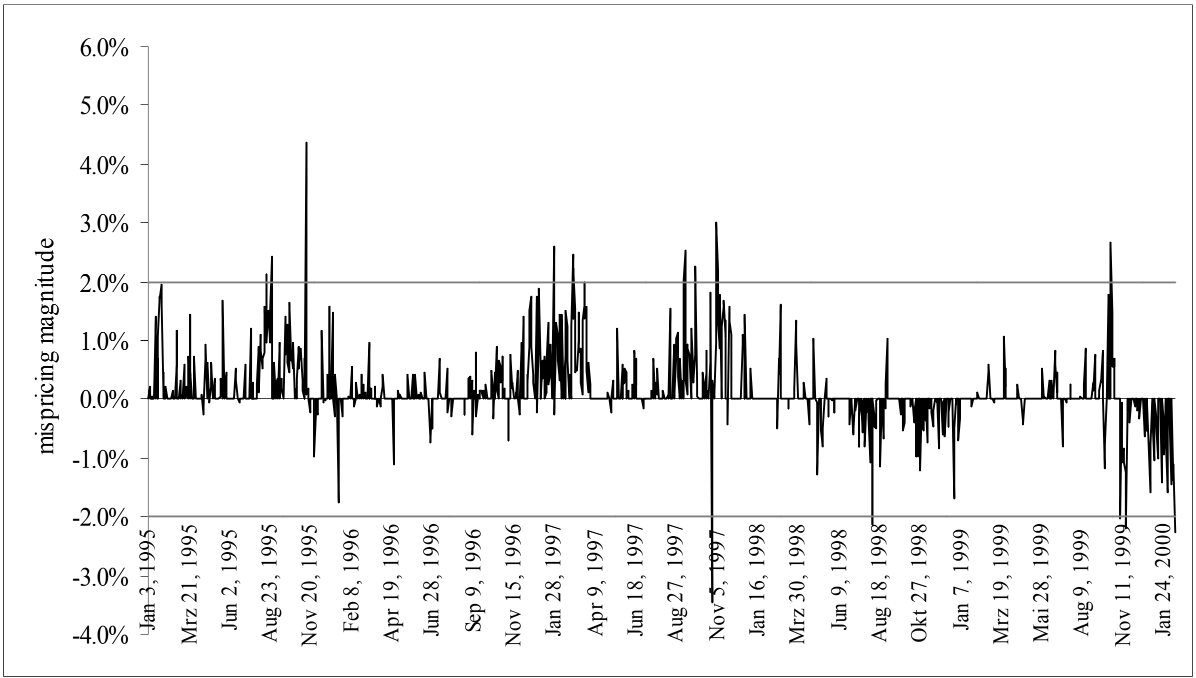


Figure 10. OEX-NYA put mispricings. *Theoretical mispricings of the relative implied volatilities of OEX-NYA puts for the period starting January 3rd, 1995 and finishing February 10th, 2000. The security margins have been included to determine the significant mispricings (gray lines).*



REFERENCES

ACKERT, L. and TIAN, Y. (1998): „The Introduction of Toronto Index Participation Units and Arbitrage Opportunities in the Toronto 35 Index Option Market"; Journal of Derivatives 5, 44-54

ACKERT, L. and TIAN, Y. (1999): „Efficiency in Index Options Markets and Trading in Stock Baskets"; Working Paper Series (Federal Reserve Bank of Atlanta), Vol. 99 Issue 5

CHUNG, P. (1991): „A Transactions Data Test of Stock Index Futures Market Efficiency and Index Arbitrage Profitability"; Journal of Finance 46, 1791-1809

FAMA, E. (1965): „The Behavior of Stock-Market Prices"; Journal of Business 38, 34-105

FIGLEWSKI, S. (1984): „Hedging Performance and the Basis Risk in Stock Index Futures"; Journal of Finance 39, 657-669

FIGLEWSKI, S. (1989): „Options Arbitrage in Imperfect Markets"; Journal of Finance 44, 1289-1311

GATEV, E., GOETZMANN, W. and ROUWENHORST, K. (1999): „Pairs Trading: Performance of a Relative Value Arbitrage Rule"; Working Paper 7032, National Bureau of Economic Research, Cambridge

GOULD, J. and GALAI, D. (1974): „Transactions Costs and the Relationship Between Put and Call Prices"; Journal of Financial Economics 1, 107-129

HARVEY, C. and WHALEY, R. (1991): „S&P 100 Index Option Volatility"; Journal of Finance 46, 1551-1561

JOHNSON, R., ZULAUF, C., IRWIN, S., and GERLOW, M. (1991): „The Soybean Complex Spread: An Examination of Market Efficiency from the Viewpoint of the Production Process"; Journal of Futures Markets 11, 25-37

KLEMKOSKY, R. and RESNICK, B. (1979): „Put-Call Parity and Market Efficiency"; Journal of Finance 34, 1141-1155

MERTON, R. (1973): „Theory of Rational Option Pricing"; Bell Journal of Economics and Management Science 4, 141-183

NCUBE, M. (1994): „Modelling Implied Volatility with OLS and Panel Data Models"; Discussion Paper No. 194, LSE Financial Markets Group Discussion Paper Series, London

NEAL, R. (1996): „Direct Tests of Index Arbitrage Models"; Journal of Financial and Quantitative Analysis 31, 541-562

POITRAS , G. (1997): „Turtles, Tails and Stereos: Arbitrage and the Design of Futures Spread Trading Strategies“; Journal of Derivatives 5.2, 71-88

SOFIANOS, G. (1993): „Index Arbitrage Profitability“; The Journal of Derivatives - Fall 1993, 6-20

STOLL, H. (1969): „The Relationship between Put and Call Options Prices“; Journal of Finance 24, 801-822