A Reexamination of Put–Call Parity on Index Futures

JOEL S. STERNBERG

Perhaps the single most controversial question in modern finance is whether markets are efficient. Prior to the 1980s a strong case (albeit indirect using low power tests) could be made for the existence of semi-strong equity market efficiency. However, with the advent of futures, options, and futures options on stock indices, a unique opportunity has been created to directly study efficiency in these markets.

Unlike most other cases of mispricing, e.g., index arbitrage, violations of put–call parity on futures can be exploited virtually without risk or large amounts of capital. Performing the arbitrage is also relatively easy because: only three instruments need be traded, all three are traded on the same exchange, and no dynamic rebalancing is necessary. Furthermore, testing put–call parity does not require a joint hypothesis of both a pricing model and market efficiency, since put–call parity is a mathematical identity.

In these many respects, the futures options traded at the Chicago Mercantile Exchange present the ideal opportunity to evaluate efficiency in a market from inception toward maturity. This market, therefore, warrants more thorough study than it has thus far received.

Preliminary evidence presented by Evnine and Rudd (1985) and Whaley (1986), in general, forms a credible case for rejecting efficiency in both the index options and futures options markets over the first year of trading. However, their results are highly dependent upon the quality

■ *Joel S. Sternberg is an Assistant Professor in the Department of Finance and Real Estate at the University of Arizona.*

The Journal of Futures Markets, Vol. 14, No. 1, 79–101 (1994)

CCC 0270-7314/94/010079-23

of the trading data supplied by the exchanges. In fact, Evnine and Rudd (1985) frankly admit that "it is inevitable that even real time prices direct from the exchanges will be somewhat non-synchronous" and grant that their "results suggest that in calm markets the errors induced by non-synchronous pricing may not be too severe." Unfortunately, it is in hectic markets that the tape is the most unreliable. And even to the extent that the tape represents a perfect record of real time trading, trades that appear even seconds apart may not have been simultaneously executable at the recorded prices during periods of high volatility.

In his survey article on options, Galai (1983) points out that the use of the trading record to test options market efficiency inevitably involves the joint hypothesis of data and trading synchronization, as well as efficiency. Errors due to data non-synchronization occur when two trades mistakenly appear to have taken place simultaneously because of incorrect time stamping. Errors due to trading non-sychronization occur when two correctly time stamped trades recorded only seconds apart nevertheless could not have been transacted simultaneously. The assumption that closely time stamped trades could have been simultaneously transacted at the listed prices may have indicated misvaluations which in fact were unexploitable.

This article attempts to overcome the problems of data and trading non-synchronization faced not only by the aforementioned authors, but by anyone studying mispricing in derivative securities markets. A data filtering technique is developed that minimizes susceptibility to these problems. As a result, a more reliable assessment of market efficiency can be made in testing the put–call parity equation.

The finding of mispricing ex-post, of course, does not assure the existence of exploitable arbitrage opportunities ex-ante, especially when transaction costs are considered. Nevertheless, this study clearly identifies misvaluations of large magnitude (even after transaction costs) and demonstrates a strong bias within the sample toward relatively overvalued futures and futures put options, and relatively undervalued futures call options. On the other hand, over the course of study there is a marked improvement in efficiency, as the number of potential opportunities for arbitrage diminishes significantly.

The article lays out the basic put–call parity relationship to be tested. It discusses the data filtering technique used to avoid problems of data and trading non-synchronization and addresses the complications presented by early exercise of futures options and its potential impact on arbitrage profits. An analysis of results is provided for the period

May–June, 1983[1] and the study is advanced through June, 1984, and tests a tax-modified put–call parity relationship. Finally, the appendix derives the tax-modified put–call parity equation, something which has been ignored heretofore in the literature.

THE PUT–CALL PARITY RELATIONSHIP

The technical aspects of the put–call parity relationship are discussed in much of the literature, including Cox and Rubinstein (1984). Adapting it to futures and futures options, $f = C - P + (K - F)e^{-rt}$, where f equals the cost of the forward contract now (or zero since the forward contract requires no initial cash outlay), P equals the cost of the put, C equals the cost of the call, F equals the market value of the forward contract, K equals the strike price of the call and put, and r is the interest rate on the t-bill whose maturity most closely corresponds to the expiration of the contracts. (If forming the hedge results in an outflow of cash, an interest rate 2% higher will be used to account for borrowing/lending spreads.) Since at maturity both sides of the equation will yield the same cash amount, they should require the same outlay. If this is not the case, then, hypothetically, an arbitrage opportunity exists which could be exploited by buying the relatively undervalued side of the equation and selling the other. The difference in value between the two sides of the equation, in turn, provides a measure of inefficiency in the market.

A problem with the put–call parity equation, as given, is that forward contracts are not traded. However, futures contracts, which are subject to daily marking-to-the-market, exist in abundance in U.S. financial markets. Jarrow and Oldfield (1981); Cox, Ingersoll, and Ross (1981); and Margrabe (1978) clarify some important differences between forward and futures contracts.

With the exception of French (1983) for the silver market, there is little empirical evidence of a significant difference in valuation between forward and futures contracts. [This is in contrast to French, Rendelman, and Carabini (1979), who demonstrated that the mathematical difference is insignificant.] Therefore, while the trading of futures contracts complicates the basic relationship, it does not render it invalid. A fairly strict pricing relationship should still exist between

[1]Simultaneous trading commenced on Friday, January 28, 1983, with the introduction of the S&P index futures options. S&P futures trading had already begun several months earlier.

these instruments. If it is violated, there is a mispricing of at least one of the instruments.

METHODOLOGY: THE DATA FILTERING TECHNIQUE

This study utilizes the time and sales record from the Chicago Mercantile Exchange for trading of the Standard and Poor's 500 futures and futures options contracts, and a record of the minute-by-minute values of the S&P index over the same periods. Using these data, the above mentioned equation was tested by combining those transactions which were recorded as taking place within a one-minute time frame on the time and sales record.

A test is performed to determine the degree of mispricing (if any) that would have been found by executing the theoretically equivalent positions of the put–call parity relationship under the assumption of costless trading. The form of the results is a printout of all possible trading combinations, along with the profits that would have been netted in each individual situation by trading *only* one of each contract at recorded prices, which would likely understate their potential magnitude. Henceforth, these profits are referred to as *hypothetical arbitrage profits* or *opportunities*.

Certain precautions must be taken, however, to ensure against errors due data and trading non-synchronization, since it is necessary to make certain that the trades comprising each combination could in fact have been made simultaneously to lock in the hypothetical arbitrage profit. The most obvious solution is to screen for those periods when the results would not be susceptible to these errors. This would be accomplished by selecting only those intervals from the time and sales record that are characterized by stability in the value of the underlying index. If only those intervals are selected, then neither inaccuracies in the time stamping nor time lapses between put, call, and futures trades would cause a problem since the underlying index upon which the put, call, and futures values are based has essentially remained the same.

Consequently, an important question is the appropriate manner in which to define stability. An obvious criterion for stability is variance of the underlying index. But variance alone may not be sufficient to screen for periods of stability. Two samples could have identical variance and yet one of these samples may demonstrate an upward trend of the index over time. Such a sample should be eliminated from the relevant data set since it may be masking autocorrelation in the underlying index (due to non-synchronous trading), while in theory the "true" value of the underlying asset is reflected in the prices of the calls, puts, and futures. The validity of any combination made during this interval would be vulnerable to trading non-synchronization through the matching within a single combination of put, call, and futures values, that are not based upon the same "true" underlying index value.

To account for this potential problem, a second criterion for stability is used. It puts a maximum value upon the range of values that can occur for the underlying index within any acceptable, selected interval. Therefore, for the purposes of this study, stability is defined as meeting both a range and variance criteria.

The maximum acceptable bounds for movements in the underlying index are a variance of 0.0002 and a range of 0.03. Given that the value of the underlying index during the May–June, 1983 period fluctuated between 160.00 and 170.00, the chosen criteria are, indeed, quite restrictive. (Even more stringent criteria are used, but do not appreciably affect the results.)

A tape with a minute-by-minute record of the value of the underlying index over the May–June, 1983 period is used to locate those periods of 10 minutes or more which meet these criteria, with the exception of the first 15 and last 40 minutes[2] of the record each day, which could reflect the effects of special order execution. For example, if minutes 16–25 for a particular day meet the dual criteria it is accepted and extended one minute at a time until it fails. If minutes 16–25 fail, the screening process resumes at minutes 17–26. This process continues until the end of the day is reached.

After these periods of stability are established, they are used to cut down the data set from the time and sales record of futures and options. To ensure that selected data actually falls within the specified interval (given the delay in time stamping), the first 20 seconds are chopped off if the first record is a quote and the first 90 seconds are chopped off if the first record is a sale. Once this is done, combinations are formed from futures, puts, and calls that share identical strike prices and expiration dates and occur within one minute on the record. Hypothetical arbitrage profits are calculated and are denoted by Arbitrage 1 in Table I.

EARLY EXERCISE

Arbitrage 1 is only intended to provide a rough estimate of the arbitrage profits that are netted in each case. Arbitrage 2 in Table I takes account

[2]The futures market at that time remained open 20 minutes later than the NY Stock Exchange.

TABLE I
May–June, 1983 (All Contracts)

	Arbitrage 1		*Arbitrage 2*	
	Sale	*Sale & Quote*	*Sale*	*Sale & Quote*
Sample size	193	256	193	256
Mean value	−31.9	−30.8	−31.8	−29.5
M-A-D	54.3	49.8	53.1	48.0
σ	62.7	57.4	61.7	56.3
T-statistic	7.06	8.59	7.15	8.39
T-probability	<0.01	<0.01	<0.01	<0.01
# Values > $70	67	79	63	72
# Negative values	122	166	122	166
Non-parametric Π	0.00015	<0.00001	0.00015	<0.00001

of the possibility that early exercise of options contracts sold by the arbitrageur may affect his profits. The early exercise of a sold option can impact the results because the equality expressed in the put–call parity equation is formulated on the basis of terminal values, i.e., the two sides of the equation provide the same cash outflow/inflow at maturity. In other words, if the options are not held to expiration, the equality will not hold precisely.

In general, an option will not be exercised very far in advance of expiration because the investor would sacrifice the premium over parity embodied in the option price. Specifically, in the case of a call option, this would be the amount by which the price of the call exceeds the difference between the price of the underlying asset and the strike price. There are times, however, when the premium over parity may fall to zero and lead to early exercise. In the case of equity call options, early exercise is feasible prior to ex-dividend dates, since holding the call past this point would entail a loss of a dividend payment.

Since futures contracts do not pay dividends, the situation here is somewhat more complicated. Still, early exercise may be advantageous. The important thing to remember here is that a futures contract is marked-to-the-market on a daily basis. Options on futures contracts, however, are not. Therefore, the only way an option holder can take advantage of the marked-to-the-market feature of futures contracts is by exercising his in-the-money option.[3] When (if) the option contract is exercised, a futures position will be opened that will immediately be

[3]This same reasoning would not apply to forward contracts since they are not marked-to-the-market. Therefore, an option on a forward contract will never be exercised early.

marked-to-the-market by the difference between the futures price and the strike price of his option. Brenner, Courtadon, and Subrahmanyam (1984) and Ramaswamy and Sundaresan (1984) demonstrate that early exercise will usually only become optimal very close to the expiration date of the option.[4]

It is impossible to know in advance whether an option contract will be exercised early. Nevertheless, a hypothetical arbitrage profit can be calculated for the following worst case scenario in which the option does get exercised against the arbitrageur. Assume that $C - P + (K - F)e^{-rt} > 0$. In other words, the futures is relatively undervalued and should be bought, while the arbitrageur should sell the call and buy the put. In addition, assume that the call which is being sold is in-the-money ($F > K$). Upon purchase, an investor could instantaneously exercise the option. (The intelligence of such a decision is another matter.)

The investor would receive the long side of a futures contract and the contract would immediately be marked-to-the-market. The difference between the futures and strike prices would be deposited in the investor's account and taken out of the arbitrageur's account (or whomever is assigned the call). Therefore the arbitrageur must forego interest on his posted margin.

The put–call parity hedge between futures and futures options is not destroyed by the exercise of the option that was written when initiating the hedge. Consequently, there is no need for the arbitrageur to liquidate his position because he already is holding the offsetting futures position. The exercisee merely relinquishes his option, leaving him exposed to the risk that the futures price may fall below the strike price of the option by expiration, a situation from which the arbitrageur could benefit.

For example, if the futures is undervalued, the arbitrageur would buy the futures and the put, and sell the call. If the call gets exercised against the arbitrageur, the short futures position that will be created in his account will be canceled by the long position in futures purchased when setting up the hedge. Therefore, the only risk for the arbitrageur arises from stochastic interest rates and their impact on the loss of interest income.

[4]Note that optimal early exercise is established by market consensus and would be reflected in the disappearance of the premium over parity. It is also important to reiterate that an individual investor, regardless of his own beliefs as to the merits of early exercise, will not do so if he can capture a premium over parity by merely selling the option.

The loss from early exercise can be closely approximated by computing interest on the amount that is marked-to-the-market at prevailing interest rates. This amount, $|F - K|(e^{rt} - 1)$, can be subtracted from the hypothetical arbitrage profit calculated for Arbitrage 1, where applicable, to yield Arbitrage 2. In effect, the result of this adjustment is to create bounds on the put–call parity relationship, in which[5]

$$-|F - K|(1 - e^{-rt}) \leq C - P + (K - F)e^{-rt} \leq |F - K|(1 - e^{-rt})$$

ANALYSIS OF ARBITRAGE VALUES FOR MAY–JUNE 1983

The results are broken up in several ways and analyzed using both t tests and a non-parametric test of probability where $E(\text{Profit}) = 0$. Since both quote (bid and ask) as well as actual trade values are used to form combinations, some of these combinations and their corresponding hypothetical arbitrage values are not valid. For example, if the value is positive, $C - P + (K - F)e^{-rt} > 0$, the correct strategy would be to sell the call, and buy the futures and the put. However, if the value of the call in the combination is actually an ask price, the arbitrageur/seller would actually receive a lower price. In this situation, an actual arbitrage opportunity may not exist since the hypothetical arbitrage profit may be no more than a reflection of the overstated call price.

To avoid this problem, all values which are in the expected direction (positive or negative), given the use of bid/ask values in the combinations, are eliminated from the analysis. A sample is then created from only those combinations which are composed solely from actual sale values, and are analyzed accordingly.

[5]Whaley, in *Futures and Options: Theory and Applications*, posits the put–call parity formula for futures options as:

$$-K(1 - e^{-rt}) \leq C - P + (K - F)e^{-rt} \leq F(1 - e^{-rt})$$

This formulation implies that if a short put is exercised against the arbitrageur, he must pay the strike price prematurely. And, if he exercises a long call, he receives premature payment of the futures price. As a result, the gain from early exercise of the long call and the resulting upper bound of put–call parity is posited as the difference between the present and future values of the futures price, while the loss from early exercise of the short put and the resulting lower bound is posited as the difference between the present and future values of the strike price. However, early exercise of the put or call does not result in the premature payment of either the futures or strike prices, but merely the opening of a long futures position and its immediate marking-to-market by the difference between the futures and strike prices. Theoretically, even if stock index futures had a delivery feature, payment of the futures and strike prices would still occur only at the delivery date, not when the option is exercised. Therefore, the above equation somewhat overstates the parity bounds.

A second sample is created which combines the first sample with those combinations that use bid/ask prices but have values in the wrong direction. Using the example from above, this would correspond to a situation where a positive mispricing is found, but the call value is a bid price.

Note that the use of bid/ask values to augment the sample allows certain improvements. The bid (ask) quote is the price that a floor broker offers to pay (accept) in exchange for the contract at that time. It, therefore, represents a price at which the hypothetical arbitrageur could surely execute the trade.

The use of bid/ask values also affords a larger sample size. (Unfortunately, a dearth of bid and ask quotes on the time and sales record makes their exclusive use impractical.) The large number of combinations partially formed with bid/ask values is seen in Tables I and IA by comparing the size of the "sale" sample to that of the "sale and quote" sample. Yet, inclusion of these additional combinations results in only a nominal reduction in the mean value.

Since the strict put–call parity equation represents a mathematical equivalence, hypothetical arbitrage profits for each combination should approximate zero. T tests indicate a significant, systematic bias toward negative hypothetical arbitrage values, i.e., *relatively* overvalued futures, undervalued calls, or overvalued puts, where $C - P + (K - F)e^{-rt} < 0$, with a mean value of about $-\$30$ and mean absolute deviations (M-A-D) of from \$50 to \$60 for most of the samples. However, the precise t-statistics may be affected by dependence of the values in the sample since the same futures contract may be used in several combinations with different puts and calls. This would result in an

TABLE IA

Data Set 1: May–June, 1983 (Near-Dated Contracts Only)

	Total Sample	
	Sale Only Records	*Sale & Quote Records*
Sample size	188	237
Mean value	−29.7	−27.9
σ	62.5	58.0
T-statistic	6.52	7.41
T-probability	<0.005	<0.005
# Negative values	117	149
Non-parametric Π	<0.0005	<0.0005
# Values > \$75	54	60
# Negative values	(51)	(57)

understatement of the true variance of the sample which would inflate the t-values. Still, the t-values are of such extraordinary magnitude that the negative bias of the values should be of little doubt.

Furthermore, non-parametric tests which directly compute the probability of discovering the same number or more negative values in a sample of similar size (given a null hypothesis of $E(\text{Profit}) = 0$ and the assumption of a binomial distribution) are similarly astonishing, although they too could be somewhat biased by dependence.

To eliminate problems of dependence, the average hypothetical arbitrage value within each interval of stability is computed. These average values are used then to generate a sample from which the t-statistics and non-parametric probabilities in Table IB are calculated. This highly conservative approach dramatically reduces the sample size and effectively complicates the task of demonstrating bias.

Nevertheless, the t test is significant at the 0.01 level while the non-parametric probability is 0.0147. Note that this conclusive overvaluation of futures contracts relative to the difference between call and put prices, $C - P$ (or undervaluation of $C - P$ relative to futures), contrasts sharply with the undervaluation of futures relative to the spot index that was observed by Figlewski (1983), among others.

The most obvious explanation for the negative bias found in the hypothetical arbitrage values involves the difference between forward and futures contracts. It is important to note that only an average 6¢ difference in valuation of a hypothetical futures contract over the corresponding forward contract would eliminate a $30 average profit (since each point on the S&P futures contract is valued at $500). Is there anything that would lead one to believe that a futures contract would be valued above a forward contract in this market? There is indeed a

TABLE IB
Data Set 1: May–June, 1983 (Near-Dated Contracts Only)

	Average Interval Values
Sample size	31
Mean value	−19.9
σ	41.3
T-statistic	2.69
T-probability	<0.01
# Negative values	22
Non-parametric Π	0.0147
# Values > $75	3
# Negative values	(3)

scenario under which it is possible. Margrabe (1978) demonstrated that forward contracts would trade at a premium over futures contracts in the event of stochastic interest rates in which the covariance of discount bond prices with spot prices exceeds the variance of bond prices.

Realistically, it is unlikely that there are any significant difference in valuation between the forward contracts and the futures contracts that are used in this study. Almost all of the combinations that comprise the samples are compiled from contracts with June, 1983 expirations. Therefore, the results are derived essentially from contracts that are close to expiration, when the difference in value between corresponding forward and futures contracts would be relatively small, if not totally insignificant.

Even if the 6¢ overvaluation could be successfully explained, it still would not account for the numerous, large hypothetical arbitrage values that fall to both sides of zero. The largest observed value for a single contract combination is negative 35¢ ($174), while about one-third of the values are in excess of $70, the approximate commissions at the time for executing the three trades in the combination through a discount broker. Differences of this magnitude could never be explained as resulting entirely from any difference between forward and futures contracts. This is most clearly borne out by the several large positive values (as big as $150) which, in contrast, involve an undervaluation of the futures contract relative to the value $(C - P)$.

Having demonstrated the presence of mispricing in the early stages of simultaneous trading of stock index futures and options, the same methodology developed here is adopted to evaluate the efficiency of these markets as they mature.

TESTING PUT–CALL PARITY MAY, 1983–JUNE, 1984 WITH TAX-MODIFIED BOUNDARY CONDITIONS: THE MOVEMENT TOWARD EFFICIENCY

With evidence of mispricing in the early stages of simultaneous trading of stock index futures and futures options, several periods are evaluated now under certain uniform conditions to observe their increasing efficiency over time.

Only early exercise adjusted values are reported, i.e., Arbitrage 2. Furthermore, to minimize the potential impact of early exercise and the futures versus forward contract dichotomy, only near-dated options within 1 1/2 months of expiration are used to test put–call parity on the March, June, September, and December futures from September, 1983 to June, 1984. The study ends with the June, 1984 contracts because the tax laws affecting futures options were changed effective July 17, 1984. Data beyond that date would not be comparable.

Prior to July 17, 1984, futures and futures options were subject to unequal tax treatment. Futures options were taxed at short-term rates if held less than one year and one day. In contrast, futures were (and still are) taxed as if any gain/loss was 60% long-term and 40% short-term, yielding a tax rate of $0.64T$, where T is the investor's marginal tax rate. As a result, there was an incentive for an investor holding a profitable position in futures options to exercise to open a futures position and receive the more favorable tax treatment that came with it. In contrast, an exercisee who had losses on his futures options position would receive less favorable tax treatment.

To more effectively deal with these potential tax effects, boundary conditions for put–call parity are constructed. As developed in the appendix, the modified put–call parity equation is:

$$-[0.36T/(1 - 0.64T)]P + X < C - P + (K - F)e^{-rt} < [0.36T/(1 - 0.64T)]C - Y$$

where

$$X = [TC - 0.64((1 - T)/(1 - 0.64T))P + T(K - F)e^{-rt}]/(e^{rt} + Te^{rt} + T)$$

and

$$Y = [0.64((1 - T)/(1 - 0.64T))C - TP + T(K - F)e^{-rt}]/(e^{rt} + Te^{rt} + T)$$

Results using this equation are presented in Tables IC–VC.

This section focuses now upon the magnitude of put–call parity violations. In doing so, an emphasis is placed on combinations that involve mispricing above a critical level. Since the multiplier on the S&P 500 futures and futures options is 500 and the tick (or minimum price change) is 5¢, a \$75 dollar level demonstrates a bias even if all prices in the combination are off by one tick from what the arbitrageur actually could attain, i.e., $(3 \times 0.05) \times 500 = \$75$. While it is possible that the actual executable prices may differ by even more than 5¢, it is important to remember that the data is screened for stability. In such periods, excessive price volatility is unlikely in the futures and futures options within the one-minute time frame used to make the combinations. In

TABLE IC
Data Set 1: May–June, 1983 (Near-Dated Contracts Only)

		Average Interval Values *Tax-Modified Sample*			
	Rate	15%	25%	35%	46%
Sale only records	Sample size	64	43	24	12
	# Negative values	42	29	19	7
	Non-parametric Π	0.0084	0.0158	0.0033	0.387
Sale & Quote records	Sample size	76	50	28	16
	# Negative values	52	36	23	11
	Non-parametric Π	0.0009	0.0013	0.0005	0.1050

fact, its observance would have no economic explanation and would indicate a mispricing between spot and futures markets.

Nevertheless, one may not want to regard any single combination as valid, even if it is over $75. Also, to the extent that dependence exists, many combinations in excess of $75 may be formed from the same contract. Therefore, a more conservative alternative, the average value for each interval, is calculated also with results appearing in Tables IB–VB.

In general, the presence of invalid trades in the sample should not affect the average of the interval, as some exert an upward bias while others exert a downward bias. In fact, the average value of the interval should be understated since bid/ask prices as well as sale prices are used in computing it. Therefore, while one may be unable to use the average

TABLE IIA
Data Set 2: August–September, 1983 (Near-Dated Contracts Only)

	Total Sample	
	Sale Only Records	*Sale & Quote Records*
Sample size	118	131
Mean value	−80.2	−74.1
σ	62.0	61.7
T-statistic	14.05	13.75
T-probability	<0.005	<0.005
# Negative values	103	115
Non-parametric Π	<0.000001	<0.000001
# Values > $75	58	60
# Negative values	(58)	(60)

TABLE IIB
Data Set 2: August–September, 1983 (Near-Dated Contracts Only)

	Average Interval Values
Sample size	16
Mean value	−47.5
σ	58.7
T-statistic	3.24
T-probability	<0.005
# Negative values	12
Non-parametric Π	0.0384
# Values > $75	4
# Negative values	(4)

interval value as a conclusive test of whether efficiency is achieved, the disappearance of average interval values over $75 can certainly be viewed as an indication of improved efficiency.

The individual results for each of the contracts and the average interval values are analyzed using t tests and non-parametric probability tests. Tax rates of 15%, 25%, 35%, and 46% are inserted into the equation. Use of the tax-modified put-call parity equation renders the nominal values of the results meaningless, so t tests are not performed on these data. For the non-parametric tests to detect directional bias in the sample, any values falling within the bounds of the tax-modified equation are considered to equal zero and subsequently omitted from the sample.

Noted earlier in Table IA, B, and C, the first sample period covering June, 1983 contracts from May 1 to the June 17 pre-expiration, reveal significant inefficiencies with an averaging of the intervals confirming

TABLE IIC
Data Set 2: August–September, 1983 (Near-Dated Contracts Only)

		Average Interval Values Tax-Modified Sample			
	Rate	15%	25%	35%	46%
Sale only records	Sample size	72	35	17	15
	# Negative values	72	35	17	15
	Non-parametric Π	<0.00001	<0.00001	<0.00001	<0.00001
Sale & Quote records	Sample size	78	41	20	16
	# Negative values	76	39	18	15
	Non-parametric Π	<0.00001	<0.00001	<0.00001	<0.00026

TABLE IIIA
Data Set 3: November–December, 1983 (Near-Dated Contracts Only)

	Total Sample	
	Sale Only Records	*Sale & Quote Records*
Sample size	42	74
Mean value	−3.6	3.3
σ	36.1	32.9
T-statistic	0.65	0.86
T-probability	<0.30	<0.20
# Negative values	18	27
Non-parametric Π	0.566	0.966
# Values > $75	3	4
# Negative values	(2)	(2)

these biases. Furthermore, three of the intervals during this period have average values of over $75. In Tables II–V A, B, and C, results for each of the successive contract periods over time are presented.

September futures and futures options trading for the period August 1 to September 16, 1983 presented in Tables IIA, B, and C, prove surprising. Far from showing any sign of the negative bias being eliminated and, thus, a movement toward efficiency, the September contract reveals an even larger negative bias. The t tests for both the sale sample and bid/ask sample are significant at the 0.005 level; whereas, the non-parametric probability tests yield even more dramatic results. For the sale only sample, 103 of the 116 combinations are negative while for the bid/ask augmented sample, 115 of 129 are negative. These numbers correspond to probabilities of less than 1 in 1 million. In addition, none of the tax rates used in the tax-modified arbitrage equation eliminated the negative bias.

TABLE IIIB
Data Set 3: November–December, 1983 (Near-Dated Contracts Only)

	Average Interval Values
Sample size	23
Mean value	0.85
σ	26.3
T-statistic	0.14
T-probability	NM
# Negative values	8
Non-parametric Π	0.76
# Values > $75	1
# Negative values	(4)

TABLE IIIC
Data Set 3: November–December, 1983 (Near-Dated Contracts Only)

		Average Interval Values Tax-Modified Sample			
	Rate	15%	25%	35%	46%
Sale only records	Sample size	17	6	0	0
	# Negative values	10	5	0	0
	Non-parametric Π	0.315	0.109	NM	NM
Sale & Quote records	Sample size	23	6	0	0
	# Negative values	14	5	0	0
	Non-parametric Π	0.315	0.109	NM	NM

Evaluation of the data by intervals also reveals an enlargement of bias. The grand mean of the average interval values is −$47.52, which is statistically significant at the 0.005 level. There are also four intervals with average values over $75, and all are negative. While increasing bias seems to deny a movement toward market efficiency (at least in these still early months), the apparent increase in mispricing may reflect a magnification of the previous mispricing due to increased volume in these markets.

In general, the data in Tables IIIA, B, and C for the December contracts drawn from November 1 to December 16, do not demonstrate any identifiable directional bias in observed values. In fact, whereas the sale only sample reveals a negative average value, the augmented bid/ask sample has a positive average value. In both cases, the average values are

TABLE IVA
Data Set 4: February–March, 1984 (Near-Dated Contracts Only)

	Total Sample	
	Sale Only Records	*Sale & Quote Records*
Sample size	189	271
Mean value	−5.0	−10.3
σ	59.9	54.1
T-statistic	1.15	3.13
T-probability	<0.20	<0.005
# Negative values	113	175
Non-parametric Π	0.0043	0.0005
# Values > $75	49	53
# Negative values	(26)	(30)

TABLE IVB
Data Set 4: February–March, 1984 (Near-Dated Contracts Only)

	Average Interval Values
Sample size	29
Mean value	3.7
σ	40.6
T-statistic	0.50
T-probability	<0.40
# Negative values	16
Non-parametric Π	0.3555
# Values > $75	2
# Negative values	(0)

quite low (−\$3.62 and \$3.29, respectively), which could be within the range of transaction costs paid by traders on the exchange floor. On the other hand, the non-parametric probability of observing the 41+ positive values out of the sample of 68 is only 0.057. While not significant at the 0.05 level, it gives the first indication thus far that, while the bias could be diminishing, it may be possible also that it is just becoming positive.

In testing the tax-modified arbitrage equation non-parametrically, none of the categories yield anything close to a statistically significant result. Any statistically insignificant bias is once again in the negative direction. The average interval values are also without significant bias. Here, the grand mean of the average interval values is a scant \$0.85 and neither t tests nor non-parametric tests are significant. Nevertheless, while there is no convincing evidence of bias, one value is greater

TABLE IVC
Data Set 4: February–March, 1984 (Near-Dated Contracts Only)

	Average Interval Values Tax-Modified Sample				
	Rate	15%	25%	35%	46%
	Sample size	82	46	24	12
Sale only records	# Negative values	43	30	11	11
	Non-parametric Π	0.37	0.849	0.729	0.0032
	Sample size	119	72	37	26
Sale & Quote records	# Negative values	77	43	24	25
	Non-parametric Π	0.0009	0.0625	0.0494	<0.00001

TABLE VA
Data Set 5: May–June, 1984 (Near-Dated Contracts Only)

	Total Sample	
	Sale Only Records	*Sale & Quote Records*
Sample size	222	363
Mean value	−7.3	−2.6
σ	35.0	32.7
T-statistic	3.31	1.51
T-probability	<0.005	<0.10
# Negative values	126	186
Non-parametric Π	0.0488	0.3832
# Values > \$75	12	15
# Negative values	(10)	(12)

than \$75. This observation along with the sample's relatively small size, however, gives reason to counsel caution in interpreting these results.

Fortunately, the fourth data set in Tables IVA, B, and C, covering February 1 to March 16 consists of a sample size that is about four times larger than that of data set 3. Once again, indications of a negative bias are observable. The sale only sample is inconclusive. It has a mean value of −\$5.00, but a very high standard deviation prevents it from being significant at the 0.05 level. However, the bid/ask augmented sample with a mean of −\$10.28 is statistically significant at the 0.01 level. This same pattern occurs in the non-parametric probability tests. For both the standard put–call parity equation and the tax-modified equation, the bid/ask augmented sample is statistically significant (with the lone exception of the $t = 25\%$ case where the probability is 0.0625), while the sale only sample is only significant for the no tax case and the 46% tax

TABLE VB
Data Set 5: May–June, 1984 (Near-Dated Contracts Only)

	Average Interval Values
Sample size	43
Mean value	−1.47
σ	25.7
T-statistic	0.36
T-probability	<0.40
# Negative values	21
Non-parametric Π	0.62
# Values > \$75	1
# Negative values	(1)

TABLE VC
Data Set 5: May–June, 1984 (Near-Dated Contracts Only)

Average Interval Values
Tax-Modified Sample

	Rate	15%	25%	35%	46%
	Sample size	49	16	13	6
Sale only records	# Negative values	22	7	6	3
	Non-parametric Π	0.804	0.773	0.709	0.656
	Sample size	56	16	13	6
Sale & Quote records	# Negative values	25	7	6	3
	Non-parametric Π	0.825	0.773	0.709	0.656

case. Note that introducing higher tax rates causes the bias to disappear but then reappear for the highest tax rate. This is further evidence that taxes do not seem to be responsible for the observed bias.

Some interesting observations can be garnered from the average interval values. The mean value of these values is actually positive, although none are statistically significant. In contrast, 16 of the 29 values are negative, although the probability of this happening is 0.3555. Evidence of bias is therefore much less clear. While there may be indications of a disappearance of bias, there are still two positive average interval values over \$75. This is not entirely surprising since the mean is positive and a majority of the values are negative.

In Tables VA, B, and C, the fifth data set covering May 1 to June 15, 1984, provides the best evidence of the elimination of inefficiency. The only evidence at all of bias is in the sale only sample of the no-tax equation. The mean of this sample is −\$7.77 which is significant at the 0.01 level in t tests and the 0.05 level in non-parametric tests. However, when the bid/ask combinations were included in the sample, the sample size the mean falls dramatically to −\$2.59 (and is not statistically significant under either test). None of the tests of the tax-modified equation are even close to significance for either the sale only or bid/ask augmented samples.

The fifth data set is characterized also by the near total disappearance of average values in excess of \$75. Although it is easily the largest sample tested both in terms of the total number of combinations and the number of intervals, there is only one average interval value in excess of \$75. Viewed alongside the results of the third and fourth data sets, it establishes a pattern of increasing efficiency in the futures and futures options markets.

SUMMARY

The reliability of efficiency tests in derivative securities markets rests heavily upon the quality of the trading record. This study develops and deploys a technique to filter the trading record to minimize the possibility of errors due to trading and data non-synchronization.

The results document the existence of market inefficiencies under the most rigorous conditions. In spite of the extremely conservative approach taken in screening for periods of stability, significant inefficiencies of two types are found. Foremost, a large number of violations of put–call parity are found that neither fall within the realm of transaction costs nor can be explained away by taxes. In addition, the results from most of the first year of trading reveals a fairly persistent bias toward relatively overvalued futures, undervalued calls, or overvalued puts.

On the other hand, a definitive improvement over the life of the study is apparent. Although the final period does not appear to signal a complete elimination of mispricing in these markets, it does demonstrate a diminished number of large hypothetical arbitrage values, as well as significantly less bias. It appears that by June, 1984 these markets are clearly achieving a heightened level of efficiency.

APPENDIX

Prior to July, 1984 and the enactment of the Deficit Reduction Act, futures and futures options were subject to different tax treatment. All futures contracts regardless of their holding period were taxed as if 60% of the gain (loss) was long-term and 40% short-term. On the other hand, taxation of the futures options, like other investments, depended entirely on whether their holding period was greater or less than one year and one day. Since very few options (and none used in this study) have maturities of over one year, futures options were effectively treated as short-term investments subject to full taxation. However, by exercising the option, thus converting it into a futures position, the futures option holder could have converted a short-term gain to a more favorable 60/40 gain. Conversely, the writer of the option could have had a short-term loss converted to a less favorable 60/40 loss.

The effective tax rate on futures contracts is $0.64T$, where T is the marginal tax bracket of the investor. Of the 60% profit/loss that is treated as long-term, only 40% of this amount is taxable, while the 40% that was treated as short-term was fully taxable. After taxes, the investor would have been left with a net profit of (GROSS PROFIT) $\times$ $(1 - 0.64T)$.

Since futures contracts do not require any initial outlay, taxes will never affect their pricing. This is because any investor could counteract the tax burden by purchasing $1/(1 - 0.64T)$ futures contracts at the outset leaving him the same after-tax profit as in a taxless world. As a result, no changes need to be made to the futures contract in the put–call parity relationship to account for any adjustment in the futures position.

Since net profits after taxes on short-term gains are (GROSS PROFIT) $\times$ $(1 - T)$, an investor buying $(1 - T)/(1 - 0.64T)$ futures contracts would be duplicating the payoff from a futures contract taxed entirely at the short-term rate. In other words, to convert the after-tax profit on an investment with 60/40 tax treatment to the after-tax profit of that same investment under short-term tax treatment, the investor would buy $(1 - T)/(1 - 0.64T)$ contracts.

Taxes may affect the pricing of futures options. Other than to accumulate interest on the amount credited when the underlying futures contract is marked-to-the-market, there is no economic reason to exercise a futures option early since it would entail a loss of the premium over parity. However, the reduction in the tax rate from early exercise may have resulted in a higher after-tax profit.

Since 100% of the profits from the futures option is taxable versus only 64% from the futures, the net difference in the tax burden would have been 0.36T $\times$ GROSS PROFIT, where gross profit on a long position equals $(F - K - C)$ for calls, and $(K - F - P)$ for puts and vice versa for short positions. It follows that an investor may rationally exercise his option if (0.36T $\times$ GROSS PROFIT) $>$ $(1 - T) \times$ (PREMIUM OVER PARITY), where the decision to exercise is a function of the tax bracket of the individual investor.

To account for the possibility of exercise, the put–call parity equation must be modified as the after-tax gain will be greater for the exerciser and the after-tax loss greater for the exercisee. (Of course, it is possible that the person assigned as the exercisee could have a profit on the position as well, since the two investors may not have opened their positions at the same time/price.) The arbitrageur could account for the possibility of exercise by transacting a lesser number of contracts that would result in the same payoff as one option contract taxed at the short-term rate. As discussed with regard to futures, this number is $(1 - T)/(1 - 0.64T)$.

In the event of exercise of the call, selling only $(1 - T)/(1 - 0.64T)$ call options, buying one put and borrowing the amount $(K - F)e^{-rt}$, while buying $(1 - .T)/(1 - 0.64T)$ futures contracts, will create a position which is unaffected by the movement of the futures contract. However, the arbitrageur will be left with an end of period flow of,

$$0.64T[(1 - T)/(1 - 0.64T)]C - TP + T(K - F)e^{-rt}$$

arising from the effect of taxes on the net amounts paid for the call and received on the put, plus interest accumulation on the borrowing position (assuming that taxes are paid at the expiration of the contract). By increasing the initial borrowing position by the amount,

$$[0.64T((1 - T)/(1 - 0.64T))C - TP + T(K - F)e^{-rt}]/(e^{rt} - Te^{rt} + T)$$

final flows will net to zero. Therefore, given the possibility of early exercise of the call, but not the put, the following bound can be established:

$$((1 - T)/(1 - 0.64T))C - P + (K - F)e^{-rt} + X < 0$$

where,

$$X = [0.64T((1 - T)/(1 - 0.64T))C - TP + T(K - F)e^{-rt}]/(e^{rt} - Te^{rt} + T)$$

Subtracting $C - P + (K - F)e^{-rt}$ from both sides the following expression is obtained:

$$C - P + (K - F)e - rt < [0.36T/(1 - 0.64T)]C + X$$

In the event of exercise of the put, having sold only $[(1 - T)/(1 - 0.64T)]$ puts, while buying one call, selling $[(1 - T)/(1 - 0.64T)]$ futures contracts, and lending the amount $(K - F)e^{-rt} + Y$, where

$$Y = [TC - 0.64T[(1 - T)/(1 - 0.64T)]P + T(K - F)e^{-rt}]/(e^{rt} - Te^{rt} + T)$$

will provide a net flow of zero at expiration. Therefore, given the possibility of exercise of the put, but not the call, a second bound can be established:

$$0 < C - [(1 - T)/(1 - 0.64T)]P + (K - F)e^{-rt} + Y$$

Subtracting $C - P + (K - F)e^{-rt}$ from both sides, the following expression is obtained:

$$-[0.36T/(1 - 0.64T)]P + Y < C - P + (K - F)e^{-rt}$$

Of course, the exercise of either option cannot be ruled out. By combining expressions for both exercise of the call and put, a single, robust relationship is obtained which must hold to preclude arbitrage,

$$-[0.36T/(1 - 0.64T)]P + Y < CP + (K - F)e^{-rt} < [0.36T/(1 - 0.64T)]C + X$$

It is this expression that is tested in the article.

BIBLIOGRAPHY

Brenner, M., Courtadon, G., and Subrahmanyam, M. (1984): "Options on Stock Indices and on Stock Index Futures," Working Paper, New York University.

Cox, J., Ingersoll, J., and Ross, S. (1981): "The Relationship between Forward Prices and Futures Prices," *Journal of Financial Economics,* 9:321–346.

Cox, J., and Rubinstein, M. (1984): *Options Markets.* Englewood Cliffs, NJ, Prentice-Hall.

Evnine, J., and Rudd, A. (1985): Index options—the early evidence, *Journal of Finance,* 3:743–756.

Figlewski, S. (1983): "Why are Prices for Stock Index Futures So Low?," Working Paper, Institute of Business and Economic Research, University of California, Berkeley.

French, K. (1983): "A Comparison of Forward and Futures Prices," *Journal of Financial Economics,* 12:311–342.

Galai, D. (1983): "A Survey of Empirical Tests of Option-Pricing Models," in Brenner, M. (ed.), *Option Pricing: Theory and Applications,* Lexington, MA, Lexington Books.

Jarrow, R., and Oldfield, G. (1981): "Forward Contracts and Futures Contracts," *Journal of Financial Economics,* 9:373–382.

Margrabe, W. (1978): "A Theory of Forward and Futures Prices," Unpublished.

Ramaswamy, K., and Sundaresan, S. (1984): "The Valuation of Options on Futures Contracts," *Journal of Finance,* 40:1319–1340.

Rendelman, R., and Carabini, C. (1979): "The Efficiency of the Treasury Bill Futures Market, *Journal of Finance,* 34:895–914.

Whaley, R. (1986): "The Valuation of American Futures Options: Theory and Empirical Tests," *Journal of Finance,* 1:127–150.