Forecasting Market Returns using the Put-Call Parity Bias of Transactions Level Options Data

Cameron Rookley
Department of Finance
University of Arizona

April, 1998
(2nd Draft)

Several methods currently exist for extracting some of the information embedded in option prices. To date, most of these methods have focused on the second or higher moments of the underlying asset's instantaneous return or value at expiration. In contrast to these methods, I present a method here which exploits transactions level options data to help forecast the first moment of lead returns in the underlying asset. This is accomplished by constructing a signal from what I call the "put-call-parity bias" of option quotes. Although put-call-parity intrinsically ties together the price of a put and call with the same strike and expiration, in the presence of transaction costs it is possible to observe bullish and bearish pressure based on the relative values of the bid and ask quotes for the put, call, index and risk free bond. Instead of working with these bid and ask prices directly though, I map these price into a set of 4 implied volatility values to form individual signals, and aggregate such signals using local polynomial regression; these aggregate signals are then used to forecast market returns. For the case of S&P 500 index options, these aggregate signals are shown to be related to recent and contemporaneous market performance, and more importantly appear to possess informational content for forecasting lead index returns.

I. Introduction

Options by there very nature are forward looking instruments. For this reason, many methods have been suggested as a means of exploiting the information content of option prices. Most of these methods focus on the second and higher moments of the underlying asset's instantaneous return or value at expiration. Lamaroux and Lastrapes (1993) for example, show how implied volatility can be used to help improve volatility forecasts based on historical data alone. Several methods also exist for estimating the risk neutral density of the underlying asset at expiration, such as Shimko (1993), Rubinstein (1994), Ait-Sahalia and Lo (1997), Jackwerth and Rubinstein (1997), Rookley (1997). Unfortunately, in a non-Black Scholes world, the relationship of the risk neutral density to the actual density remains a relatively open question. In fact Longstaff (1994) even suggests there is empirical evidence to reject the martingale restriction all together, implying certain options may not even be priced in the risk neutral measure.

In contrast to these other studies, very little research has been done on the use of options to forecast the first moment of the underlying asset's instantaneous return or value at expiration; this is to be expected given the dominant paradigm of risk neutral pricing. If options are priced in the risk neutral measure, without knowledge of the actual expected return to the underlying asset, all we really know with confidence is that a delta-hedged portfolio should earn an instantaneous expected return equal to the risk free rate. In fact, by making use of put-call parity option prices are often used to estimate implied risk free rates. Furthermore, although call ownership is in general a rather bullish strategy, while put ownership is in general a bearish strategy, it is difficult - although not impossible as we shall see in this paper - to extract first moment information based upon the relative prices of calls and puts as these prices are intrinsically tied together by put-call parity.

Although little research has been done on the use of option prices to infer first moment properties of the underlying asset, the ability to do so is of obvious interest to portfolio managers and individuals who implement active asset allocation models. Two studies which attempt to use options data to forecast market returns are Billingsley and Chance (1988), and Chance (1990). In both cases the authors focus on the relative volume of put and call options and suggest using put-call volume ratios as a contrarian indicator for investing in the underlying index, or rebalancing an actively managed portfolio. The authors find some evidence that a high put-call volume ratio, using OEX or total CBOE option volume, positively effects the conditional returns of the S&P 500 index. Schaeffer (1997) advocates monitoring commonly observed values from the options markets such as volume and open interest, in order to extract contrarian market signals. Although he presents very little systematic evidence to suggest such an approach is fruitful, he does introduce a lot of anecdotal evidence to support one of the major themes in his book: call option buyers tend to be overly enthusiastic during market tops and put buyers tend to be overly pessimistic regarding the underlying asset during market bottoms.

In this paper, I move beyond readily observed statistics from the options markets such as volume and open interest and construct a signal using transactions level options data. This signal is based upon what I call the “put-call-parity bias” of option prices. Unlike the few previous studies in this area, I have no a-priori belief as to whether the signal I construct will be contrarian in nature, or actually indicate foresight on the part of options market participants.

For European options, put-call-parity is often presented as a functional identity, however in the presence of transactions costs we know that put-call-parity should only hold approximately. If we explicitly use the bid and ask prices for all four assets involved (the call, the put, the underlying and a bond), put-call-parity implies a set of two inequalities which must hold to rule out static arbitrage opportunities. Within the bounds of transactions costs, it is then possible for a put and call option with the same strike and expiration to reveal upward or downward pressure within these allowable bounds. This could occur for example if the mid-point of the call and put’s bid and ask prices are misaligned according to put-call parity, yet based on all four asset’s bid and ask prices no static arbitrage opportunities exist. The signal I use in this paper is based upon this intuition, however instead of using these prices directly I convert option prices to implied volatilities and form a signal in implied volatility space. Since options data contain a high level of noise, several individual put-call-parity signals are aggregated using local polynomial regression to form a set of four daily signals for each trading day; this aggregation allows us successfully integrate over much of the noise contained in individual prices. For the case of S&P 500 index options, the level of this daily signal is shown to be statistically and economically significant in forecasting lead returns over several day horizons.

The rest of the article is organized as follows: in Section II I review put-call-parity in the context of implied volatility; in Section III I discusses how to construct signals using S&P 500 index options; the data and results are presented in Sections IV and V; Section VI concludes.

II. Put-Call Parity in Implied volatility Space

The familiar put-call parity formula which ties together the prices of European Puts and Calls, such as S&P 500 index options, is usually expressed as an equality as follows :

$$P \ + \ S \ = \ C \ + \ X(1+r)^{-\tau} \ + \ Divs$$
$$Divs \ = \ \sum_{t=1}^{T} D_t (1+r)^{-(\frac{t}{365})} \qquad \textbf{(1)}$$

where P is the price of a put with strike X and τ years to expiration, S is the current price of the underlying asset, C is the price of a call with strike X and τ years to expiration, and Divs is the present value of expected dividend payments between now and expiration. This equality rules out two possible pure arbitrage strategies:

1) Purchase the put and the underlying asset now, invest any dividends between now and expiration at the risk free rate and sell the call and a bond with face value $(X+Divs(1+r)^{\tau})$ at a greater price than the cost of the put and underlying asset.

2) Short the put and underlying asset, pay back all dividend payments with interest at expiration, and purchase the call and bond at a price lower than the proceeds from the put and underlying asset.

Recall that these strategies, if feasible, would be risk free arbitrage opportunities as the portfolios on the left and right side of (1) have identical payoffs at expiration regardless of the value of the underlying asset. If we introduce transactions costs in the form of bid-ask spreads on all four assets involved, put-call parity takes the form of the following inequalities:

$$P_{ask} \ + \ S_{ask} \ \geq \ C_{bid} \ + \ X(1+r_{borrow})^{-\tau} \ + \ (\frac{(1+r_{lend})}{(1+r_{borrow})})^{\tau} Divs_{lend}$$
$$P_{bid} \ + \ S_{bid} \ \leq \ C_{ask} \ + \ X(1+r_{lend})^{-\tau} \ + \ (\frac{(1+r_{borrow})}{(1+r_{lend})})^{\tau} Divs_{borrow} \qquad \textbf{(2)}$$
$$where: \quad P_{ask} \geq P_{bid} \ , \ C_{ask} \geq C_{bid} \ , \ S_{ask} \geq S_{bid} \ , \ r_{borrow} \geq r_{lend}$$

We now have two prices for the put, call, and underlying asset, and two risk free rates which can be calculated from the bid and ask prices of a t-bill maturing the same week the options expire.

The rate we use will then depend upon whether the arbitrager is buying or selling the bond. The first inequality involves buying the assets on left hand side, and selling the assets on the right hand side, which includes a bond with face value $X+(1+r_{lend})^{\tau}Divs_{lend}$. Since we are selling the bond, we are borrowing at the higher rate r_{borrow} ; at the same time, we assume the stock holder reinvests all dividends received between now and expiration at the rate r_{lend} . $Divs_{lend}$ therefore represents the present value of all dividends paid, where the discounting is done with the lower rate r_{lend}, and $(1+r_{lend})^{\tau}Divs_{lend}$ represents the future value of accumulated dividends with interest. The second inequality involves selling the assets on the left hand side and buying the assets on the right hand side. In this case, the short seller of the stock is responsible for all dividend payments with interest. Since the stock holder is short, we use r_{borrow} to calculate the accumulated interest on these payments.

If we focus on the bid and ask prices of the puts and calls only, it is possible to reduce the put-call-parity inequalities to expressions such as:

$$\begin{gathered} P_{ask}-C_{bid}\geq a \\ P_{bid}-C_{ask}\geq b \\ P_{ask}\geq P_{bid} \quad , \quad C_{ask}\geq P_{ask} \end{gathered} \tag{3}$$

where the constants a and b depend upon the bid and ask prices of the other assets involved. Even this set of expressions is difficult to work with though due to the constants a and b. Also as we compare options with differing times to expiration and degrees of moneyness, the absolute spreads will tend to show a lot of variation. For this reason, it is more convenient to standardize the option bid and ask prices by converting these values to Black-Scholes implied volatilities. Furthermore, the use of implied volatilities is ideal as regardless of the validity of the Black-Scholes assumptions, put-call parity requires a put and call option with the same strike and expiration to have the same implied volatility. The proof of this proposition is given below.

Proposition 1: Ignoring transactions costs, put-call-parity requires the implied volatility of a European call option with the same strike and expiration as a European put option, to have the same implied volatility. i.e. $\sigma_c=\sigma_p$

Proof: Proposition 1 is best illustrated by constructing a synthetic put option from the other three assets. Letting $C(\sigma_c)$ and $P(\sigma_p)$ represent the Black-Scholes values for a call and put option with implied volatilities of σ_c , and $\sigma_{p.}$ respectively, it is possible to construct a synthetic put option with price P_{syn} using put-call-parity:

$$C(\sigma_c) = (S-Divs)N(d_1(\sigma_c)) - X(1+r)^{-\tau}N(d_2(\sigma_c))$$

$$P(\sigma_p) = X(1+r)^{-\tau}N(-d_2(\sigma_p)) - (S-Divs)N(-d_1(\sigma_p)) \quad \textbf{(4)}$$

$$P_{syn} = C(\sigma_c) - (S-Divs) + (1+r)^{-\tau}X$$

Now substituting the Black-Scholes value of the call into the synthetic put expression, and collecting terms:

$$P_{syn} = (S-Divs)N(d_1(\sigma_c)) - X(1+r)^{-\tau}N(d_2(\sigma_c)) - (S-Divs) + (1+r)^{-\tau}X$$

$$P_{syn} = X(1+r)^{-\tau}(1 - N(d_2(\sigma_c))) - (S-Divs)(1 - N(d_1(\sigma_c))) \quad \textbf{(5)}$$

$$P_{syn} = X(1+r)^{-\tau}N(-d_2(\sigma_c)) - (S-Divs)N(-d_1(\sigma_c)) = P(\sigma_c)$$

Where the last line is from the symmetry of the normal distribution. It therefore follows that if put-call parity implies P_{syn}=$P(\sigma_c)$, it must be the case that $P(\sigma_p)$=$P(\sigma_c)$; if this were not the case, risk free arbitrage opportunities would exist by buying the actual put and selling the synthetic put or vice-versa. Finally since we are treating all other inputs other than implied volatility as constant, σ_c=σ_p naturally follows.

Once we introduce transaction costs, put-call-parity is no longer a simple equality in implied volatility space. It is possible to proceed though by explicitly considering the two possible arbitrage strategies discussed above. I therefore construct implied volatilities corresponding to all four option quotes where the other inputs are chosen based upon the attempted arbitrage strategy involved. The table below summarizes the inputs used in calculating implied volatilities for each quote. The face value of the bond in each case will be equal to the strike price plus the future value of dividends either received or owed.

Table 1: Inputs for Implied Volatility Calculations

Strategy	Implied Volatility	Option Quote	Underlying Asset Price	Risk Free Rate	Future Value of Dividends
Buy Put and the Underlying	σ_{Pask}	Put Ask	Ask Price	Borrowing	$(1+r_{lend})^{\tau}Divs_{lend}$
Sell Call and the Bond	σ_{Cbid}	Call Bid	Ask Price	Borrowing	$(1+r_{lend})^{\tau}Divs_{lend}$
Sell Put and Underlying	σ_{Pbid}	Put Bid	Bid Price	Lending Rate	$(1+r_{borrow})^{\tau}Divs_{borrow}$
Buy Call and Bond	σ_{Cask}	Call Ask	Bid Price	Lending Rate	$(1+r_{borrow})^{\tau}Divs_{borrow}$

Note that holding put and call prices constant, either a larger underlying asset price or larger risk free rate corresponds to lower implied volatilities for calls and higher implied volatilities for puts. For either puts or calls, a higher option price always corresponds to a higher implied volatility, ceteris paribus. It naturally follow then that $\sigma_{Pask}>\sigma_{Pbid}$, and $\sigma_{Cask}>\sigma_{Cbid.}$ Furthermore, to rule out static arbitrage opportunities it must be the case that $\sigma_{Pask}>\sigma_{Cbid}$, and $\sigma_{Cask}>\sigma_{Pbid.}$ To see why these last two inequalities must hold, once again consider the sale of a synthetic put option, which involves selling the call, selling the bond (borrowing), and buying the stock.

Proposition 2a: $\sigma_{pask}>\sigma_{cbid}$

Proof: The proceeds from selling a synthetic put option P_{SYNBid} ,when transactions costs are explicitly considered, is given as follows:

$$P_{SYNBid} = S_{ask} - X(1 + r_{borrow})^{-\tau} - C(\sigma_{Cbid}) + Divs^*$$

$$where: \quad Divs^* = \frac{(1+r_{lend})^{\tau} Divs_{lend}}{(1+r_{borrow})^{\tau}}$$

$$P_{SYNBid} = S_{ask} - X(1 + r_{borrow})^{-\tau} - (S_{ask} - Divs^*)N(d_1(\sigma_{Cbid}))$$

$$+ X(1 + r_{borrow})^{-\tau} N(d_2(\sigma_{Cbid})) + Divs^* \qquad \textbf{(6)}$$

$$P_{SYNBid} = X(1 + r_{borrow})^{-\tau}(1 - N(d_2(\sigma_{Cbid})) - (S_{ask} - Divs^*)(1 - N(d_1(\sigma_{Cbid}))$$

$$P_{SYNBid} = X(1 + r_{borrow})^{-\tau}(N(-d_2(\sigma_{Cbid})) - (S_{ask} - Divs^*)(N(-d_1(\sigma_{Cbid}))$$

$$P_{SYNBid} = P(\sigma_{Cbid})$$

$$\therefore \quad \sigma_{Cbid} > \sigma_{Pask} \Rightarrow P_{SYNBid} > P_{ask}$$

It naturally follows then, that $\sigma_{Cbid} < \sigma_{Pask}$, otherwise risk free arbitrage opportunities would exist by purchasing the put at a price P_{ask} and selling a synthetic put at a higher price P_{SYNBid}.

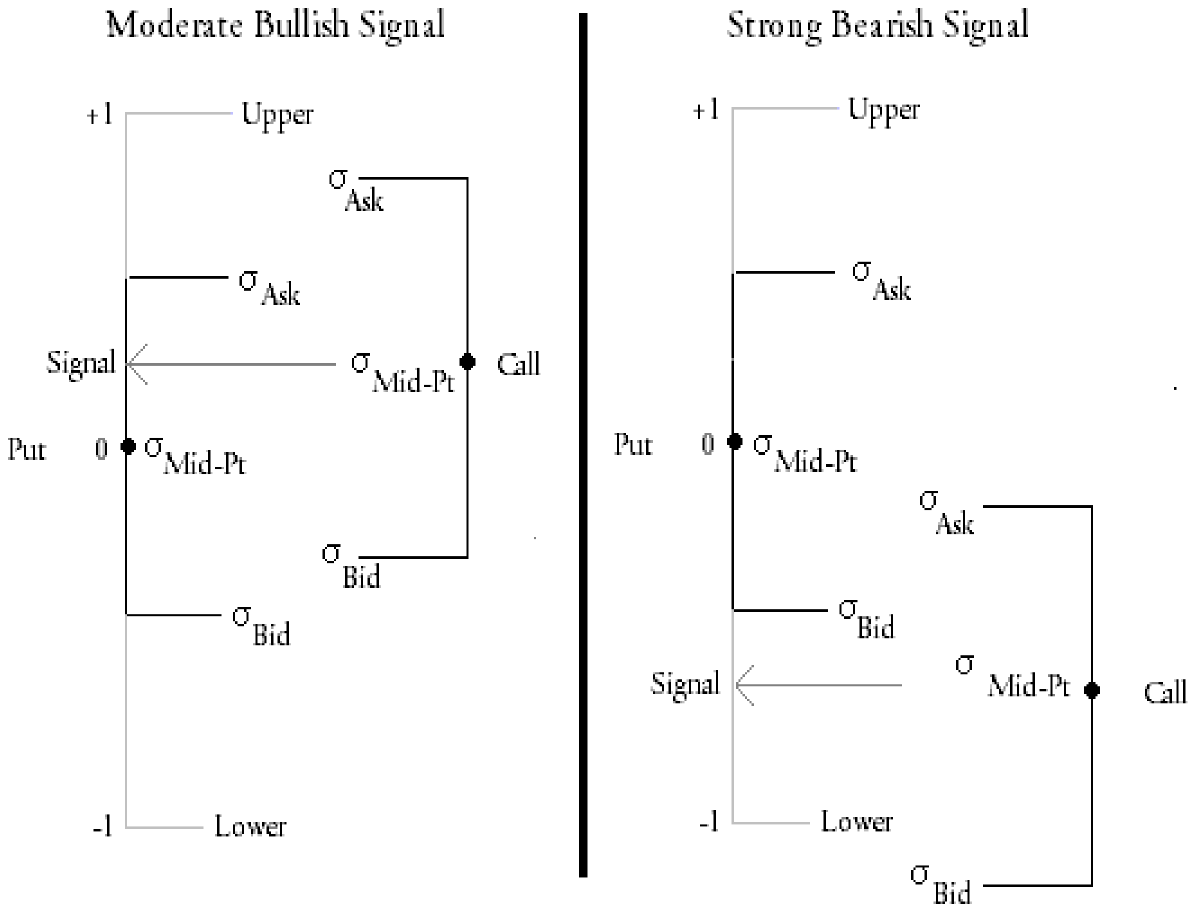
Proposition 2b: $\sigma_{cask} > \sigma_{pbid}$

Proof: The proof is analogous to Proposition 2a. In this case it can be shown that the cost of purchasing a synthetic put option must be less than the proceeds from selling the put option outright.

III. Constructing the Signal

Based on the previous discussion, we have four inequalities describing the put-call-parity relationship in implied volatility space. $\sigma_{Cbid} > \sigma_{Pask}$, $\sigma_{Cask} > \sigma_{Pbid}$, $\sigma_{Cask} > \sigma_{Cbid}$ & $\sigma_{Pask} > \sigma_{Pbid}$. Diagram 1 (below) illustrates two examples of what I call “put-call-parity bias” where the bid-ask midpoints of the put and call do not line up in implied volatility space, yet based on the inequalities above, they are not misaligned enough to offer any risk free arbitrage opportunities.

Diagram 1: Construction of Put-Call-Parity Bias Signals



Each of the two panels above are plotted in implied volatility space, as opposed to price space. The left panel depicts a moderate bull signal while the right panel is an example of what can be viewed as a strong bear signal. In each panel the dark C-shaped objects represent the bid-ask spread on the put, while the dark inverse C-shaped objects represent the bid-ask spread on a call. Given a particular bid-ask spread on the put, the grey brackets extending to the words "upper" and "lower", represent the upper and lower arbitrage free bounds for the call's bid-ask spread mid-point. Specifically:

$$upper = \sigma_{Pask} + 0.5(\sigma_{Cask} - \sigma_{Cbid})$$
$$lower = \sigma_{Pbid} - 0.5(\sigma_{Cask} - \sigma_{Cbid}) \qquad \textbf{(7)}$$

To quantify the extent of put-call-parity bias, I assign a value of -1 to “lower”, a value of 0 to the put’s bid-ask midpoint, and a value of +1 to “upper”. Therefore a call bid-ask midpoint at +1 (-1) is as high (low) as it can be without violating static arbitrage opportunities. The put-call parity bias signal for intermediate values is then given as:

$$signal = \frac{0.5(\sigma_{Cbid} + \sigma_{Cask}) - Lower}{(upper - lower)} \qquad \textbf{(8)}$$

When using transactions level data, the implied volatilities for individual put and call option pairs and their corresponding signals will in general be rather noisy. In fact it is not uncommon to observe individual signals which exceed one in absolute value, thereby indicating the existence of static arbitrage opportunities. In order to exploit the information content of multiple put-call parity pairs, and as a means of integrating over much of this noise, I carefully screen put and call quote pairs with the same strike and expiration, and aggregate individual signals into stronger signals using local linear regression. The exact details of the screening process are described in the data section which follows.

Once the data is in place, I calculate individual signals and smooth them with respect to call option moneyness and time of day using local polynomial regression, so as to obtain aggregate signals; call option moneyness is defined in this case as the dividend adjusted index price divided by strike M=(S-Divs)/X. This provides a mean signal surface, from which I can extract a fitted put-call-parity signal for each trading day corresponding to a particular time of day and level of moneyness; for the analysis which follows I choose the fitted signal corresponding to noon, and moneyness exactly equal to 1. A secondary signal, which I refer to as the put-call-parity spread is also obtained from the smooth corresponding to the difference between the noon signal with 105% moneyness less the noon signal with 95% moneyness. Recall that a large put-call-parity signal implies the price of call options is high relative to that of put options, and can be viewed as a bullish signal or as an indication of positive market sentiment among option traders. The put-call-parity spread on the other hand should be inversely related to market sentiment; if market participants are bullish, out of the money call options will be popular relative to in-the-money call options as out of the money calls offer more of a leverage effect should the market turn up.

While the two put-call parity signals I describe above measure the relative price of puts and calls, the absolute level of prices is well captured by the implied volatility values themselves. Therefore, in order to control for the absolute level of option prices , I also estimate an implied volatility surface which is consistent with the estimated signal surface. To do this, I average all four implied volatilities used to create individual signals and smooth these average implied volatility values with respect to call option moneyness in both directions and time of day in one direction. From this surface I then extract a fitted implied volatility value corresponding to an at-the-money option as of noon on each day. I also extract an implied volatility spread which is equal to the difference between the 105% noon moneyness and 95% noon moneyness implied volatility values in the surface.[1]

In order to arrive at the put-call-parity signal surface and implied volatility surfaces I describe above, I use a variation of local polynomial regression which smooths in one direction for the time dimension and in both directions for call option moneyness. Specifically let (M_{targ},τ) represent a target point of interest, where M is option moneyness, and t is time measured in minutes past 8:00 a.m. In order to estimate a function at this point, I fit a local linear regression at the target using nearby observations which are weighted by their distance to the point. In this case I have a trivariate data set which consists of n observations on moneyness, time to expiration and either individual signals or implied volatility values (which I'll denote as S_j for both in the exposition which follows) $\{M_j,t_j,S_j\}$, j=1,....n. I then obtain the fitted value corresponding to the

[1]The implied volatility spread is essentially measuring the slope of the familiar volatility "smile", which is related to deviations of the risk neutral density from log-normality. Implied volatility values which are relatively higher for large values of moneyness, imply a risk neutral density which is more skewed to the left than a log-normal distribution.

target point as:

$$E(S(M_{targ},\tau)) \;=\; \hat{S}(M_{targ},\tau) \;=\; e^{/}\hat{\beta}_i \;=\; e^{/}(X_i^{/}\;W_i\;X_i)^{-1}X_i^{/}\;W_i\;Y$$

$$e^{/}=(1\;\;0\;\;0)$$

$$X_i \;=\; \begin{array}{ccc} 1 & (M_1-M_{targ}) & (t_1-\tau) \\ 1 & (M_2-M_{targ}) & (t_2-\tau) \\ \\ .... & .... & ... \\ 1 & (M_n-M_{targ}) & (t_n-\tau) \end{array} \qquad \textbf{(9)}$$

$$diag(W_i) \;=\; (\varphi((M_1-M_{targ})/h_M)\cdot\varphi((t_1-\tau))/h_t)\cdot D(t_1<\tau),...,\varphi((M_n-M_{targ})/h_M)\cdot\varphi((t_n-\tau)/h_n))\cdot D(t_n<\tau)$$

$$D(t<\tau) \;=\; \begin{array}{l} 1\;\;if\;\;t<\tau \\ 0\;\;if\;\;t>\tau \end{array}$$

$$Y^{/}=(S_1,S_2,..,S_n)$$

The estimate of $E(S(M_{targ},\tau))$ is therefore the constant term from a WLS regression where the weights along the diagonal of W are given by a modified multiplicative Gaussian kernel $\varphi((M\text{-}M_{targ})/h_{M}.)\varphi((t\text{-}\tau)/h_t)D(t<\tau)$, where the kernel is modified for this particular case so that zero weight is attached to observations which occur after the target time. Larger values of either bandwidths (h_m and h_t) lead to higher weights being attached to observations which are farther away from the target, as φ represents the Gaussian function. By increasing the values of h_m or h_t it is therefore possible to obtain increased levels of smoothing and indeed the choice of the bandwidth vector $h_m|h_t$ is often the key decision to make in the context of nonparametric regression. Since we need to smooth every day in our sample, I smooth using three different subjective bandwidth vectors . These bandwidths are arbitrarily selected, and set to $h_m|h_t$=0.1|30, $h_m|h_t$=0.2|60 and $h_m|h_t$=0.4|120, where h_t is measured in minutes past 8:00 a.m.; these bandwidths will be referred to as low, medium and high smoothing levels for the results which follow.[2]

IV. Data

The previous three sections outline a methodology for constructing aggregate signals from transactions level options data. In this section I describe the specific inputs used to construct such signals for the case of S&P 500 index options. In order to construct the signal described above, I use several sources of data spanning close to five years from January 1989 to October 1993. This data set includes: every recorded quote for S&P 500 index options on the CBOE as taken from the Berkeley Options Database; every recorded quote and trade for S&P 500 Futures contracts on the CME; the S&P 500 daily dividend record. and finally daily secondary market t-bill rates as recorded from the Wall Street Journal.

In order to arrive at implied volatility values it is necessary to observe all of the inputs which enter the Black Scholes formula. These include: strike price, time to expiration, index price, expected dividend payments, risk free rate and option prices. Strike prices and option prices are directly observable from the BODB tapes. To calculate time to expiration I count the number of calender days until expiration, including the current day and expiration date, and divide this difference by 365. Borrowing and lending risk free rates are obtained from the bid and ask prices of secondary market t-bill quotes as reported in the Wall Street Journal, where the expiration date of the bill is chosen to match as close as possible the expiration date of the option in question. Since expected dividend payments are not directly observable, I use actual dividend payments on the S&P 500 index instead, as reported in the S&P 500 daily dividend record, from the current date to option expiration date. The use of actual dividends is somewhat justified as it is relatively easy to accurately forecast dividend payments based on the current trend and seasonality of

---

[2]I should point out that several methods for automatically selecting the bandwidth exist. For a nice introduction to Kernel Smoothing see Wand and Jones (1996). Also, Ruppert (1997) presents a method of local bandwidth choice which is data-driven; in Rookley (1997) Ruppert's empirical method of local bandwidth choice is used for smoothing intraday implied volatility surfaces and in estimating risk neutral densities. Unfortunately many of these methods are computationally intensive and their advantage over what's loosely referred to as the "eyeball metric" is often minimal.

historical dividend payments. Finally, instead of using the actual S&P 500 index value as reported on the Berkeley Options Database tapes, I construct a minute by minute implied index series using all available S&P 500 futures quotes, actual dividends paid for the S&P 500 index, and a set or repo rates which I imply from the data (see Rookley, 97 for details).

Once the basic Black Scholes inputs are place, I look for matching put and call option pairs and apply a screening process before including them in the overall database. This screening process is aimed at obtaining as much data as possible while attempting to maintain the synchroneity of both option and implied index quotes. It is also important to avoid quotes which generate unreasonable implied volatility values due to illiquidity, little time to expiration, or other factors. With these goals in mind, I begin by specifying a set of target times across each day which are ten minutes apart and search for put and call option quotes with the same strike and expiration as close to these target times as possible. If the corresponding put and call quotes occur within ten minutes of each other and if the value of the underlying index at these corresponding times deviates by less than 0.1%, these quotes are considered as candidates for inclusion in the overall data set. Since options with the nearest maturities tend to be more liquid, for the purpose of what follows I only use candidate option quotes corresponding to the closest expiration date with more than 14 calender days until expiration. Finally, when I map asset bid and ask prices to the four implied volatility values needed to construct individual signals, I require the output of this mapping to yield reasonable implied volatility values where for the purpose of this study I define "reasonable" as being between 0 and 1.

After imposing the filters described above the four implied volatilities for the bid and ask price of each put and call option pair are calculated. These 4 implied volatilities are then converted to individual signals according to Equations 7 and 8, and smoothed according to Equation 9, so as to obtain the necessary signal surface. Implied volatility foursomes are also averaged and an implied volatility surface is estimated for each trading day as well. From these two surfaces I then extract the put-call-parity signal, the put-call parity signal spread, at-the-money implied volatility, and the implied volatility spread for subsequent analysis.

As well as calculating the signals of interest, several volume statistics are also extracted from the Berkeley tapes. These include: total option volume, total call volume and total put volume for each trading day. Also as another means of identifying buying and selling pressure for puts and calls, the percentage of option volume and percentage of option trades occurring at or above the ask price, and the percentage of option volume and percentage of option trades occurring at or below the bid price are recorded. Intuitively if investors are dumping call options, for example, we would expect to see a high percentage of call volume and total call trades occurring at or below the bid price, whereas call buying pressure would manifest itself in a high percentage of call volume and trades occurring at or above the ask price.

V. Results

Tables 2 & 3 (attached) provide simple correlation statistics between lagged and lead index returns and the variables described in the data section above. In brackets next to each of these variables are market sentiment tags indicating the relationship between its level and the optimism of option market participants. For example, a high value of the put-call-parity signal is an indication of bullish behavior on the part of option traders, and this variable is tagged accordingly. In Table 2 the first four rows correspond to the aggregated variables which are obtained by smoothing data points occurring before noon. It is therefore natural to look at the correlation between these variables and noon-to-noon index returns.[3] Since the volume statistics are based on all transactions occurring between open and close of the option markets, correlations for these statistics are made with respect to the open to close returns of the actual index. By examining the correlations between lagged returns and the variables down the leftmost column of the table, we note relatively strong relationships between previous market performance and the current levels of some of these variables. First, the relative price of calls to puts tends to rise due to previous market rallies as indicated by the positive correlations between the put-call parity signal and lagged market returns; the positive relationship between the put-call-parity spread and lagged returns is somewhat counter-intuitive however given this first

---

[3]Actually, instead of using the S&P 500 index returns, I continue to use implied index values, which are representative of the returns one would earn holding S&P 500 futures contracts.

result. Continuing down the list, during market downturns it is common for the overall level of implied volatility to rise and this is captured by the negative correlation between lagged market returns and at-the-money implied volatility; the implied volatility spread on the other hand appears to show little relationship with market returns at most leads and lags. The behavior of the volume statistics also suggests somewhat of a bandwagon effect. Based on the correlation coefficients, when markets are rising interest in call options appears to surge as a large percentage of total trades and total volume occurs at or above the ask for calls, while puts are dumped with a large percentage of trades occurring at or below the bid, and vice-versa during market downturns. This story is particularly compelling when we examine the very strong correlation coefficients between the volume statistics and concurrent return data. Finally, overall interest in options appears to rise during market downturns, presumably due to an interest in options as "portfolio insurance". The negative correlation between concurrent returns and put volume indicates overall interest in put options rises during downturns.

The first 4 columns of Table 2 illustrate how option markets react to the overall performance of the underlying asset, bidding up call prices during rallies and bidding up put prices (and the overall price of all options as captured by implied volatility) during market declines. Is such behavior rational though? More importantly do the relative prices of put and call options provide any information regarding the conditional return of the index in the future? Looking at the last 6 columns of Table 2, we see a very weak relationship if any between signal levels and lead returns which is consistent with a belief in market efficiency. Although weak, two questions immediately come to mind. First, are these correlation coefficients statistically significant? Second, are these correlation coefficients economically significant? The remainder of this paper will tend to focus on these two questions. Since the correlation coefficients for leads or lags greater than one involve overlapping data, standard statistical test cannot be employed. However with reference to the 1 day lead column, given 1185 daily returns the standard error of a sample correlation coefficient equal to 0.05 in absolute value is approximately equal to 0.029. Therefore the first three variables are on the threshold of being statistically significant using traditional confidence levels in forecasting one day lead returns. Moreover, the statistical results which follow will indeed indicate the informational content of these variables for forecasting. Before proceeding to these results, it is interesting to examine the content of Table 3; here we observe rather weak relationships between the put-call-parity signal, implied volatility, and the volume statistics. This is a potentially positive feature if one wishes to incorporate multiple sources of information into signals for asset allocation models. Not surprisingly the volume statistics themselves are highly correlatedb with each other.

Moving beyond the simple correlation coefficients, Tables 4 through 6, and Figures 1 and 2 present multi-variate regression results, where lead noon to noon returns are regressed on the put-call-parity signal, the put-call-parity signal spread, implied volatility and the implied volatility spread, using three levels of smoothing for all four of these aggregated variables. Since these smoothing levels are arbitrarily chosen, I present full results for all three bandwidth choices in Table 4, and Figures 1 and 2. In order to fully exploit the entire universe of returns, overlapping returns were used for lead lengths greater than one which requires the use of adjusted standard errors for the regression coefficients; these are obtained using a Newey-West (1987) estimator of the variance covariance matrix, with a bandwidth chosen using the method of Andrews (1991). All t-statistics in Table 4 are therefore based on a heteroscedastic and autocorrelated consistent estimate of the variance covariance matrix.[4] Looking at the sign of the individual regression coefficients, we note that the put-call-parity signal coefficient is positive up to 30 leads for all three smoothing levels, the put-call-parity signal spread coefficient is negative (or zero) at all leads, while the implied volatility coefficient is positive at all leads. The implied volatility spread is neither consistently positive or negative and has very low t-stats at virtually all leads and bandwidth choices. Based on this preliminary observation it appears that the option market correctly "predicts" index returns, according to the put-call-parity signal and signal spread. The positive relationship between implied volatility and market returns can be interpreted as a risk-premium effect, where risk in this case might be partially captured by implied volatility. Figures

[4]See Hamilton (1994), pages 280-290 for an overview of the mechanics of this estimator. Gauss code for this type of estimator can also be found on my webpage at: http://www.goodnet.com/~dh74673/gcode.htm

1 and 2 focus on the magnitude of the put-call-parity signal coefficients and the signal spread coefficients.

First in reference to Figure 1-A, the put call parity spread coefficient has the nice property of increasing almost linearly with the lead length out to about 14 lags. This is more or less true for all three bandwidth levels. In terms of statistical significance, the t-stats plotted in Figure 1-B exceed 2 at several lead lengths for the case of a low bandwidth (little aggregation of individual signals), and is similarly shaped but relatively lower for the other 2 bandwidth levels. The t-statistic of 2.30 corresponding to the put-call-parity signal for 1 lead length, using low smoothing is very suggestive of the informational content of this signal, particularly when viewed in light of the coefficient's steady climb in magnitude out to 14 day leads.

Figures 2-A and 2-B present similar results, but for the put-call parity spread. For this coefficient the direction is negative and increasing in absolute magnitude out to about 9-10 lead lengths. In contrast to the put-call-parity signal itself, the higher bandwidth (higher levels of aggregation) tend to produce more significant t-stats, while all three smoothing levels produce a negative and significant t-stat at lead length 4.

Although the magnitudes and t-stats of these regression statistics are compelling, it might be the case that their values are being generated by only a few outliers. Tables 5 and 6 present results which are more robust to this possibility. In Table 5 the relationship between fitted conditional expected return values and actual average return values for this multi-variate regression are investigated, for various lead lengths and levels of smoothing. For each lead length and level of smoothing, OLS coefficients are estimated and fitted values for the expected conditional return are calculated. These fitted values are then sorted from lowest to highest and placed into quartiles. The first two columns of each bandwidth panel give the mean and median expected conditional return for each quartile and lead length. The third and fourth columns of each bandwidth panel give the actual mean and median returns for each quartile and lead length. Although the ordinal ranking of actual mean and median returns is not perfect for the second and third quartiles in each box, in every single case the mean and median actual return for the fourth quartile is higher than the first. More importantly the difference between these values widens as we consider longer lead lengths. Although out of sample testing is still necessary, these results are encouraging when contemplating applying the signal to asset allocation decisions.

Table 6 is similar in spirit to Table 5; in this case a "hit table" is generated from the same multi-variate regression of lead returns regressed on the four independent variables. Once again I sort fitted values from lowest to highest and arrange these values into quartiles. I then calculate the percentage of individual actual returns which are greater than the median actual return for each lead length and quartile. If we have a random signal, or a signal which was being generated by a few outliers we would expect the percentages of this table to be approximately 50% across the board. If on the other hand the signal tends to add forecasting value on a day-by-day basis we would expect percentages below 50% for the first two quartiles and percentages greater than 50% for the third and fourth. Although the results for the second and third quartiles are far from perfect, in every single case for all leads and bandwidths every first quartile percentage is less than 50% and every fourth quartile percentage is greater than 50%, once again adding to the efficacy of using the signal in practice.

Most of the evidence presented so far has addressed the statistical significance issue. The last set of tests I conduct seeks to address the question of economic significance. Given fitted values for conditional returns, is it possible to use these fitted values as a signal for timing the market for purposes of active asset allocation? Figures 3-A through 3-G give the cumulative wealth associated with investing $1 in the S&P 500 index at the beginning of 1989 and shifting all cumulative wealth between the S&P 500 index and t-bills according to the daily signal received. In particular, I regress 1 day lead returns on the four independent variables which were formed using a low bandwidth for aggregation purposes; the percentiles corresponding to the fitted values of this regression are then calculated and used to form buy/sell/hold signals which depend upon a particular market timing strategy. Since the market timing rule is somewhat arbitrary I consider seven possible rules, which entails specifying a buy percentile threshold and a selling percentile threshold. If the fitted value for conditional returns passes above the buy threshold then a position in the index is assumed until a sell signal is received due to fitted values falling below the sell percentile threshold. Given that stocks have historically outperformed t-bills, it may be desirable to have a trading rule with a “hold the index bias” which assumes a position in the index by default unless a sell signal is currently in force. In this case we would hold the index unless our signal provides a good reason to do otherwise.

Table 7 below summarizes the performance and number of trades involved for seven arbitrarily chosen trading rules. The last three trading rules have a “hold the index bias”, while the first four can lead to rather lengthy periods where t-bills are being held as opposed to the index; these periods show up in Figures 3-A to 3-G in the form of relatively linear lines across multiple trading days.[5]

In interpreting these results several caveats are in order: first, these are in-sample results and more stringent out of sample tests are required before concluding the signal is tradable. This is important as individual signals borrow information contained in subsequent time observations through the values of the regression coefficients and the values of the percentile cutoffs. The performance of the trading signal is likely biased upward for this reason. On the other hand, the index alternative of investing 100% of cumulative wealth in t-bills is somewhat restrictive as most asset allocation rules allocate wealth between stocks and bonds as well as other possible asset classes. Since bonds have historically returned more than t-bills, the choice of t-bills as the alternative investment vehicle might have the effect of biasing the results downwards. Also, 100% shifts from one asset class to the other is rather uncommon in practice, and most active asset allocation rules tend to “tilt” the portfolio from one asset class to another when a strong enough signal is received.   Finally since the use of transactions level options data is a rather unique source of information, it might be the case that the information content of signals derived from this data is orthogonal to other trading signals, such as those derived from dividend discount models. It might therefore be possible to effectively include put-call-parity type signals

[5]With reference to the cumulative wealth diagrams, the buy and hold paths exhibit an unusual vertical jump in returns around trading day 750. This jump occurs due to the unavailability of CME futures data for several days. This particular positive return is therefore a multiple day return which happened to coincide with a rather bullish week in the market near the end of 1991.

within larger forecasting models.

With all of the necessary caveats out of the way, we now refer to Table 7 below, where 6 out of 7 of the arbitrarily chosen trading strategies is shown to outperform a buy and hold strategy. The best performing rule is a 50/50 rule (Figure 3-A) which outperforms buying and holding the index by 3.6% on an annualized basis (over 4.8 years).

Table 7

Performance Summary for Various Trading Rules

Buy Sell Rule	# of Trades	Final Wealth from Initial Dollar Investment	Percent Improvement over Buy and Hold	Annualized Percent Improvement over Buy and Hold
Buy & Hold	1	$1.67	0	0
50/50	324	$1.98	18.7%	3.6%
75/25	97	$1.82	9.3%	1.9%
90/10	37	$1.60	-3.9%	-0.8%
95/5	24	$1.76	5.5%	1.1%
25/25	242	$1.87	12.2%	2.4%
10/10	158	$1.67	0.1%	0.0%
5/5	88	$1.72	3.3%	0.7%

Unfortunately, once we explicitly consider transactions costs we have to face the fact that the most successful strategies tend to involve the most frequent trading and depending upon the institution doing the trading, frequent rebalancing has the potential to wipe out such gains.[6] The second and third best strategies (25/25 and 75/25) also outperform a buy and hold strategy by an economically significant amount, but once again involve a substantial number of trades. One

[6]It is interesting to note that many firms involved in index trading for large institutional investors often have active asset allocation groups which can benefit from internally crossing orders; this ability to cross orders drastically reduces and almost eliminates the transactions costs associated with actively managing a portfolio. For example, Barclay's Global Investors of San Francisco manages more than $400 Billion in index funds and crosses close to 90% of its trades internally. The ability to exploit the churning of large index fund orders for the purpose of reallocating the smaller actively managed portfolios in a strategic manner can almost eliminate the cost of frequent reallocations.

encouraging result is the 95/5 signal, which involves only 24 trades, yet provides an average annualized return 1.1% above and beyond a buy and hold strategy. Based on these in sample results, it does appear plausible that transactions level options data can be used to generate an economically significant signal for the purpose of forecasting lead returns. Moreover, since t-bills are essentially a risk free asset, any strategy which involves holding t-bills part of the time, will be less risky than a strategy which involves buying and holding the index exclusiviely. Based on the evidence presented so far, it appears that many of these market timing strategies mean-variance dominate a strategy of buying a holding.

Given that this study is based on data which ends October 1993, the more than 4 years which have since elapsed provides an excellent opportunity for testing the signal out of sample. This remains an exercise for future research.

VI Conclusion

This paper has investigated the put-call-parity relationship between put and call options with the same strike price and expiration in implied volatility space. Given the presence of transactions costs, I show how it is possible to map the bid and ask prices of the four assets involved into four implied volatility values and demonstrate the four inequalities which must hold to rule out static arbitrage opportunities. Using this mapping, the relative bullishness of option market participants can be quantified and aggregated using local polynomial regression to arrive at a set of daily “signals” for the purpose of forecasting market returns. First we note how the apparent bullishness of option traders is positively related to recent market performance, where call option prices tend to be bid up relative to put options after market rallies, and vice-versa during declines. More importantly, the results of Section V are suggestive that the signals we construct from transactions level options data are both economically and statistically significant in helping to forecast market returns and can be used for active asset allocation strategies.

References

Ait-Sahalia, Y. and A. Lo. “Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices”, working paper, 1995.

Arnott, R. and F. Fabozzi, *Active Asset Allocation: State-of-the-Art Portfolio Policies, Strategies and Tactics*. Chicago: Probus Publishing, 1992.

Black, F. and M. Scholes, 1973, “The Pricing of Options and Corporate Liabilities.” *Journal of Political Economy*, 81, 637-659.

Billingsley, Randall S., and Don M. chance, “Put-Call Ratios and Market Timing Effectiveness.” *The Journal of Portfolio Management*, Fall 1988, pp. 25-28.

Chance, Don M., “Option Volume and Stock Market Performance”, *The Journal of Portfolio Management*, Summer 1990, pp. 42-51.

Chriss, N. *Black Scholes and Beyond*, Chicago: Irwin Professional Publishing, 1997.

Gibson, Roger. *Asset Allocation: Balancing Financial Risk*. Chicago: Irwin, 1996.

Grinhold, R. and R. Kahn, *Active Portfolio Management: Quantitative Theory and Applications*. Chicago: Probus Publishing, 1995.

Hardle, W. *Applied Nonparametric Regression*. Cambridge: Cambridge University Press, 1989.

Hull, J. *Options, Futures and Other Derivative Securities*. 2nd e. Englewood Cliffs: Prentice Hall, 1993.

Jackwerth, J. C., and M. Rubinstein, 1996, “Recovering Probability Distributions from Contemporaneous Security Prices.” *Journal of Finance*, 51, 5, 1611-1631.

Kritzman, Mark. *Asset Allocation for Institutional Portfolios*. Homewood IL: Business One Irwin, 1990.

Lamoureux, C. G., and W.D. Lastrapes, 1993, “Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities,” *Review of Financial Studies*, 6, 293-326.

Longstaff, F, “Option Pricing and the Martingale Restriciton”, *Review of Financial Studies*, 8, 1091-1124.

Rookley, Cameron. "Fully Exploiting the Information Content of Intra-day Option Quotes", Working Paper, Department of Finance, University of Arizona, 1997.

Rubinstein, M., 1994, "Implied Binomial Trees" *Journal of Finance*, 49, 3, 771-818.

Ruppert, D. (1997). "Empirical-bias Bandwidths for Local Polynomial Nonparametric Regression and Density Estimation", *JASA*, 92, 1049-1062.

Schaeffer, B. *The Option Advisor*. New York: John Wiley & Sons Inc, 1997.

Shimko, D., 1993, "Bounds of Probability" *Risk*, 6, 33-37.

Wand, M.. and C. Jones, *Kernel Smoothing*. London: Chapman & Hall, 1995.

Table 2

Correlations Between Current Signal Levels and Market Returns at Various Leads & Lags

Variable/Return Lead or Lag	10 Day lag	5 Day lag	1 Day Lag	Concurrent Returns	1 Day Lead	2 Day Lead	3 Day Lead	4 Day Lead	5 Day Lead	10 Day Lead
Put-Call Parity Signal (Bull)	0.18	0.17	0.14	0.05	0.05	0.04	0.04	0.03	0.05	0.05
Put-Call Parity Signal Spread (Bear)	0.05	0.08	0.09	-0.01	-0.05	-0.03	-0.06	-0.08	-0.07	-0.07
At the Money Implied Volatility (Neut)	-0.21	-0.15	-0.08	-0.03	0.05	0.06	0.06	0.06	0.07	0.07
Implied Volatility Spread (Bear)	-0.03	-0.04	0.01	0.00	0.01	0.08	0.05	0.06	0.07	0.02
% Call Trades Above Ask (Bull)	0.17	0.21	0.19	0.61	0.04	0.00	-0.04	0.00	0.00	-0.03
% Call Volume Above Ask (Bull)	0.18	0.21	0.18	0.57	0.04	0.01	-0.03	0.00	-0.01	-0.03
% Put Trades Above Ask (Bear)	-0.11	-0.15	-0.14	-0.66	-0.04	0.00	0.04	0.01	0.00	0.05
% Put Volume Above Ask (Bear)	-0.13	-0.15	-0.13	-0.62	-0.04	0.01	0.04	0.02	0.02	0.06
% Call Trades Below Bid (Bear)	-0.10	-0.14	-0.12	-0.61	-0.04	0.01	0.03	-0.01	-0.01	0.02
% Call Volume Below Bid (Bear)	-0.09	-0.14	-0.11	-0.58	-0.04	0.01	0.04	0.00	0.01	0.02
% Put Trades Below Bid (Bull)	0.14	0.18	0.21	0.66	0.04	0.00	-0.05	0.00	0.00	-0.05
% Put Volume Below Bid (Bull)	0.13	0.17	0.20	0.62	0.05	-0.01	-0.04	0.01	-0.02	-0.05
Total Option Volume (Neut)	-0.04	-0.04	0.01	-0.11	0.02	0.00	-0.01	0.02	0.02	-0.03
Total Call Volume (Neut)	-0.03	-0.02	0.02	-0.05	0.02	0.00	-0.01	0.02	0.03	-0.02
Total Put Volume (Neut)	-0.04	-0.05	0.00	-0.16	0.01	0.00	0.00	0.02	0.00	-0.04

Notes: # of observations=1186-(# of leads or lags)

Correlations for first 4 items based on noon to noon returns, all others based on open to close.

Table 3

Cross Correlations Among Contemporaneous Signals

	Put-Call Parity Signal	At the Money Implied Volatility	% Call Trades Above Ask	% Put Trades Above Ask	% Call Trades Below Bid	% Put Trades Below Ask	Total Option Volume
Put-Call Parity Signal	1.00	-0.05	0.02	-0.01	-0.02	0.00	-0.05
At the Money Implied Volatility	-0.05	1.00	-0.04	0.03	0.03	0.01	-0.01
% Call Trades Above Ask	0.02	-0.04	1.00	-0.76	-0.93	0.83	0.00
% Put Trades Above Ask	-0.01	0.03	-0.76	1.00	0.82	-0.92	0.08
% Call Trades Below Bid	-0.02	0.03	-0.93	0.82	1.00	-0.81	0.07
% Put Trades Below Ask	0.00	0.01	0.83	-0.92	-0.81	1.00	0.02
Total Option Volume	-0.05	-0.01	0.00	0.08	0.07	0.02	1.00

Note: # of observations=1186

Table 4

Multivariate Regression Results, Returns at Various Leads Regressed on Signals in Levels

(First Row for Each Group is Coefficient, Second is t-stat)

	Low Bandwidth for Aggregating Signals					Medium Bandwidth for Aggregating Signals					High Bandwidth for Aggregating Signals				
Lead	Constant	PCP Signal	PCP Signal Spread	Implied Volatility	I.V. Spread	Constant	PCP Signal	PCP Signal Spread	Implied Volatility	I.V. Spread	Constant	PCP Signal	PCP Signal Spread	Implied Volatility	I.V. Spread
1	-0.002	0.001	-0.001	0.012	-0.001	-0.002	0.001	-0.001	0.011	0.002	-0.002	0.001	-0.002	0.011	0.002
	-1.700	2.300	-1.820	1.450	-0.116	-1.570	1.440	-1.910	1.360	0.248	-1.500	1.260	-2.200	1.360	0.202
2	-0.003	0.001	0.000	0.020	-0.001	-0.003	0.001	-0.001	0.022	-0.005	-0.003	0.001	-0.002	0.022	-0.007
	-1.450	1.620	-0.728	1.550	-0.096	-1.410	1.400	-1.170	1.630	-0.284	-1.340	1.420	-1.280	1.630	-0.436
3	-0.004	0.002	-0.001	0.029	-0.004	-0.004	0.002	-0.003	0.030	-0.011	-0.004	0.001	-0.004	0.031	-0.013
	-1.380	1.770	-1.580	1.480	-0.161	-1.190	1.280	-1.980	1.510	-0.458	-1.120	1.100	-2.400	1.500	-0.550
4	-0.005	0.002	-0.002	0.038	-0.013	-0.004	0.002	-0.005	0.038	-0.023	-0.004	0.002	-0.006	0.038	-0.025
	-1.210	2.010	-2.200	1.430	-0.473	-0.888	1.180	-2.480	1.360	-0.750	-0.849	1.020	-2.780	1.360	-0.806
5	-0.007	0.003	-0.001	0.044	-0.002	-0.006	0.002	-0.004	0.045	-0.017	-0.006	0.002	-0.006	0.046	-0.019
	-1.310	1.980	-1.470	1.350	-0.066	-1.060	1.550	-2.000	1.350	-0.484	-1.000	1.390	-2.390	1.340	-0.537
6	-0.007	0.003	-0.002	0.049	-0.007	-0.006	0.003	-0.005	0.050	-0.023	-0.006	0.003	-0.007	0.050	-0.028
	-1.140	1.960	-1.500	1.280	-0.194	-0.910	1.770	-1.980	1.250	-0.598	-0.836	1.720	-2.370	1.240	-0.727
7	-0.007	0.003	-0.002	0.054	-0.005	-0.007	0.003	-0.006	0.055	-0.016	-0.007	0.003	-0.008	0.056	-0.020
	-1.030	1.830	-1.590	1.200	-0.137	-0.841	1.440	-2.200	1.160	-0.385	-0.816	1.450	-2.530	1.160	-0.472
8	-0.007	0.003	-0.002	0.055	-0.008	-0.006	0.003	-0.006	0.056	-0.018	-0.006	0.002	-0.009	0.056	-0.021
	-0.808	1.820	-1.400	1.030	-0.183	-0.658	1.400	-1.870	1.000	-0.397	-0.640	1.310	-2.200	0.999	-0.459
9	-0.007	0.004	-0.002	0.055	-0.005	-0.006	0.003	-0.006	0.056	-0.019	-0.006	0.002	-0.009	0.056	-0.018
	-0.699	1.590	-1.240	0.901	-0.121	-0.557	1.210	-1.820	0.880	-0.395	-0.548	1.090	-1.990	0.872	-0.379
10	-0.007	0.004	-0.003	0.059	-0.002	-0.007	0.003	-0.006	0.060	-0.014	-0.007	0.003	-0.008	0.061	-0.015
	-0.672	1.510	-1.380	0.875	-0.038	-0.563	1.220	-1.700	0.854	-0.267	-0.569	1.200	-1.730	0.852	-0.273
15	-0.010	0.006	-0.001	0.101	-0.053	-0.008	0.006	-0.003	0.096	-0.069	-0.008	0.006	-0.003	0.096	-0.076
	-0.571	2.190	-0.479	0.957	-0.811	-0.432	1.770	-0.649	0.897	-0.925	-0.412	1.810	-0.562	0.892	-1.000
20	-0.012	0.003	-0.001	0.137	-0.044	-0.010	0.002	-0.005	0.135	-0.052	-0.010	0.002	-0.005	0.135	-0.059
	-0.519	0.961	-0.352	0.985	-0.532	-0.441	0.539	-0.745	0.962	-0.556	-0.432	0.609	-0.654	0.966	-0.640
30	-0.020	0.004	-0.001	0.174	0.033	-0.019	0.004	-0.004	0.175	0.024	-0.019	0.005	-0.004	0.176	0.013
	-0.705	1.330	-0.349	1.070	0.269	-0.672	1.210	-0.579	1.080	0.167	-0.674	1.350	-0.426	1.090	0.092
40	-0.015	0.000	0.000	0.219	-0.016	-0.016	-0.001	-0.003	0.219	-0.011	-0.016	0.000	-0.005	0.223	-0.027
	-0.486	-0.104	0.051	1.290	-0.088	-0.474	-0.231	-0.418	1.290	-0.054	-0.486	0.054	-0.492	1.320	-0.124
50	-0.012	-0.005	-0.006	0.228	0.012	-0.012	-0.006	-0.013	0.228	0.026	-0.014	-0.006	-0.017	0.230	0.027
	-0.324	-0.769	-1.080	1.320	0.050	-0.319	-0.918	-1.380	1.310	0.093	-0.350	-0.863	-1.370	1.330	0.095

Table 5
Expected Returns and Actual Returns Based on Signal Quartiles

Lead Length	Quartile	Low Bandwidth for Aggregating Signal				Medium Bandwidth for Aggregating Signal				High Bandwidth for Aggregating Signal			
		Average Expected Conditional Return		Actual Return		Average Expected Conditional Return		Actual Return		Average Expected Conditional Return		Actual Return	
		Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
1	1	-0.047%	-0.037%	-0.019%	-0.002%	-0.039%	-0.029%	0.020%	-0.037%	-0.042%	-0.032%	-0.004%	-0.037%
	2	0.026%	0.026%	0.003%	-0.021%	0.026%	0.027%	-0.021%	0.018%	0.027%	0.027%	-0.022%	0.027%
	3	0.071%	0.068%	0.042%	0.048%	0.070%	0.069%	0.022%	0.043%	0.071%	0.071%	0.046%	0.028%
	4	0.141%	0.129%	0.163%	0.284%	0.133%	0.122%	0.169%	0.272%	0.134%	0.123%	0.169%	0.270%
2	1	-0.013%	-0.006%	-0.050%	-0.047%	-0.027%	-0.014%	-0.030%	-0.033%	-0.032%	-0.019%	0.018%	0.009%
	2	0.063%	0.064%	0.089%	0.128%	0.063%	0.064%	0.058%	0.095%	0.063%	0.064%	0.063%	0.107%
	3	0.115%	0.114%	0.176%	0.168%	0.121%	0.119%	0.127%	0.116%	0.124%	0.124%	0.076%	0.056%
	4	0.213%	0.197%	0.162%	0.165%	0.220%	0.208%	0.222%	0.251%	0.222%	0.206%	0.220%	0.256%
3	1	-0.022%	-0.007%	-0.006%	0.016%	-0.045%	-0.018%	0.017%	0.016%	-0.059%	-0.032%	0.047%	0.056%
	2	0.096%	0.097%	0.123%	0.193%	0.100%	0.100%	0.065%	0.156%	0.098%	0.098%	0.036%	0.164%
	3	0.173%	0.171%	0.169%	0.241%	0.183%	0.181%	0.172%	0.202%	0.189%	0.191%	0.178%	0.153%
	4	0.311%	0.289%	0.273%	0.171%	0.320%	0.299%	0.305%	0.257%	0.330%	0.299%	0.297%	0.257%
4	1	-0.042%	-0.018%	-0.012%	0.062%	-0.068%	-0.025%	0.018%	0.105%	-0.086%	-0.036%	0.008%	0.062%
	2	0.131%	0.130%	0.150%	0.261%	0.135%	0.136%	0.078%	0.168%	0.133%	0.131%	0.119%	0.204%
	3	0.238%	0.238%	0.140%	0.260%	0.248%	0.244%	0.183%	0.274%	0.254%	0.252%	0.153%	0.162%
	4	0.412%	0.386%	0.461%	0.361%	0.422%	0.391%	0.460%	0.481%	0.437%	0.405%	0.458%	0.523%
5	1	-0.023%	-0.002%	0.047%	0.140%	-0.051%	-0.013%	-0.002%	0.132%	-0.074%	-0.031%	0.095%	0.166%
	2	0.161%	0.160%	0.063%	0.075%	0.170%	0.171%	0.116%	0.189%	0.166%	0.165%	-0.020%	0.110%
	3	0.283%	0.279%	0.378%	0.501%	0.297%	0.294%	0.277%	0.236%	0.307%	0.308%	0.291%	0.155%
	4	0.498%	0.462%	0.430%	0.412%	0.502%	0.469%	0.528%	0.562%	0.519%	0.477%	0.552%	0.626%
10	1	0.103%	0.140%	0.129%	0.198%	0.072%	0.125%	0.120%	0.221%	0.049%	0.099%	0.163%	0.262%
	2	0.358%	0.357%	0.247%	0.334%	0.368%	0.373%	0.177%	0.221%	0.363%	0.367%	-0.008%	0.127%
	3	0.526%	0.517%	0.689%	0.714%	0.540%	0.537%	0.695%	0.730%	0.553%	0.557%	0.789%	0.779%
	4	0.816%	0.772%	0.738%	0.598%	0.823%	0.783%	0.811%	0.628%	0.838%	0.782%	0.858%	0.679%
15	1	0.148%	0.188%	0.359%	0.318%	0.157%	0.210%	0.356%	0.322%	0.143%	0.201%	0.331%	0.322%
	2	0.529%	0.529%	0.462%	0.290%	0.540%	0.549%	0.331%	0.234%	0.537%	0.540%	0.387%	0.312%
	3	0.787%	0.783%	0.676%	0.922%	0.791%	0.787%	0.827%	0.922%	0.798%	0.805%	0.795%	0.800%
	4	1.212%	1.136%	1.179%	0.871%	1.188%	1.117%	1.162%	0.935%	1.198%	1.128%	1.163%	1.091%
20	1	0.303%	0.329%	0.444%	0.415%	0.287%	0.353%	0.359%	0.399%	0.280%	0.328%	0.320%	0.408%
	2	0.706%	0.716%	0.633%	0.658%	0.722%	0.735%	0.527%	0.523%	0.722%	0.725%	0.501%	0.391%
	3	0.963%	0.963%	0.905%	1.188%	0.970%	0.968%	1.075%	1.257%	0.975%	0.975%	1.169%	1.314%
	4	1.560%	1.426%	1.550%	1.339%	1.553%	1.401%	1.572%	1.413%	1.555%	1.389%	1.542%	1.413%
30	1	0.567%	0.612%	1.092%	0.930%	0.565%	0.614%	0.937%	0.710%	0.549%	0.604%	0.981%	0.857%
	2	1.008%	1.012%	0.893%	0.681%	1.022%	1.034%	0.981%	0.815%	1.029%	1.040%	0.932%	0.677%
	3	1.359%	1.349%	0.992%	0.977%	1.365%	1.357%	0.885%	0.974%	1.376%	1.365%	0.959%	1.056%
	4	2.259%	2.141%	2.216%	2.425%	2.240%	2.122%	2.390%	2.561%	2.239%	2.095%	2.321%	2.414%
40	1	0.845%	0.884%	1.133%	1.146%	0.849%	0.894%	1.116%	1.095%	0.831%	0.871%	1.102%	1.046%
	2	1.407%	1.419%	1.533%	1.375%	1.413%	1.424%	1.514%	1.375%	1.425%	1.426%	1.441%	1.482%
	3	1.785%	1.780%	0.759%	1.205%	1.779%	1.780%	0.771%	1.212%	1.790%	1.792%	0.922%	1.388%
	4	2.860%	2.582%	3.476%	3.623%	2.856%	2.594%	3.500%	3.601%	2.852%	2.597%	3.436%	3.454%
50	1	1.195%	1.310%	1.465%	1.399%	1.156%	1.275%	1.430%	1.557%	1.145%	1.271%	1.461%	1.615%
	2	1.805%	1.812%	1.837%	1.881%	1.802%	1.802%	1.525%	1.449%	1.808%	1.809%	1.250%	1.370%
	3	2.206%	2.198%	1.127%	1.350%	2.223%	2.209%	1.397%	1.319%	2.231%	2.206%	1.772%	2.001%
	4	3.399%	3.090%	4.177%	4.045%	3.424%	3.068%	4.253%	4.187%	3.422%	3.113%	4.123%	4.101%
60	1	1.354%	1.564%	1.683%	1.713%	1.300%	1.490%	1.609%	1.713%	1.293%	1.477%	1.490%	1.666%
	2	2.172%	2.181%	1.950%	1.723%	2.187%	2.189%	1.777%	1.769%	2.189%	2.176%	1.617%	1.657%
	3	2.704%	2.705%	1.938%	2.194%	2.730%	2.727%	2.140%	2.131%	2.732%	2.720%	2.551%	2.614%
	4	4.114%	3.692%	4.776%	4.165%	4.127%	3.686%	4.822%	4.305%	4.129%	3.720%	4.688%	4.027%

Table 6

Percentage of Market Returns above Median Market Return by Signal Strength Quartiles at Various Lead Lengths

Lead Length	Low Bandwidth				Medium Bandwidth				High Bandwidth			
	Quartile				Quartile				Quartile			
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
1	44.9%	44.9%	49.8%	60.1%	43.2%	46.6%	48.8%	61.1%	43.6%	47.0%	48.5%	60.8%
2	43.6%	51.7%	53.0%	51.7%	45.3%	50.3%	51.0%	53.4%	47.3%	50.7%	48.6%	53.4%
3	44.7%	50.7%	53.7%	50.7%	45.4%	50.0%	52.0%	52.4%	46.8%	50.3%	50.0%	52.7%
4	44.1%	50.7%	50.7%	54.6%	45.1%	47.3%	51.0%	56.6%	44.1%	49.3%	49.3%	57.3%
5	44.1%	45.4%	55.4%	54.9%	43.1%	48.8%	49.7%	58.3%	46.4%	46.1%	47.3%	60.0%
6	43.1%	48.5%	53.6%	54.9%	43.4%	50.2%	49.5%	56.9%	45.1%	47.5%	50.5%	56.9%
7	45.6%	46.1%	52.2%	55.9%	43.5%	49.5%	49.2%	57.6%	44.2%	47.8%	50.5%	57.3%
8	45.2%	46.1%	53.6%	55.1%	43.2%	48.1%	52.5%	56.1%	42.9%	45.4%	55.9%	55.8%
9	45.2%	45.2%	54.6%	54.8%	44.6%	45.2%	54.2%	55.8%	43.2%	45.6%	54.9%	56.1%
10	43.2%	48.0%	55.1%	53.7%	44.6%	45.2%	55.1%	55.1%	44.9%	43.2%	55.8%	56.1%
15	43.5%	45.4%	54.9%	56.0%	43.5%	44.0%	56.3%	56.0%	43.2%	46.1%	53.9%	56.7%
20	41.6%	48.6%	55.5%	54.3%	40.9%	48.3%	55.8%	55.0%	41.9%	46.6%	57.2%	54.3%
25	42.8%	44.5%	55.3%	57.2%	41.0%	45.5%	55.7%	57.6%	43.1%	43.8%	54.3%	58.6%
30	46.7%	42.9%	47.8%	62.6%	43.3%	46.7%	46.4%	63.7%	45.0%	44.6%	47.4%	63.0%
35	43.9%	48.6%	45.1%	62.2%	42.5%	46.9%	47.6%	62.8%	41.5%	44.8%	50.7%	62.8%
40	43.0%	47.4%	44.3%	65.4%	42.0%	47.4%	44.9%	65.7%	41.3%	48.1%	45.6%	65.0%
45	40.0%	51.9%	45.8%	62.1%	43.5%	43.9%	49.3%	63.2%	43.5%	44.6%	48.3%	63.5%
50	41.2%	48.2%	46.8%	63.7%	43.3%	45.4%	47.2%	64.1%	43.0%	44.4%	50.0%	62.7%
55	38.3%	48.8%	51.9%	60.8%	39.7%	45.9%	51.6%	62.5%	38.7%	43.8%	57.2%	60.1%
60	44.1%	45.0%	48.6%	62.3%	45.2%	44.7%	47.9%	62.3%	44.1%	43.3%	51.1%	61.6%
65	43.2%	51.1%	43.8%	61.8%	41.8%	52.1%	44.8%	61.1%	41.1%	51.4%	47.7%	59.6%
70	38.4%	54.1%	43.7%	63.8%	37.3%	52.7%	47.0%	63.1%	38.0%	51.6%	49.8%	60.6%

Figure 1—A
Regression Coefficients for Put Call Parity Signal, from
Multi—variate Regression of Lead Returns on 4 Aggregate Variables

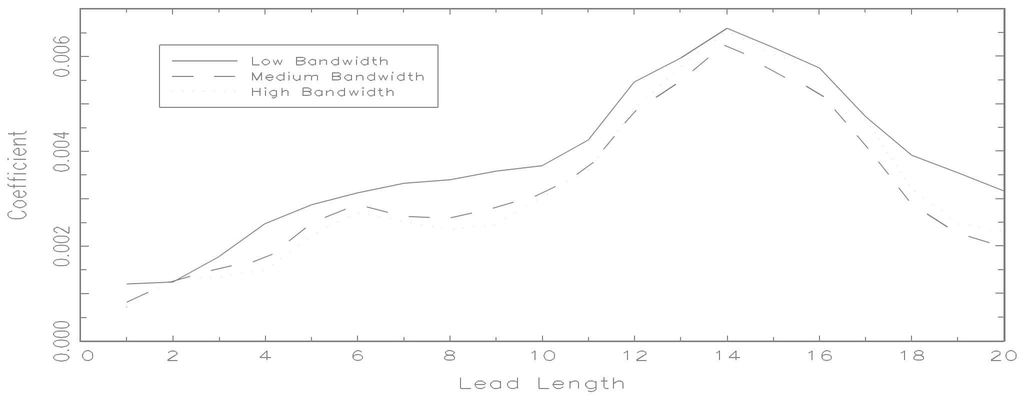


Figure 1—B
t—stats for Put Call Parity Signal Coefficient from
Multi—variate Regression of Lead Returns on 4 Aggregate Variables

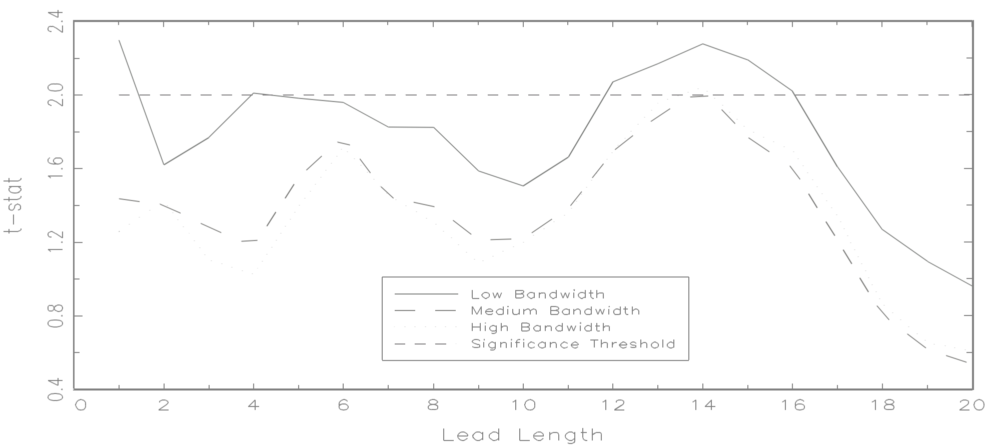


Figure 2–A
Regression Coefficients for Put–Call–Parity Spread from
Multivariate Regression of Lead Returns on 4 Aggregate Variables

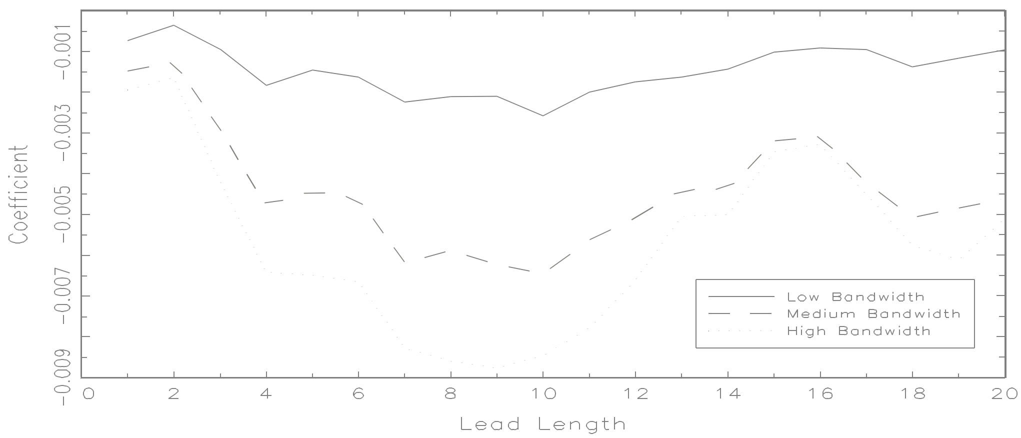
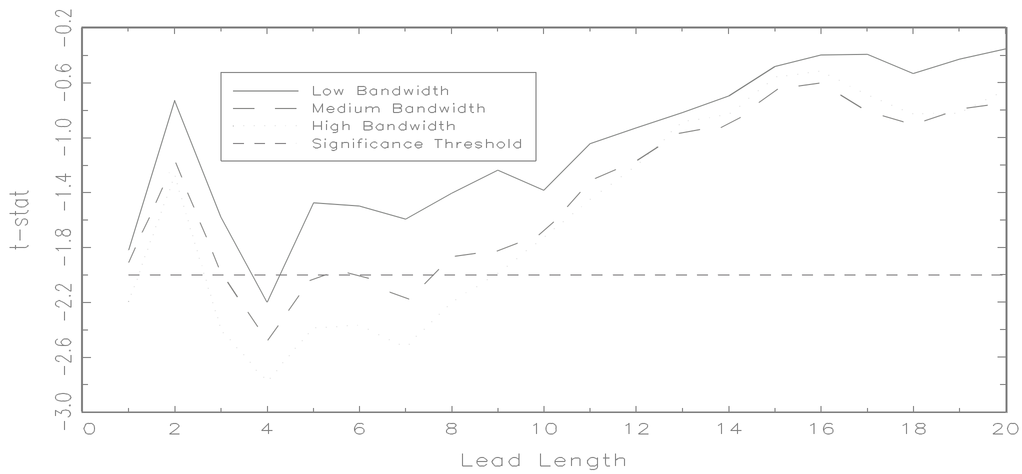


Figure 2–B
t–stats for Put–Call–Parity Spread from
Multi–variate Regression of Lead Returns on 4 Aggregate Variables



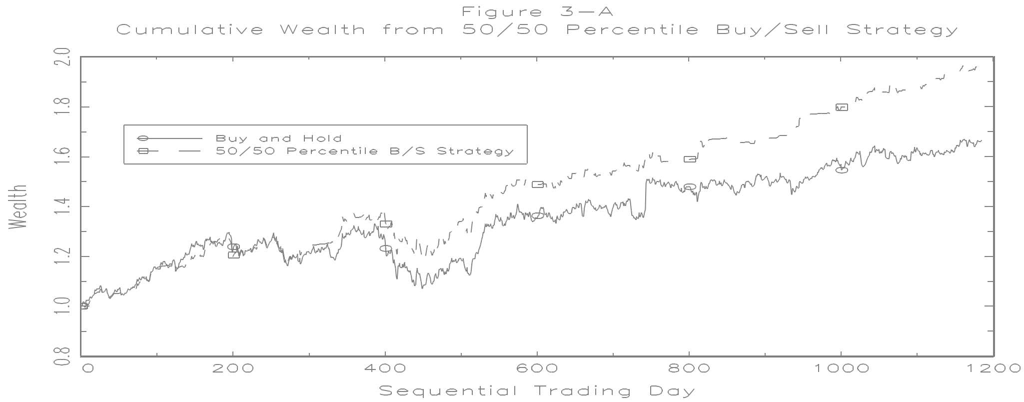


Figure 3—A
Cumulative Wealth from 50/50 Percentile Buy/Sell Strategy
Buy and Hold
50/50 Percentile B/S Strategy
Wealth
Sequential Trading Day

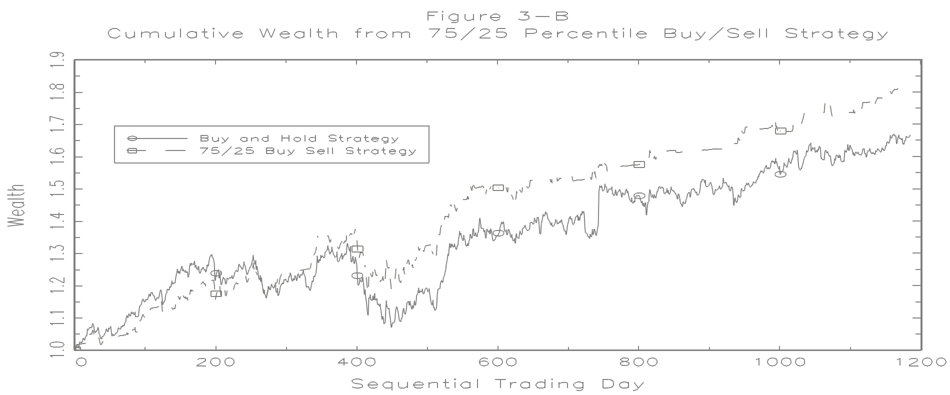


Figure 3—B
Cumulative Wealth from 75/25 Percentile Buy/Sell Strategy
Buy and Hold Strategy
75/25 Buy Sell Strategy
Wealth
Sequential Trading Day

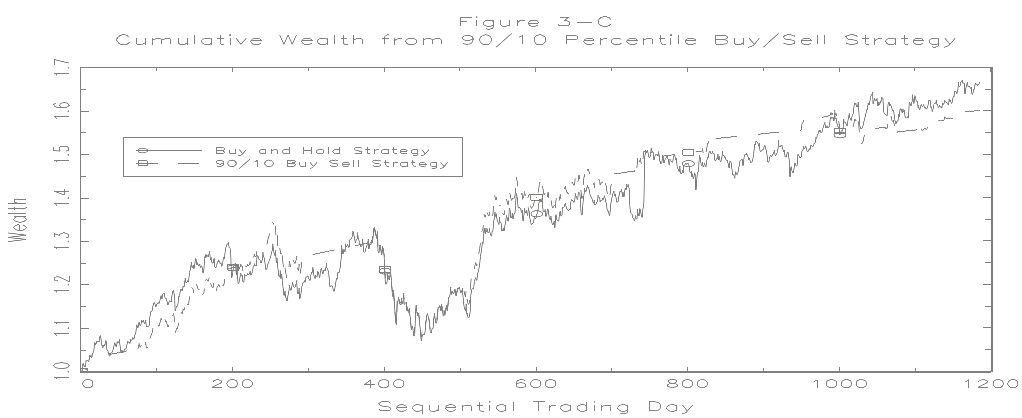


Figure 3—C
Cumulative Wealth from 90/10 Percentile Buy/Sell Strategy
Buy and Hold Strategy
90/10 Buy Sell Strategy
Wealth
Sequential Trading Day

Figure 3—D
Cumulative Wealth from 95/5 Percentile Buy/Sell Strategy

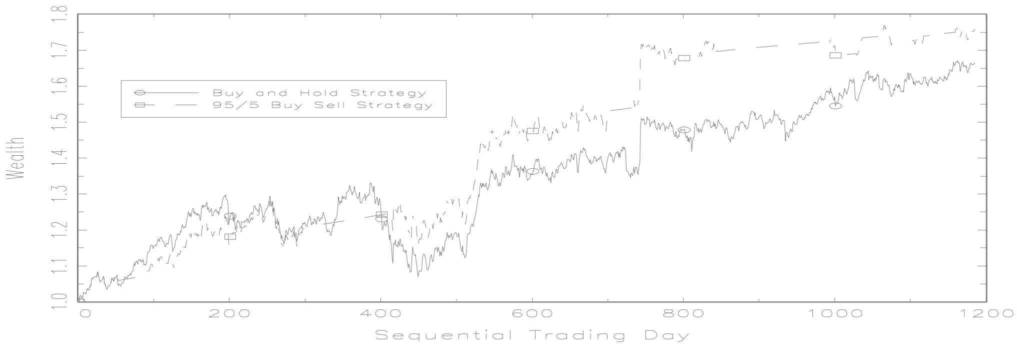


Figure 3—E
Cumulative Wealth from 25/25 Percentile Buy/Sell Strategy

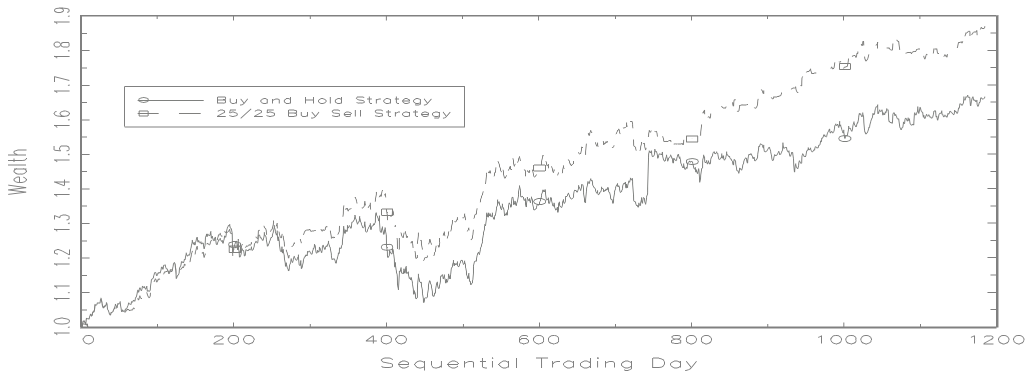
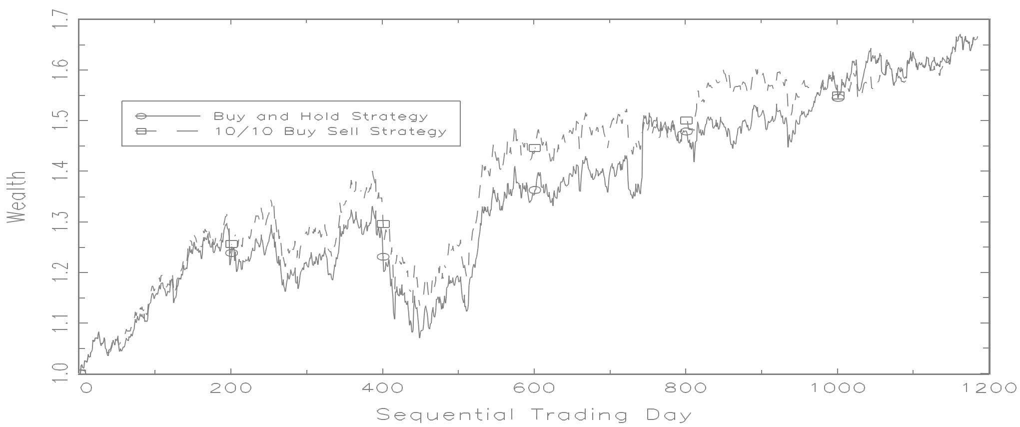


Figure 3—F
Cumulative Wealth from 10/10 Percentile Buy/Sell Strategy



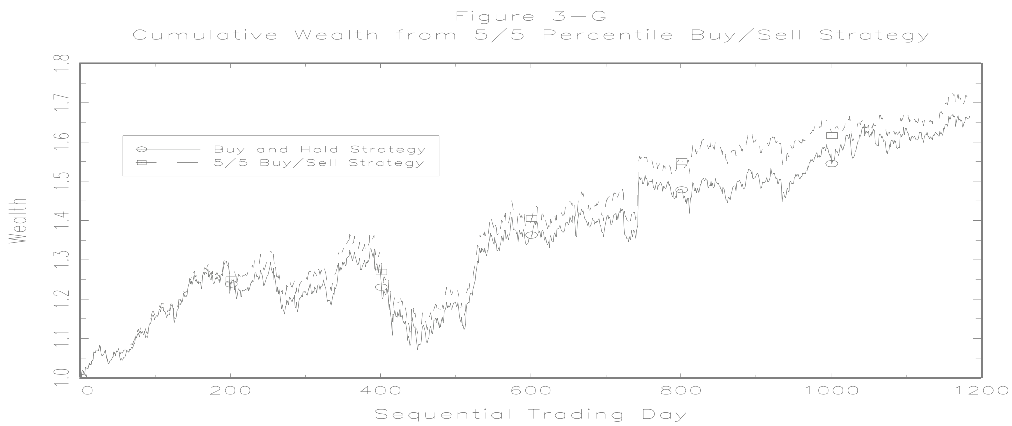


Figure 3-G
Cumulative Wealth from 5/5 Percentile Buy/Sell Strategy
Buy and Hold Strategy
5/5 Buy/Sell Strategy
Wealth
Sequential Trading Day
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
0
200
400
600
800
1000
1200