Informed Trading and Option Spreads *

Gautam Kaul
Michigan Business School
Ann Arbor, MI 48109
Email: kaul@umich.edu
PH. (734) 764-9594.

M. Nimalendran
Warrington College of Business
University of Florida
Gainesville, FL 32611.
Email: nimal@dale.cba.ufl.edu
PH. (352) 392-9526.

Donghang Zhang
Moore School of Business
University of South Carolina
Columbia, SC 29208
Email: zhang@moore.sc.edu
Ph. (803) 777-0242

.

Current Draft: April 2004

* We wish to thank participants at the Market Microstructure Conference sponsored by the *Journal of Financial Markets* and Yale International Center for Finance for their comments and suggestions. We also wish to thank the seminar participants at University of Houston and Virginia Tech for helpful comments.

Informed Trading and Option Spreads

Abstract

We assess the presence and nature of strategic trading by informed investors in the options market. Specifically, we develop and test a model for the spread of an option that directly captures the effects of strategic trading by informed traders. We show that the underlying stock's spread has an important impact on the option spreads due to the hedging activities of option market makers. The initial hedging costs explain half the effective spread of at-the-money or in-the-money options. For out-of-the-money options, initial hedging costs explain less than one third of the spread, but nevertheless play an important economic role. Rebalancing costs associated with hedging are much smaller than the theoretical values, however. This suggests that although option dealers hedge their positions, they do not hold their positions for long. We also find that the adverse selection component of the underlying stock's spread explains a significant fraction of the option spread. More importantly, contrary to conventional wisdom, adverse selection costs are higher for (the most actively traded) at-the-money or slightly out-of-the-money contracts relative to out-of-the money options. The results of the array of tests conducted in this paper, taken together, suggest that informed traders trade strategically in options markets, recognizing the trade-off between leverage and transactions costs associated with option contracts of different moneyness.

1. Introduction

The objective of this paper is to test for the presence and nature of strategic trading by informed investors in the options market. To this end, we develop and test a model for the spread in the options market. Black and Scholes (1973) show that in a "perfect" market the payoff to an option can be replicated by continuously rebalancing a portfolio of stocks and bonds. If the conditions necessary for a “perfect” market hold, then option spreads should only compensate market makers for order processing costs. However, when there are market frictions such as transaction costs, it is no longer possible to replicate the option payoff using a dynamic strategy involving continuous rebalancing. Therefore, option market makers must be compensated for the costs associated with rebalancing at discrete time intervals.

Unlike in stock market making, hedging costs play an important role in determining option spreads. These costs consist of the cost of setting up and liquidating the initial delta neutral position, and the costs of rebalancing the portfolio to maintain the delta neutral position. Several papers, including Leland (1985), Merton (1990), and Boyle and Vorst (1992), theoretically examine the impact of stock bid-ask spreads on the hedging costs imposed on option dealers due to discrete rebalancing. They show that the option spread (the difference between the prices of long and short calls) due to discrete rebalancing is positively related to the proportional spread on the underlying asset, inversely related to the revision interval, and positively related to the sensitivity of the option to changes in volatility (vega).

The microstructure of the options market, however, will also be affected by the strategic behavior of informed agents. Most adverse selection models of bid-ask spreads assume that informed agents can trade only in a single market.[1] If informed agents could trade simultaneously in both the stock and option markets, however, their behavior would impact bid-ask spreads on both the stock and the option markets. Adverse selection costs play an important role in determining stock spreads;[2] but, in the options market, the extant evidence is mixed. If informed agents trade strategically using stocks and options to maximize their returns from

[1] Easley, O’Hara, and Srinivas (1998) investigate the informational role of transaction volume in the option market through an asymmetric information model in which informed traders may simultaneously trade in the option and equity markets. They focus, however, on the price discovery process, and not the spreads, in the two markets.

[2] The bid-ask spreads on stocks compensate market makers for order processing, inventory [Garman (1976), Stoll (1978), Ho and Stoll (1981)], and adverse selection costs [Bagehot (1971), Glosten and Milgrom (1985), Kyle (1985), Easley and O'Hara (1987)].

private information, and if option market makers cannot instantaneously hedge the exposure of their positions to adverse selection, then they will face the same information disadvantage as the market makers in the underlying stocks. In this study, we consequently examine the relation between option spreads and proxies for adverse selection costs in the underlying stocks.

The advantage of explicitly linking the adverse selection costs in the stock and option markets is that it sheds light on the price discovery process in the two markets. Market(s) where informed agents trade could have important implications for the price discovery process. As Easley, O'Hara, and Srinivas (1998) point out: "If there are alternative markets in which informed traders can profit from their information, then where informed agents choose to trade may have important implications not only for security price movements, but for the behavior of related prices as well." (p. 431) Black (1975) argues that informed agents might prefer the options market for its high leverage, but Easley, O'Hara, and Srinivas (1998) suggest that informed agents may trade in both the option and the stock markets simultaneously. The empirical evidence on this issue is mixed, however. For example, Vijh (1990) and Cho and Engle (1999) find that option market makers do not face significant adverse selection costs, while Easley, O'Hara, and Srinivas (1998) and Cao, Chen, and Griffin (2003) find evidence consistent with informed trading on the options market. We attempt to shed light on this specific issue by explicitly modeling and estimating the link between option spreads and the adverse selection component of the spread of the underlying stock.

Another important issue that is not addressed by previous researchers is that the options market, relative to the stock market, provides strategic flexibility to informed agents because they can trade contracts on the same underlying asset but with different maturities and exercise prices. Following Black (1975), the conventional wisdom is that informed agents would trade out-of-the-money option contracts for the high leverage. High transaction costs and lack of liquidity, however, may offset the benefits of the high leverage provided by these contracts. This aspect of the options market may have important implications for the strategic trading of informed investors. Again, by explicitly examining how option spreads of different option contracts are affected by adverse selection, we infer the strategic behavior of informed trading in options markets.

To empirically evaluate the importance of the different determinants of option spreads, we use more than two million observations on options traded on 573 stocks during February 1995 on the Chicago Board of Options Exchange (CBOE). We find that the initial (delta neutral) hedging costs explain two percent of the option spread for out-of-the-money options, and approximately 64 percent of the spreads of deep-in-the-money options.[3] Surprisingly, however, discrete rebalancing costs incurred to maintain the delta neutral position accounts for only 6.9% of the option spread, which is much smaller than the theoretical values predicted by Leland (1985) and Boyle and Vorst (1992). This evidence suggests that option market makers hedge their positions, but they do not have to incur large rebalancing costs either because they may not hold their positions very long or they can effectively diversify their portfolios.

The adverse selection component of the stock spread explains six to 22 percent of the spreads for options contracts with different maturities and moneyness. This evidence (a) shows that option market makers do face significant adverse selection costs because informed agents trade on the options market, and (b) indicates that the option market makers are not able to completely hedge the risk arising from trading with informed agents. Our results are consistent with the findings of Easley, O'Hara, and Srinivas (1998) and Cao, Chen, and Griffin (2004).

Our analysis also shows an interesting pattern in the behavior of informed trading in options markets. It appears that informed traders choose to camouflage their trading by participating in the more liquid contracts that have high levels of noise/liquidity trading. Specifically, for actively traded options with maturity less than 60 days, we find that at-the-money and slightly out-of-of-the-money contracts have much higher adverse selection costs relative to out-of-the-money option contracts. For the overall sample, adverse selection contributes an average of 4.41 cents, or 34.26%, to the option spread for at-the-money and slightly out-of-the-money options. For the out-of-money options, however, adverse selection accounts for only 1.29 cents, or 10.05%, of the option spread. This finding is in contrast to the conventional wisdom that informed agents prefer out-of-the-money contracts for their inherently higher leverage. This evidence suggests that informed agents face a trade-off between the higher leverage of out-of-the-money options and the higher transaction costs and lower liquidity associated with them. It is possible, therefore, that the optimal contracts for informed agents are

[3] Consistent with option market makers' hedging behavior, Mayhew and Mihov (2004) find that the listing of an option increases the underlying stock's trading volume.

at-the-money and slightly out-of-the-money contracts that have the highest delta to option price ratios (and not the out-of-the-money contracts). These options have the highest sensitivity to stock price changes, and also the highest trading volume which, presumably, makes it easier for the informed agents to hide their identity.

Estimates of our model for the spreads in options markets have some ancilliary implications. Specifically, we find that trading volume has a significant negative impact on option spreads. For example, an increase in volume by 439 contracts (the average volume) will decrease spreads by four percent, which is consistent with larger volume having lower order processing costs. Also, the measure for volatility skew is statistically significant but has very little economic effect, which suggests that option dealers adjust the price instead of the spread to control for leverage effects.

Our work is related to the research on the microstucture of the options market. Jameson and Wilhelm (1992) examine the effects of an option's gamma and vega on option spreads using a sample of the most active option contracts on 40 large underlying stocks. They argue that the option gamma is a proxy for the error in delta hedging, and vega captures the uncertainty in volatility that cannot be hedged. In our paper, the option’s vega impacts its spread via the hedging behavior of market makers. In other related literature, researchers have examined the information linkage between the stock and options markets by exploring any lead-lag relations between the two markets. The empirical results from studies by Manaster and Rendleman (1982), Vijh (1988), Stephan, and Whaley (1990), and Diltz and Kim (1996) are inconclusive. Our paper sheds more light on this issue by gauging the adverse selection faced by option market makers. In related work, Mayhew (2002) finds that competition due to cross-listing helps to reduce option spreads. He also finds that different market structures, namely open outcry and DPM (Designated Primary Market Maker) that co-exist on the Chicago Board of Options Exchange (CBOE), affect quoted spreads but not effective spreads. [See also Battalio, Hatch, and Jennings (2004).] Regulations of options trading are evolving, and these studies provide insight on how these institutional arrangements may affect option spreads and information flow across markets.

The rest of the paper is organized as follows. In Section 2, we develop an empirical model of the determinants of option spreads. The differential effects of informed trading on option spreads of contracts with different moneyness are also modeled. Section 3 provides a description of the data, and the results and their analysis are provided in Section 4. Section 5 contains a brief summary and conclusion.

2. Determinants of Option Bid-Ask Spreads

The voluminous literature on the microstructure of financial markets suggests that the bid-ask spread for a stock consists of three components: order processing, adverse selection, and inventory costs. While it is reasonable to assume that option market makers should be compensated similarly, market-making in options markets have some unique features, such as the ability of the market maker to hedge her position using the underlying stock. In this section, we build and test a model for option spreads that is quite general, and which incorporates the unique features of options markets.

2.1 Hedging Costs

Initial Hedging Cost

To hedge her position, an option market maker would set up a delta neutral position by purchasing Δ shares of the stock at the ask price and close the position by selling the shares at the bid price. This would lead to a cost,

$$IC = kS\Delta, \tag{1}$$

where IC represents the initial hedging cost, k is the proportional bid-ask spread of the underlying stock, S is the stock price, and Δ is the option delta.

Rebalancing Cost

The initial hedging cost, however, does not include the cost to the market maker of rebalancing her portfolio to maintain a delta-neutral position. Leland (1985) and Boyle and Vorst (1992) provide a method to estimate the expected rebalancing cost in the presence of proportional transaction costs for the underlying asset. Following this literature, we define the rebalancing cost as follows:

$$RC = 2vk/\sqrt{2\pi(\delta t)}, \tag{2}$$

where RC denotes the rebalancing cost, v is the option vega, k is the proportional spread of the underlying stock, and δt is the rebalancing interval.

The rebalancing cost is directly proportional to the option's vega and the spread of the underlying stock, and is inversely related to the rebalancing interval. The expression for the rebalancing cost also has an intuitive explanation: the bid-ask spread of the stock gives rise to additional volatility when the option is replicated. Consider the replication of a long call option. When the stock price increases, rebalancing would require you to purchase more stock. But this has to be done at the ask price. Similarly, when the stock price falls, the stock has to be sold at the bid price to maintain a delta neutral position. This effectively increases the volatility of the asset, and this increase in volatility would be proportional to the bid-ask spread [see Roll (1984)]. Also, since vega is highest when the stock price is equal to the present value of the exercise price, *ceteris paribus*, we would expect at-the-money options to have the highest rebalancing costs.

Since we cannot observe the rebalancing frequency δt, we assume that it is the same across all option contracts. Consequently, the expression for rebalancing becomes

$$RC = \nu k. \tag{3}$$

2.2 Adverse Selection Cost Proxies

Several models have been suggested to estimate the adverse selection cost faced by stock market makers. We use estimates from two models to proxy for adverse selection costs, denoted by *AS*, in the options market (see Appendix A for details). If informed traders behave strategically, and trade both in the underlying stock as well as in the option(s) on the stock, then the adverse-selection faced by the stock and options market makers would be related. We therefore assume that the adverse selection component for a stock would be a good proxy for the adverse selection cost in the corresponding options market.

2.3 Order Processing Costs

Since order processing costs are likely to be fixed for any particular transaction, these costs should decrease as the *expected* trading volume increases. Copeland and Galai (1983) suggest a negative relation between bid-ask spreads and trading volume in the long run, and Easley and O'Hara (1992) develop a formal model that implies that the bid-ask spread will decrease with an increase in expected trading volume of the stock. We consequently use trading volume of the option contract (that is, number of contracts traded), denoted as *TV*, to proxy for order processing costs. Since we control for adverse selection, we expect the trading volume to be negatively related to spreads.[4]

2.4 Volatility Skew (Model Misspecification)

The empirical literature suggests that the implied volatility estimated from equity options is skewed with respect to the exercise price. This volatility skew indicates that in practice traders make adjustments beyond what an option pricing model can capture. This, in turn, implies that we will under- (over-)estimate hedging costs when volatility skew is high (low). To capture the volatility skew effect in our empirical model, we include a variable that is positively related to it. Namely,

$$VS = \frac{(S - X)}{X}, \tag{4}$$

where VS denotes the volatility skew variable, S is the stock price, and X is the exercise price of the option on the stock.

2.5 A Model of Option Spread

Based on the preceding discussion, we propose the following empirical model for option spreads:[5]

$$OS = \beta_0 + \beta_1(IC) + \beta_2(RC) + \beta_3(AS) + \beta_4(TV) + \beta_5(VS) + \varepsilon, \tag{5}$$

where OS is the dollar effective option spread, ε is the error term, and all other variables are defined above.

Before we empirically evaluate the relative importance of the determinants of bid-ask spreads in options markets, it is appropriate to modify (5) to allow for any inherent preference(s) informed traders may have for the type of option contracts they trade. The conventional wisdom is that informed agents should trade out-of-the-money option contracts for their higher leverage [see, for example, Black (1975)]. This, however, may not always be the case, as there is a

---

[4] Large trading volume may also reduce the need to rebalance the hedging position since it is more likely that different transactions may offset each other when trading volume is large. This phenomenon would also lead to a negative relation between option spreads and trading volume.

[5] We do not include the inventory costs as a determinant for two reasons. First, the literature on stock spreads suggests that its magnitude is small [Stoll (1989) and Smidt (1991)]. Second, option market makers rarely take directional risks; even if they carry inventory, it is likely to be hedged.

tradeoff between the higher leverage of out-of-the-money options and their higher transaction costs and lower liquidity. It is therefore conceivable that informed agents may choose to refrain from trading out-of-the-money contracts.

To investigate where informed agents are more likely to trade, we employ dummies based on an option contract's moneyness. The dummies capture how deep an option contract is in the money, and are defined differently for call and put options. We group option contracts based on moneyness into three categories: in-the-money, at-the-money (or slightly out-of-the-money), and out-of-the-money options. We employ two dummies: $D1 = 1$ only if an option belongs to the second category, and $D2 = 1$ only if an option belongs to the third category. The model specification with the dummy variables becomes,[6]

$$OS = \beta_0 + \beta_1(D1) + \beta_2(D2) + \beta_3(IC) + \beta_4(RC) + \beta_5(AS) + \beta_6(AS * D1) + \beta_7(AS * D2) + \beta_8(TV) + \beta_9(VS) + \varepsilon. \tag{6}$$

3. Data Description

We use the resorted Berkeley Option Database to obtain transaction level data on options. This database includes time-stamped transaction prices, trading volume, and bid/ask quotes on option contracts traded on the CBOE. Transaction level data for stocks are obtained from the Trade and Quote (TAQ) Database, which includes all trades and quotes for stocks traded the on the NYSE, AMEX, and NASDAQ markets. The dividend data are obtained from the Center for Research in Security Prices (CRSP).

We use data for all the options traded in February 1995, and we impose four criteria in selecting our specific sample. First, we only include equity options. We exclude all index options because market makers usually use the futures market to hedge, and the market microstructure is likely to be different. Second, we exclude all LEAPs, which are long-term options with maturities more than one year. Third, we require that data from the Berkeley Options, TAQ, and CRSP Databases are available for both the options and their underlying stocks. We however

[6] A market maker could reduce her risk exposure by holding a portfolio of options of different classes (puts and calls), different exercise prices and/or different maturities. Such behavior would mitigate the risks that she needs to manage through explicit hedging, and this in turn would reduce the hedging costs. We however deal with the hedging costs at an individual contract level, and ignore the effects of any potential diversification by the market maker. First, it is impossible to figure out the actual holdings of each option market maker. Second, as long as there is no systematic cross-sectional difference in portfolio holding for different option contracts, estimates of (6) can still provide a clear picture of the cross-sectional importance of the various factors that affect option spreads.

could not match some option contracts in Berkeley Option Database with TAQ and/or CRSP Databases because of stock ticker symbol changes and problems with the secondary option ticker symbols.[7] The above criteria give us 573 stocks that have options listed on the CBOE. The criteria used to match option symbols to the underlying stock symbols are detailed in Appendix B.

There are total of 182,605 transactions in our sample. In Table 1, Panel A, we report summary statistics on the dollar option spread, the initial hedging cost (*IC*), the rebalancing cost (*RC*), the option trading volume (*TV*), and the volatility skew (*VS*) for the entire sample. In Table 1, and in all our subsequent analyses, we use the effective spreads as our measure of the true cost of transacting for all trades. The effective option spreads are approximately half the quoted spreads, which suggests that a significant number of transactions are executed inside the quoted spread.

In Table 1, Panel B, we report the mean dollar and proportional effective spreads (spreads scaled by the mid-point of bid-ask quotes) sorted by moneyness and maturity, the two most important features that characterize an option contract. Specifically, to gauge how option spreads and the factors that determine them vary with moneyness and maturity, we categorize all option contracts into 35 (7*5) groups based on moneyness and maturity. Define the moneyness of an option contract as $\frac{S-X}{X}*100\%$ (for call) or $\frac{X-S}{X}*100\%$ (for put), where S is the stock price and X is the exercise price. The cut-offs for the seven moneyness groups are 50%, 30%, 10%, -10%, -30% and -50%, while those for the five maturity groups are 30, 60, 90 and 180 days. The option spreads are positively correlated with moneyness. In particular, the spreads are generally decreasing with increases in moneyness.[8] This is consistent with spreads being related to delta hedging costs.

However, the average spread on the underlying stocks is only 17 cents (see Table 2) compared to a spread of 21 to 29 cents for deep in-the-money options (the first column in Table 1, Panel B). Since the delta-hedging cost is equal to the stock spread times the delta of the option, we would expect this cost to be zero for a well out-of-the-money option, and 17 cents for

---

[7] We further exclude two underlying stocks because of lack of data to obtain reasonable estimates of some stock market microstructure variables.

[8] For out-of-the-money options of most maturities, the far out-of-the-money options have slightly larger spreads compared to groups that are 30 – 50% out-of-the-money.

a well in-the-money option. Therefore, option spreads are much higher than the magnitudes we would expect due to delta hedging alone. The discrepancy is quite significant for out-of-the-money options. For example, when the option is more than 50 percent out-of -the money, we still have a spread of 8 to 15 cents (the last column in Panel B of Table 1) even though the delta ranges from 0.0149 to 0.1265. This clearly indicates that delta hedging alone cannot completely explain the effective option spreads.

Table 1, Panel B, shows no discernible relation between option spreads and maturity, though the spread increases monotonically with maturity for at-the-money options and the vega increases from 1.77 (<30 days) to 9.19 (180-270 days). This suggests that rebalancing costs could be a significant factor for determining the spread of at-the-money options. More interestingly, the proportional effective spread increases from less than 5% for well in-the-money options with maturity less than 60 days (the most actively traded contracts as indicated by trading volume) to more than 80% for out-of-the-money options! This large increase is clearly driven by the very small option price, and is also affected by price discreteness.

In Table 1, Panel B, we also report the option 'Greeks,' including delta, vega, and gamma. Since all CBOE contracts are American options, we use the binomial-tree approach and the central difference formula to numerically calculate these first or second derivatives of the option price with respect to different underlying parameters.[9] Estimates of the option 'Greeks' are consistent with what we would expect for the different groups. The delta is increasing with moneyness, and the vega and gamma are highest for at-the-money options. The initial cost to set up a delta hedge (*IC*) is 0.40 to18 cents depending on moneyness and maturity. The rebalancing cost (*RC*) does not include the revision frequency and the constant term ($\frac{2}{\sqrt{2\pi(\delta t)}}$). If we assume that the dealers rebalance every trading day (i.e., $\delta t = \frac{1}{252}$, and $\frac{2}{\sqrt{2\pi(\delta t)}} = 12.67$) and

[9] For example, to estimate the delta of an option, the first derivative of the option price with respect to the underlying stock price, we first use the binomial-tree method to obtain option prices at $S-\delta$ and $S+\delta$, where S is the underlying stock price, and δ is a small increment in the stock price. Denote these two option prices as $f(S-\delta)$ and $f(S+\delta)$, respectively. Then DELTA = $\frac{f(S+\delta)-f(S-\delta)}{2\delta}$. For dividend paying stocks, we use a continuous dividend yield in our binomial-tree calculation. We use the last dividend before our sample period to calculate the dividend yield for each stock, and assume that the dividend, if any, is paid continuously.

hold the contract until maturity, then the rebalancing cost is 4.2 cents for in-the-money options and 37.5 cents for at-the-money option contacts of 60-90 day maturity. If we include the initial cost, the total hedging cost for an at-the-money option is 46 cents, which is much larger than the effective spread of 14.7 cents. Even if we assume that the dealers rebalance weekly, the total hedging cost for an at-the-money option is still 25.4 cents. Therefore, for any reasonable rebalancing frequency, effective spreads will not even cover the hedging cost if the dealers hold the contract till maturity. In practice, option market makers may not hold a position very long, and may diversify by holding a portfolio of assets. Consequently, whether a significant proportion of the option spread is due to hedging costs becomes an empirical issue.

Three months' transaction level data from TAQ (from November 1994 to January 1995) are used to calculate the stock-related variables. In Table 2, we report the relevant summary statistics for the 573 stocks. The mean effective dollar spread for the stocks is 17 cents which, on average, is 1% of the stock price. In contrast, the effective spread for in-the-money options, which would be similar to stocks, is 21 to 29 cents. The option effective spread is therefore about 50% higher than the stock spread. Similarly, the percentage effective spread for in-the-money options is from 2.96% to 4.75% (the first column in Table 1, Panel B), depending on maturity, which is more than three times the spread of an equivalent stock for most maturity groups. In Table 2, we also report the mean adverse selection fraction of the spread for stocks. We use two measures of the adverse selection component: the GKN measure proposed by George, Kaul, and Nimalendran (1991) and the LSB measure proposed by Lin, Sanger, and Booth (1995). Apart from the different methodologies used in the two studies (see Appendix A for details), GKN estimate the adverse selection component of the qouted spread, while LSB estimate the adverse selection component of the effective spread. The GKN measure indicates that 45 percent of the quoted spread of stocks is due to adverse selection.[10] Since options are traded on relatively volatile stocks, we would expect this component to be high. The LSB measure indicates that the adverse selection is only 3.18 percent of the effective spread. Since there are no studies that have compared these two measures, we use both of them in our analysis.

[10] We also use a third measure of the adverse selection component of a stock's spread, proposed by Neal and Wheatley (1998), herefater NW. As described in Appendix A, this measure is a variant of the GKN measure. We do not report the tests based on the NW measure because, for the most part, they are qualtitatively similar to the ones based on the GKN measure.

4. Results and Analysis

Estimates of models (5) and (6) are reported and analyzed in this section. We estimate model (5) using aggregated information for the groups, as well as using the individual contracts.

4.1. Group Regression

We first group option contracts based on moneyness and maturity into 35 (7*5) groups. We then calculate the means of the dependent and independent variables of model (5), for each group and for each underlying stock. We treat call and put options separately, and estimate (5) using 8641 observations.[11] The group regression results are presented in Table 3.

Three versions of the base model (5) are estimated using three different measures for the adverse selection component of the spread of the underlying stock. The estimates in Table 3 show that initial hedging costs have a positive and significant impact on option spreads. The estimated coefficients of *IC* are very robust in that they vary between 0.73 and 0.78, with very large t-statistics. This suggests that for deep-in-the-money options (with delta close to 1); the initial hedging cost is 74% (model I) of the underlying stock spread. Since the average spread on the underlying stocks is 17 cents, the delta hedging cost accounts for 0.74*17 cents = 12.6 cents of the option spreads of 21 to 29 cents for in-the-money options.

We also find that rebalancing costs, *RC*, have a statistically significant coefficient in all the models. The economic significance of the rebalancing costs, however, is very small. For example, for an at-the-money option with 60 – 90 days maturity, the expected cost due to rebalancing is less than one cent.[12] The results for the initial hedging and the rebalancing costs together indicate that market makers in the options market hedge their positions, but do not face significant rebalancing costs.

To examine the effect of adverse selection costs on option spreads, we use three measures of adverse selection. All three measures are based on the underlying stocks' adverse selection

---

[11] Theoretically, we should have 573 (number of underlying stocks) x 35 (number of groups) x 2 = 40,110 observations. Our sample size is reduced dramatically because a lot of option contracts are not traded.

[12] That is, 0.2065 * 0.0295, using the coefficient in model I. Note that the coefficient already captures the constant term, and therefore we do not need to explicitly account for the $\frac{2}{\sqrt{2\pi(\delta t)}}$ factor. We also interact the rebalancing cost variable with option volume (not reported in the table) to test the hypothesis that, if volume is high, option market makers may not have to hold their positions for long. We find that in two of our models this coefficient is negative, but not statistically significant.

costs, which we believe is a good measure of the information asymmetry between informed investors and option market makers. We find that all three measures are highly significant and positive determinants of option spreads.[13] The expected impact on the option spread is the same for all the groups and is equal to 1.78 cents (0.0393*0.4522) based on the GKN measure. This cost is a larger fraction of the spread for at- and out-of-the-money options, whose effective spreads are only 8 to 18 cents.[14]

We find that the order processing costs are inversely related to volume of trade, which is consistent with the findings of Neal (1992); the coefficient estimates are statistically significant in all the models. However, the economic significance of order processing costs is low because the expected cost is less than one cent. Finally, the volatility skew measure is significant and positive, which is consistent with model misspecification. However, its economic impact is again very small.

4.2 Individual Contracts: OLS Regression and Ordered Probit Models

While the group regression results enable us to provide a clear picture of the cross-sectional determinants of option spreads, they suffer from some drawbacks. First, aggregation within groups leads to loss of information. Second, consistent with the findings of previous researchers, we too find that option quotes and transaction prices are clustered, which renders OLS estimation of models (5) and (6) inefficient. Using individual contracts enables us to employ ordered-probit models, in addition to OLS models, to gauge the impact of these drawbacks. Since the results based on the ordered-probit model are qualitatively similar to the OLS evidence, for brevity we only report the OLS results in Table 4.

The results in Tables 3 and 4 are qualitatively similar. The major differences are that the coefficient estimates for the rebalancing cost and transaction volume variables are larger. It appears that without aggregation within groups, the covariance between these variables and the option spreads, especially for the at-the-money groups where these factors may change more dramatically, shows up more readily. These differences, however, do not affect the statistical patterns and interpretations; these factors are all statistically positive and significant in both the

---

[13] Note that the coefficient estimates of GKN and LSB are different because they measure the adverse selection components of the quoted and effective spreads, respectively.

[14] In the next subsection, we provide a more detailed discussion of the effect of information asymmetry on option spreads for contracts with different moneyness and maturity.

group and individual contract regressions.[15] Since 82,605 observations are used in estmating the regressions reported in Table 4, it is not surprising that the t-statistics are much larger than the corresponding ones in Table 3.

Instead of relying solely on the statistical significance of the various determinants of option spreads, we calculate the economic importance of each variable. We report these estimates in Table 5 for the whole sample using the individual contract OLS regression coefficients, Panel A, and for each group using the group regression coefficients, Panel B. We only report estimates for Model I in which the GKN measure of adverse selection is used.

From Table 5, Panel A, we can see that for the entire sample, the initial hedging costs (*IC*) on average contribute about half of the mean option spread (6.33 cents out of 12.87 cents).[16] The adverse selection cost, using the GKN measure, accounts for 1.15 cents (8.95%) of the mean option spread. Rebalancing costs, *RC*, and the transaction volume, *TV*, also play an economically important role; *RC* adds 0.89 cents (or 6.93%) to the option spread, while *TV* reduces the spread by 0.51 cents (or 3.96%). On the other hand, although volatility skew, *VS*, is statistically significant in the regressions, it has a very marginal economic effect on option spreads.

Table 5, Panel B, reports the economic importance of the same determinants sorted by the 35 maturity-moneyness groups.[17] Initial hedging costs (*IC*) are still the most important component of the option spread; for most option contracts they account for more than 40% of the spread. Rebalancing costs (*RC*) add the most, from 1.72% to 6.48%, to the spread for at-the-money contracts and contracts with long maturities, which is consistent with the change in vega. Due to the aggregation of the dependent and independent variables within groups, the transaction volume (*TV*) plays a very marginal role in the group results, reducing the spread by less than 1% for all groups. The volatility skew (*VS*) becomes economically important for deep in-the-money and very out-of-the-money option contracts. This is understandable because the volatility skew, as a measure of model-specification error, should be most severe at both ends of the moneyness

---

[15] Furthermore, the interaction variable between the rebalancing cost and the transaction volume is significant and negative (not reported). Therefore, higher volume leads to lower rebalancing costs. This again is consistent with the hypotheses that market makers do not hold their positions for long, or that they are able to hedge their risk exposure.

[16] The 6.33 cents estimate is calculated as 0.8501*$0.0745. The coefficient, 0.8501, is from Table 4, and the mean value of initial costs, $0.0745, is from Table 1, Panel A. All estimates reported above are calculated using a similar procedure.

[17] The estimates in Panel B are based on the regression coefficients from Table 3, Column 1, and the mean values of the relevant variables are obtained from Table 1, Panel B.

spectrum. This pattern disappears when we study the whole sample because both deep in-the-money and way out-of-the-money option contracts are not heavily traded.

Adverse selection costs, measured by the GKN variable, contribute 1.78 cents to the spread in the group regression. The economic importance of these costs, measured in percentage terms, varies from 6.15% for in-the-money options with 60-90 days maturity to 21.94% for out-of-the-money options with 180-270 days maturity. This pattern suggests that informed traders may trade some option contracts more actively than others. Conversely, however, this variation could simply result from changes in the spreads due to reasons other than information asymmetry. It is therefore interesting to shed further light on the issue of whether informed agents prefer to trade certain types of contracts. In the next subsection, we estimate model (6) to further investigate this issue.

4.3 What Contracts Do Informed Agents Trade?

To estimate model (6), we employ two dummy variables based on an option contract's moneyness. We divide both the group and individual samples into two sub-samples based on maturity, and estimate the regression models separately for each sub-sample. Informed agents should trade option contracts that are sensitive to changes in the underlying stock price, (assuming, of course, that they have inside information about future stock price movements). The change of the option price in response to changes in the stock price is proportional to $\frac{\text{Delta}}{\text{Option Price}}$. Based on the estimates reported in Table 1, Panel B, it is clear that for options with less than 60 days' maturity, this ratio is highest for at-the-money and slightly out-of-the-money contracts (Moneyness Groups 4 and 5). On the other hand, for options with more than 60 days' maturity, the out-of-the-money contracts (Moneyness Groups 6 and 7) have the highest ratios.

Based on the above observation, we divide all observations in both the group and individual contract samples into three categories, and define two dummy variables to distinguish the option contracts' moneyness. The first dummy, *D1*, equals ONE if an observation belongs to groups 4 or 5 on the moneyness scale [that is, the variable *MONEY* belongs to the (-30%, 10%) range], and ZERO otherwise. The second dummy variable, *D2*, equals ONE only if an observation belongs to Moneyness Groups 6 or 7 [that is, *MONEY* is less than or equal to –30%], and ZERO otherwise. For each model specification, we divide the observations into two sub-samples, the ones with maturity less than 60 days and the ones with maturity of 60 – 270 days.

Table 6 contains estimates of the coefficients of the independent variables in model (6). In all regressions, we only use the GKN measure of adverse selection.[18] We are most interested in the coefficients of *GKN, D1*GKN,* and *D2*GKN*. The coefficient of *GKN* is positive in all the regressions over all different sub-samples. In regressions for option contracts with less than 60 days' maturity, the coefficient of *D1*GKN* is positive and consitently larger than the coefficient of *D2*GKN*. This indicates that, contrary to conventional wisdom, informed trading in the most actively traded option contracts is most intense in at-the-money and slightly out-of-the-money contracts. Conversely, however, for options with more than 60 days' maturity, the coeffecient of *D2*GKN* is positive and consistently larger than the coefficient for *D1*GKN*. This suggests that, for option contracts with longer maturities, adverse selection is higher for out-of-the-money contracts relative to in-the-money options.

In Table 7, we report the economic contributions of the adverse selection component to the option spread in different sub-samples based on the OLS regressions. The economic contribution is calculated as the product of the mean GKN measure (reported in Table 2) and the corresponding coefficients in Table 6. For in-the-money options, the coefficient is simply the coefficient for *GKN*; for at-the-money and slightly out-of-the-money options, the coefficient is the sum of the coefficients of *GKN* and *D1*GKN*; and for out-of-the-money options, the coefficient is the sum of the coefficients for *GKN* and *D2*GKN*. For the less than 60 days' maturity sub-sample, the individual contract regressions suggest that the adverse selection contributes 2.24, 4.41, and 1.29 cents to the spreads of in-the-money, at-the-money and slightly out-of-the-money, and out-of-the-money contracts, respectively. Translated into percentages, the economic contributions are 17.39%, 34.26%, and 10.05%, respectively. For the group regression, the estimates are 1.39, 2.04, and 1.76 cents, respectively, ranging in percentage terms from 5.36% to 23.23%.

This evidence suggests that at-the-money and slightly out-of-the-money contracts do face more severe adverse selection, with a nontrivial economic consequence. For example, based on the individual contract regressions, the adverse selection costs for these contracts could be higher by as much as 3.12 cents, which amounts to 24.24% of the option spread. For the sub-sample with more than 60 days' maturity, both the individual contract and group regressions suggest that

[18] We also use the LSB and the NW measures to run the same analyses. While we do not report the results, the LSB measure shows the same (and stronger in the sense of supporting our arguments) pattern, while the NW measure does not show a clear pattern.

the adverse selection component contributes more to the spreads for out-of-the-money contracts, although not by a very significant magnitude.

In summary, and in contrast to conventional wisdom, we find that informed trading is more severe for the most actively traded contracts (which are the ones with less than 60 days' maturity). The informed trade out-of-the-money contracts more often only when their maturity is long. This suggests that informed traders seek liquid contracts, presumably to hide their identity by trading with the crowd.

5. Conclusion

In this paper, we develop and estimate a model for option spreads to gauge the presence and nature of informed trading in the options market. We specifically examine the role of transactions costs and adverse selection costs in determining the spreads of options. Our results and analysis lead to three important conclusions about the behavior of spreads in options markets, and the relation between options trading and the characteristics of the underlying stocks. First, option market makers hedge their positions, and hedging costs constitute a substantial part of option spreads. This also suggests that option market makers' hedging activities may be an important channel for information flow between the stock and options markets. Second, there is strong evidence of informed trading in the options market in that adverse selection costs play an important role in determining option spreads. Third, informed traders appear to trade strategically in options markets, recognizing the trade off between leverage and transactions costs. Specifically, informed traders do not trade the most leveraged out-of-the-money contracts; instead, they trade more actively traded at-the-money and slightly out-of-the-money contracts with less than 60-days to maturity.

Appendix A: Estimation of the Adverse Selection Component of the Spread

A.1 Method I - George, Kaul, and Nimalendran (1991)

Three important assumptions underlie the spread decomposition method of George, Kaul, and Nimalendran (1991) (hereafter, GKN). First, they decompose the spread into only two components (adverse selection and order processing), because they assume that the inventory component is small enough to ignore.[19] Second they assume that the sequence of buy and sell orders is serially uncorrelated: regardless of the most recent order's type, the probability of "buy" and a "sell" on the next order both equal 0.50. Finally, they assume that the quoted spread is constant across transactions.

GKN compute two different return series for each stock--one based on transaction prices and the other based on quote midpoints. Let R_{it}^{T} be the return to stock i at time t, based on transaction prices. Correspondingly, define R_{it*}^{Q} as the return to security i at time t*, based on the midpoint of the bid and ask quotes. The time subscripts on these returns differ because GKN assume that the quotes are updated following each transaction. Hence, t* > t. Next define $R_{it}^{D} = R_{it}^{T} - R_{it*}^{Q}$ as the difference in returns based on the transaction prices and quote midpoints for security i at time t. Finally, let S_i is the quoted spread, and π_i is the fraction of the quoted spread due to order processing costs. [Of course, this makes $(1-\pi_i)$ the fraction due to adverse selection costs.] GKN show that

$$\pi_i S_i = 2\sqrt{-[\mathrm{Cov}(R_{it}^{D}, R_{it-1}^{D})]}\ . \qquad \text{(A-1)}$$

GKN use daily data and end-of-day prices to estimate spread components, while here we use intra-day transaction prices and quotes.

Neal and Wheatley [1998] implement GKN's methodology in a slightly different manner. In particular, they allow the proportional spread to vary through time, and they do not impose the restriction that the probability of a buy or sell is 0.50. Under these conditions, the following regression model can be used to estimate the adverse selection component of the spread:

$$2RD_t = \pi_0 + \pi_1(s_{qt}Q_t - s_{qt-1}Q_{t-1}) + \varepsilon_t \qquad \text{(A-2)}$$

where s_{qt} is the quoted proportional spread at time t, Q_t is a +1/-1 buy/sell indicator variable, RD_t is the difference between the transaction price based return and the quote based return, and ε_t is

[19] Stoll (1989) documents that the inventory cost component is a small fraction of the total spread (less than 10%), and Madhavan and Smidt (1991) find that inventory effects are economically and statistically insignificant.

an error term. The estimate of $(1-\pi_1)$ again measures the fraction of the spread due to adverse selection. We do not report the results based on the NW measure because they are qualitatively similar to those based on the GKN measure.

A.2 Method II - Lin, Sanger and Booth (1995)

Lin, Sanger and Booth (LSB) employ a regression method to estimate the proportion of the *effective* spread that can be attributed to adverse selection. Their approach is based on Stoll (1989) and Huang and Stoll (1994). The main idea underlying this decomposition method is that quote revisions will reflect the adverse selection component of the spread, while changes in transaction prices will reflect order processing costs and bid-ask bounce. As in the GKN model, LSB assume that the market maker's inventory cost is zero. Unlike GKN, however, LSB estimate an order persistence parameter, which measures the probability that a buy (sell) order will be followed by another buy (sell).

Let

P_t = transaction price at time t,

Q_t = quote midpoint,

$Z_t = P_t - Q_t$., one half the effective spread,

λ = proportion of the effective spread due to adverse selection, and

$\delta = (\theta + 1)/2$ = order persistence parameter.

The adverse selection and order persistence parameters are estimated from the following pair of equations:

$$Q_{t+1} - Q_t = \lambda Z_t + \varepsilon_{t+1}, \qquad \text{(A-3)}$$

$$Z_{t+1} = \theta Z_t + \eta_{t+1} \qquad \text{(A-4)}$$

In this model, ε_{t+1} and η_{t+1} are noise terms, while λ measures the fraction of the *effective* spread, which is due to the market maker's adverse selection costs. By contrast, GKN's $(1-\pi_i)$ measures adverse selection costs as a fraction of the *quoted* spread.

Appendix B: Matching Options and Underlying Stock

The Chicago Board of Options Exchange (CBOE) uses a unique three-letter symbol to identify a series of option contracts on the same underlying security. They occasionally introduce a new symbol for option contracts on the same underlying security when the price of that underlying security changes so that 26 characters of the alphabet are not enough to indicate all the contracts with different strike prices. In addition, CBOE also uses different symbols for LEAPs. To identify the underlying securities for each option contract traded on the CBOE, we use the CBOE's list "Historical Equity Options Listings for All Domestic Options Exchanges: April 1973 thru August 1996." This list includes all historical equity options symbols and the name of the companies for which that option was traded. Using the Berkeley Database, we compile a list of all option symbols traded on the CBOE during February 1995. We identify 837 option symbols on the Berkeley List and, by comparing the 837 symbols with the CBOE list; we are able to match 603 stocks that were on both lists. The sample is further reduced to 575 stocks when we match the stock symbols with the TAQ and CRSP databases. We use all available information in the name structure of the CRSP header information and some CBOE historical publications to ensure that our match is correct. Finally, we exclude two of the 575 stocks because we could not obtain estimates of the market microstructure variables. This gives us the final sample of option contracts on 573 underlying stocks.

References

Bagehot, Walter, 1971, "The only game in town," *Financial Analysts Journal* 27, 12-14, 22.

Battalio, Robert, Brian Hatch, and Robert Jennings, 2004, "Toward a national market system for U.S. exchange-listed equity options" *Journal of Finance* forthcoming

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

Black, Fisher, and Myron Scholes, 1973, "The pricing of options and corporate liabilities," *Journal of Political Economy* 81, 637-659.

Boyle, Phelim, and Ton Vorst, 1992, "Option replication in discrete time with transaction costs," *Journal of Finance* 47, 271-293.

Cao, Charles, Zhiwu Chen, and John M. Griffin, 2003, "The informational content of option volume prior to takeovers," *Journal of Business*, forthcoming.

Copeland, T., and D. Galai, 1983, "Information effects and the bid-ask spread," *Journal of Finance* 38, 1457-1469.

Cho, Young-Hye, and Robert F. Engle, 1999, "Modeling the impacts of market activity on bid-ask spreads in the option market," NBER Working Paper # 7331.

Diltz, J., and S. Kim, 1996, "The relationship between stock and option price changes," *The Financial Review* 31, 499-519

Easley, David, and Maureen O'Hara, 1987, "Price, trade size, and information in securities markets," *Journal of Financial Economics* 19, 69-90.

Easley, David, and Maureen O'Hara, 1992, "Adverse selection and large trade volume: Implications for market efficiency," *Journal of Financial and Quantitative Analysis* 27, 185-208.

Easley, David, Maureen O'Hara, and P. S. Srinivas, 1998, "Option volume and stock prices: Evidence on where informed traders trade," *Journal of Finance* 53, 431-465.

Garman, Mark B., 1976, "Market microstructure," *Journal of Financial Economics* 3, 257 -275.

George, Thomas J., Gautam Kaul and M. Nimalendran, 1991, "Estimation of the bid-ask spread and its components: A new approach," *Review of Financial Studies* 4, 623-656.

Glosten, Lawrence, and Paul R. Milgrom, 1985, "Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders," *Journal of Financial Economics* 13, 71-100.

Ho, Thomas, and Hans R. Stoll, 1981, "Optimal dealer pricing under transactions and return uncertainty," *Journal of Financial Economics* 9, 47-73.

Huang, Roger D., and Hans R. Stoll, 1994, "Market microstructure and stock return predictions," *Review of Financial Studies* 7, 179-213.

Jameson, Mel, and William Wilhelm, 1992, "Market making in the options markets and the costs of discrete hedge rebalancing," *Journal of Finance* 47, 765-779.

Kyle, Albert S., 1985, "Continuous auctions and insider trading," *Econometrica* 53, 1315-1335.

Lee, Charles M. C., and Mark J. Ready, 1991, "Inferring trade direction from intraday data," *Journal of Finance* 46, 733-746.

Leland, Hayne E., 1985, "Option pricing and replication with transactions costs," *Journal of Finance* 40, 1283-1301.

Lin, Ji-chai, Gary C. Sanger, and G. Geoffrey Booth, 1995, "Trade size and components of the bid-ask spread," *Review of Financial Studies* 8, 1153-1183.

Madhavan, A., and S. Smidt, 1991, "A Bayesian model of intraday specialist pricing," *Journal of Financial Economics* 30, 99-134.

Manaster, S., and R. J. Rendleman, 1982, "Option prices as predictors of equilibrium stock prices," *Journal of Finance* 37, 1043-1057

Mayhew, Stewart, 2002, "Competition, market structure, and bid-ask spreads in stock option markets," *Journal of Finance* 57, 931-958

Mayhew, Stewart, and Vassil Mihov, 2004, "How do exchnages select stocks for option listing?" *Journal of Finance* 59, 447-472.

Merton, Robert C., 1990, "Continuous Time Finance", *Basil Blackwell Ltd., Oxford*.

Neal, Robert, 1992, "A comparison of transaction costs between competitive market maker and specialist market structures," *Journal of Business* 65, 317-334.

Neal, Robert, and Simon M. Wheatley, 1998, "Adverse selection and bid-ask spreads: Evidence from closed-end funds," *Journal of Financial Markets* 1, 121-149.

Roll, Richard, 1984, "A simple implicit measure of the effective bid-ask spread in an efficient market," *Journal of Finance* 39, 1127-1139.

Stephan, J. A., and R. E. Whaley, 1990, "Intraday price change and trading volume relations in the stock and stock option markets," *Journal of Finance* 45, 191-220

Stoll, Hans R., 1978, "The supply of dealer services in securities markets," *Journal of Finance* 33, 1133-1151.

Stoll, H. R., 1989, "Inferring the components of the bid-ask spread: Theory and empirical tests," *Journal of Finance* 44, 115-134.

Vijh, Anand M., 1988, "Potential biases from using only trade prices of related securities on different exchanges," *Journal of Finance* 43, 1049-1055

Vijh, Anand M., 1990, "Liquidity of the CBOE equity options," *Journal of Finance* 45, 1157-1179.

White, H., 1980, "A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica* 48, 817-838.

Table 1: Summary Statistics on Options

Panel A reports the overall summary statistics for our option sample, and Panel B reports the summary statistics for 35 groups based on moneyness and maturity. Moneyness is defined as,

$$MONEY = \frac{S - X}{X} * 100\% \text{ if the option is a call, and}$$

$$MONEY = \frac{X - S}{X} * 100\% \text{ if the option is a put,}$$

where S is the stock price and X is the strike price. The options are assigned to seven groups based on 'MONEY'. The cut-off points for the 'MONEY' groups are 50%, 30%, 10%, -10%, -30%, and -50%. Within each moneyness group, options are assigned to five groups based on maturity. The cut-offs for the maturity groups are 30, 60, 90, and 180 days. The statistics reported here are based on transactions and their matched quotes (if needed. The matched quote for a transaction is the quote right before that transaction for the exactly same contract specification). The sample consists of 182,605 transactions. For the group statistics, we first aggregate all the transactions (and hence effective spreads) of the call or put options on the same underlying stock within each group. The numbers of observations reported in Panel B are numbers of means of effective spreads after such aggregation, not numbers of individual transactions, for each group.

Panel A: Summary Statistics for the Overall Sample

Variable	Mean	Std. Dev	Min	Max
OS ($)	0.1287	0.1453	0.0000	13.5000
IC ($)	0.0745	0.0509	0.0000	0.5600
RC ($)	0.0176	0.0136	0.0000	0.1671
TV(1000's)	0.4394	0.9565	0.0010	10.4020
VS	-0.0034	0.1109	-0.7438	1.5250

Panel B: Summary Statistics by Group

		'MONEY' Groups						
Maturity	**Variable**	**1(in)**	**2**	**3**	**4(at)**	**5**	**6**	**7(out)**
	OS ($)	0.259	0.232	0.213	0.132	0.088	0.116	0.145
	OS (%)	3.33	4.08	6.50	21.85	48.22	74.14	80.11
	Delta	0.9985	0.9942	0.9535	0.5576	0.1052	0.0228	0.0149
	Gamma	0.0006	0.0059	0.0412	0.2220	0.0800	0.0179	0.0047
	Vega	0.0228	0.0424	0.4059	1.7693	0.9080	0.1935	0.1960
≤30 days	TV(1000's)	0.1021	0.1358	0.2768	1.2503	0.4059	0.1174	0.1076
	RC ($)	0.0002	0.0005	0.0033	0.0110	0.0073	0.0017	0.0020
	IC($)	0.1569	0.1723	0.1684	0.0975	0.0207	0.0039	0.0042
	VS	0.4417	0.2522	0.0571	0.0049	-0.0128	0.1139	0.7050
	Number of Obs.	42	143	536	894	357	44	5
	OS ($)	0.211	0.225	0.211	0.134	0.099	0.087	0.112
	OS (%)	2.96	3.96	5.82	13.47	30.17	51.33	81.01
	Delta	0.9875	0.9690	0.8653	0.4770	0.1875	0.0574	0.0317
	Gamma	0.0048	0.0222	0.0693	0.1401	0.0870	0.0382	0.0150
	Vega	0.3306	0.4212	1.6308	3.6607	2.1752	0.6216	0.8620
30-60 days	TV(1000's)	0.0273	0.0539	0.1312	0.6359	0.2644	0.0914	0.0409
	RC($)	0.0021	0.0040	0.0120	0.0218	0.0169	0.0068	0.0047
	IC($)	0.1806	0.1648	0.1478	0.0831	0.0371	0.0105	0.0067
	VS	0.4422	0.2156	0.0727	-0.0016	-0.0173	0.0146	0.3010
	Number of Obs.	22	91	407	865	463	71	9

(Cont. on the next page)

Table 1 (Cont.)

Maturity	Variable	IN- / AT- / OUT-OF-THE-MONEY						
		1(in)	2	3	4(at)	5	6	7(out)
60-90 days	OS ($)	0.289	0.266	0.221	0.147	0.114	0.089	0.089
	OS (%)	3.30	3.78	5.19	10.49	26.83	46.29	51.17
	Delta	0.9823	0.9485	0.8400	0.4824	0.2253	0.0824	0.0254
	Gamma	0.0052	0.0233	0.0658	0.1086	0.0803	0.0516	0.0104
	Vega	0.4802	0.9661	2.6203	5.3739	3.2631	1.0583	0.6442
	TV(1000's)	0.0237	0.0327	0.0968	0.3884	0.2000	0.1426	0.0731
	RC($)	0.0033	0.0080	0.0176	0.0295	0.0239	0.0113	0.0051
	IC($)	0.1825	0.1825	0.1421	0.0841	0.0432	0.0148	0.0058
	VS	0.4011	0.2062	0.0679	-0.0053	-0.0350	-0.0680	0.3708
	Number of Obs.	25	61	232	565	328	66	8
90-180 days	OS ($)	0.236	0.261	0.238	0.160	0.118	0.093	0.098
	OS (%)	3.57	4.45	5.35	9.06	15.98	29.67	49.47
	Delta	0.9578	0.9067	0.7818	0.4816	0.2760	0.1470	0.0829
	Gamma	0.0209	0.0363	0.0638	0.0887	0.0709	0.0549	0.0316
	Vega	0.9678	1.7443	4.1037	6.7556	5.0146	2.0277	1.4115
	TV(1000's)	0.0368	0.0518	0.1016	0.3464	0.2173	0.1378	0.1185
	RC($)	0.0082	0.0168	0.0281	0.0393	0.0349	0.0217	0.0160
	IC($)	0.1438	0.1596	0.1276	0.0831	0.0504	0.0249	0.0155
	VS	0.5475	0.2220	0.0601	-0.0074	-0.0324	0.0014	0.0380
	Number of Obs.	35	129	398	869	647	133	23
180-270 days	OS ($)	0.280	0.178	0.247	0.180	0.137	0.110	0.081
	OS (%)	4.75	3.61	5.50	7.68	12.00	16.98	27.05
	Delta	0.9004	0.8172	0.7275	0.4895	0.3237	0.2380	0.1265
	Gamma	0.0269	0.0664	0.0635	0.0685	0.0621	0.0603	0.1027
	Vega	1.9469	2.2472	5.4487	9.1895	6.2734	2.9549	0.9782
	TV(1000's)	0.0250	0.0304	0.0488	0.1391	0.1023	0.0753	0.0285
	RC($)	0.0174	0.0245	0.0356	0.0472	0.0419	0.0333	0.0254
	IC($)	0.1253	0.1164	0.1110	0.0814	0.0537	0.0376	0.0236
	VS	0.3525	0.1025	0.0453	-0.0089	-0.0390	-0.0189	0.2479
	Number of Obs.	5	22	178	556	361	45	6

Table 2: Summary Statistics of Underlying Stocks

This table reports summary statistics on 573 underlying stocks for which we have options traded on the CBOE in our sample. The variables reported in this table are defined as follows:

ES: Effective stock spread, defined as twice the absolute value of transaction price minus its matched prevailing quote. A matched prevailing quote for a transaction is defined as the latest quote before the transaction that is at least 5 seconds earlier according to Lee and Ready (1991).

PES: Percentage effective spread, defined as effective dollar spread divided by the middle price of the matched quote.

GKN: Adverse selection component defined in George, Kaul, and Nimalendran (1991).

LSB: Adverse selection component defined in Lin, Sanger, and Booth (1995).

The summary statistics reported here are based on transaction data from the TAQ database over the three-month period from November 1994 to January 1995.

Variable	Mean	Std. Dev.	Minimum	Maximum
ES ($)	0.17	0.08	0.06	0.56
PES (%)	1.00	0.80	0.14	4.55
GKN (Fraction of Spread)	0.4522	0.2167	0	0.8085
LSB (Fraction of Spread)	0.0318	0.0439	0	0.4771

Table 3: Group Regressions of the Model for Option Spreads

The regression model is as follows:

$$OS_{ijkl} = \beta_0 + \beta_1(IC_{ijkl}) + \beta_2(RC_{ijkl}) + \beta_3(AS_j) + \beta_4(TV_{ijkl}) + \beta_5(VS_{ijkl}) + \varepsilon_{ijkl}.$$

In the model, i denotes either call ($i=1$) or put ($i=2$), and j denotes the underlying stock j. The subscripts k and l denote moneyness and maturity groups respectively. The sample consists of 8,641 observations. We first aggregate all the transactions of the call or put options on the same underlying stock within each group. We could only have a maximum of two observations for each stock within a group, if both calls and puts are traded for that group. We report t-statistics in parentheses below the coefficients, adjusted for heteroskedasticity according to White (1980).

	I	II
Intercept	0.0702	0.0838
	(20.72)	(33.30)
Hedging cost		
IC	0.7435	0.7348
	(27.75)	(27.55)
RC	0.2065	0.1767
	(2.49)	(2.14)
Adverse Selection		
GKN	0.0393	
	(8.14)	
LSB		0.1466
		(5.44)
Order Processing		
TV	-0.0020	-0.0022
	(-4.23)	(-4.27)
Model Misspecification		
VS	0.0411	0.0401
	(5.16)	(5.06)
R^2	0.21	0.21

Table 4: Individual Contract Regressions of the Model for Option Spreads

The regression model is as follows:

$$OS_{ijm} = \beta_0 + \beta_1(IC_{ijm}) + \beta_2(RC_{ijm}) + \beta_4(AS_j) + \beta_3(TV_{ijm}) + \beta_5(VS_{ijm}) + \varepsilon_{ijm}.$$

In the model, *i* denotes either call (*i =1*) or put (*i =2*), *j* denotes the underlying stock *j*, and *m* denotes the mth individual transaction. The sample consists of 182,605 transactions. We report t-statistics in parentheses, adjusted for heteroskedasticity according to White (1980).

	I	II
Intercept	0.0502	0.0553
	(52.89)	(67.34)
Hedging cost		
IC	0.8501	0.8390
	(84.37)	(83.55)
RC	0.5064	0.5193
	(16.00)	(16.48)
Adverse Selection		
GKN	0.0346	
	(25.13)	
LSB		0.3577
		(24.64)
Order Processing		
TV	-0.0116	-0.0111
	(-48.31)	(-46.80)
Model Misspecification		
VS	0.0597	0.0608
	(17.21)	(17.52)
R^2	0.12	0.14

Table 5: Contributions of the Determinants of Option Spreads Regression Model I -- the GKN measure for adverse selection

This table reports the economic importance of the independent variables – measured by the dollar amount (in cents), and the corresponding percentage (reported in parentheses), of each independent variable's contribution to the option spread. In this table we only report numbers for Regression Model I in which the GKN measure is used for the adverse selection costs. The dollar amount of each independent variable's contribution is calculated as the product of the corresponding coefficient and the mean value of that variable. For example, the contribution of the initial hedging cost (*IC*) for the whole sample, is 6.33 cents, which is simply the product of the coefficient for *IC*, 0.8501, as reported in Column I of Table 4 and the mean value of *IC* for the whole sample, $0.0745 (7.45 cents), as reported in Panel A of Table 1. Consequently, it contributes 49.21% (6.33 cents divided by 12.87 cents, the mean option spread for the whole sample as reported in Panel A of Table 1) to the option spread.

		Independent Variables				
		IC ¢ (%)	**RC** ¢ (%)	**GKN** ¢ (%)	**TV** ¢ (%)	**VS** ¢ (%)
Panel A **Whole Sample (Individual Contract Regression)**		6.33(49.21)	0.89(6.93)	1.56(12.12)	-0.51(-3.96)	-0.02(-0.16)
Panel B **By Groups (Group Regression)**						
Maturity	**Money Group**					
	1 (In-the-money)	11.67(45.04)	0.00(0.02)	1.78(6.86)	-0.02(-0.08)	1.82(7.01)
	2	12.81(55.22)	0.01(0.04)	1.78(7.66)	-0.03(-0.12)	1.04(4.47)
	3	12.52(58.78)	0.07(0.32)	1.78(8.34)	-0.06(-0.26)	0.23(1.10)
≤30 days	**4 (at-the-money)**	7.25(54.92)	0.23(1.72)	1.78(13.46)	-0.25(-1.89)	0.02(0.15)
	5	1.54(17.49)	0.15(1.71)	1.78(20.19)	-0.08(-0.92)	-0.05(-0.60)
	6	0.29(2.50)	0.04(0.30)	1.78(15.32)	-0.02(-0.20)	0.47(4.04)
	7 (out-of-the-money)	0.31(2.15)	0.04(0.28)	1.78(12.26)	-0.02(-0.15)	2.90(19.98)
	1 (In-the-money)	13.43(63.64)	0.04(0.21)	1.78(8.42)	-0.01(-0.03)	1.82(8.61)
	2	12.25(54.46)	0.08(0.37)	1.78(7.90)	-0.01(-0.05)	0.89(3.94)
	3	10.99(52.08)	0.25(1.17)	1.78(8.42)	-0.03(-0.12)	0.30(1.42)
30-60 days	**4 (at-the-money)**	6.18(46.11)	0.45(3.36)	1.78(13.26)	-0.13(-0.95)	-0.01(-0.05)
	5	2.76(27.86)	0.35(3.53)	1.78(17.95)	-0.05(-0.53)	-0.07(-0.72)
	6	0.78(8.97)	0.14(1.61)	1.78(20.43)	-0.02(-0.21)	0.06(0.69)
	7 (out-of-the-money)	0.50(4.45)	0.10(0.87)	1.78(15.87)	-0.01(-0.07)	1.24(11.05)
	1 (In-the-money)	13.57(46.95)	0.07(0.24)	1.78(6.15)	0.00(-0.02)	1.65(5.70)
	2	13.57(51.01)	0.17(0.62)	1.78(6.68)	-0.01(-0.02)	0.85(3.19)
	3	10.57(47.81)	0.36(1.64)	1.78(8.04)	-0.02(-0.09)	0.28(1.26)
60-90 days	**4 (at-the-money)**	6.25(42.54)	0.61(4.14)	1.78(12.09)	-0.08(-0.53)	-0.02(-0.15)
	5	3.21(28.17)	0.49(4.33)	1.78(15.59)	-0.04(-0.35)	-0.14(-1.26)
	6	1.10(12.36)	0.23(2.62)	1.78(19.97)	-0.03(-0.32)	-0.28(-3.14)
	7 (out-of-the-money)	0.43(4.85)	0.11(1.18)	1.78(19.97)	-0.01(-0.16)	1.52(17.12)

(Cont. on the next page)

Table 5 (Cont.)

		Independent Variables				
		IC ¢ (%)	**RC** ¢ (%)	**GKN** ¢ (%)	**TV** ¢ (%)	**VS** ¢ (%)
Panel B	**(Cont.)**					
By Groups (Group Regression)						
Maturity	**Money Group**					
	1 (In-the-money)	10.69(45.3)	0.17(0.72)	1.78(7.53)	-0.01(-0.03)	2.25(9.53)
	2	11.87(45.46)	0.35(1.33)	1.78(6.81)	-0.01(-0.04)	0.91(3.50)
	3	9.49(39.86)	0.58(2.44)	1.78(7.47)	-0.02(-0.09)	0.25(1.04)
90-180 days	**4 (at-the-money)**	6.18(38.62)	0.81(5.07)	1.78(11.11)	-0.07(-0.43)	-0.03(-0.19)
	5	3.75(31.76)	0.72(6.11)	1.78(15.06)	-0.04(-0.37)	-0.13(-1.13)
	6	1.85(19.91)	0.45(4.82)	1.78(19.11)	-0.03(-0.3)	0.01(0.06)
	7 (out-of-the-money)	1.15(11.76)	0.33(3.37)	1.78(18.13)	-0.02(-0.24)	0.16(1.59)
	1 (In-the-money)	9.32(33.27)	0.36(1.28)	1.78(6.35)	-0.01(-0.02)	1.45(5.17)
	2	8.65(48.62)	0.51(2.84)	1.78(9.98)	-0.01(-0.03)	0.42(2.37)
	3	8.25(33.41)	0.74(2.98)	1.78(7.19)	-0.01(-0.04)	0.19(0.75)
180-270 days	**4 (at-the-money)**	6.05(33.62)	0.97(5.41)	1.78(9.87)	-0.03(-0.15)	-0.04(-0.2)
	5	3.99(29.14)	0.87(6.32)	1.78(12.97)	-0.02(-0.15)	-0.16(-1.17)
	6	2.80(25.41)	0.69(6.25)	1.78(16.16)	-0.02(-0.14)	-0.08(-0.71)
	7 (out-of-the-money)	1.75(21.66)	0.52(6.48)	1.78(21.94)	-0.01(-0.07)	1.02(12.58)

Table 6: What Contracts Do Informed Agents Trade?

This table reports regression results with dummies for different moneyness groups. For the group and individual contract OLS regressions, the model is as follows:

$$OS = \beta_0 + \beta_1(D1) + \beta_2(D2) + \beta_3(IC) + \beta_4(RC) + \beta_5(GKN + \beta_6(GKN * D1) + \beta_7(GKN * D2)$$
$$+ \beta_8(TV) + \beta_9(VS) + \varepsilon.$$

In all regressions, we only use the GKN measure for the adverse selection costs. The subscripts, which are different for the group regression and individual contract regressions, and are the same as in their corresponding regressions without the dummies as reported in Tables 3 and 4, are omitted. For the group OLS and the individual contract OLS regressions, we divide the observations into two sub-samples, the ones with maturity less than 60 days and the ones with maturity of 60 – 270 days. We estimate the regression model separately for the two sub-samples. The first dummy, *D1*, equals ONE if an observation belongs to 4 or 5 of the *Moneyness* groups, that is, this observation's moneyness, *MONEY,* belongs to (-30%, 10%). [We use the variable *MONEY* to decide the value of *D1* for individual contracts.] This dummy equals ZERO otherwise. The second dummy, *D2*, equals ONE only if an observation belongs to 6 or 7 of the *Moneyness* groups, that is, this observation's moneyness, *MONEY,* is less than or equal to –30%. *D2* equals ZERO otherwise. All other variables are the same as in previous corresponding regressions.

	Group Regression		**Individual Contract Regression**	
	< 60 Days Maturity	60 – 270 Days Maturity	< 60 Days Maturity	60 – 270 Days Maturity
Intercept	0.1088	0.1344	0.0852	0.1041
	(11.46)	(10.97)	(17.96)	(14.05)
D1	-0.0269	-0.0608	-0.0246	-0.0366
	(-2.02)	(-4.31)	(-3.13)	(-4.59)
D2	-0.0458	-0.0596	-0.0331	-.0394
	(-5.02)	(-4.86)	(-7.34)	(-5.51)
Hedging cost				
IC	0.5846	0.57899	0.7585	0.9392
	(12.73)	(8.02)	(59.24)	(28.43)
RC	-0.0216	0.3891	0.2507	0.1417
	(-0.11)	(2.18)	(4.22)	(1.81)
Adverse Selection				
GKN	0.0307	0.0395	0.0495	0.0201
	(1.98)	(1.84)	(4.97)	(1.47)
D1*GKN	0.0145	-0.0387	0.0480	0.0066
	(0.32)	(-1.44)	(1.04)	(0.41)
D2*GKN	0.0083	0.0022	-0.0209	0.0143
	(0.51)	(0.10)	(-2.08)	(1.04)
Order Processing				
TV	-0.0007	-0.0051	-0.0088	-0.0439
	(-4.01)	(-5.38)	(-38.36)	(-28.88)
Model Misspecification				
VS	0.0156	0.0220	0.0119	0.0526
	(1.27)	(2.11)	(2.32)	(11.80)
R^2	0.30	0.21	0.13	0.12

Table 7: Economic Contribution of the Adverse Selection Component (GKN Measure) to the Spreads of Option Contracts with Different Moneyness and Maturity

This table reports the economic contributions of the adverse selection costs (the GKN measure) to the option spreads, based on estimates reported in Table 6. The economic contribution is calculated as the product of the mean GKN measure, 0.4522 as reported in Table 2, and the corresponding coefficients in Table 6. For in-the-money options, the coefficient is simply the coefficient for *GKN*. For at-the-money and slightly out-of-the-money options, the coefficient is the sum of the coefficients for *GKN* and *D1*GKN*. For out-of-the-money options, the coefficient is the sum of the coefficients for *GKN* and *D2*GKN*. The percentage is calculated by dividing the dollar amount of the economic contrinution by the corresponding mean option spreads. For the group regression, since each maturity sub-sample has more than two or more moneyness groups that have the same coefficent(s) but different mean spreads, we report the range of the percentage of the spreads due to the adverse selection costs.

Maturity	Moneyness	Group		Individual Contract	
		Cents	Percentage	Percentage	Percentage
< 60 Days	In-the-Money Groups MONEY>10% Group 1,2 & 3	1.39	5.36-6.58	2.24	17.39
	At-the-money and Slightly out-of-the-money Groups -30% < MONEY < 10 % Group 4 & 5	2.04	15.48-23.23	4.41	34.26
	Out-of-the-money Groups Money < -30 % Group 6 & 7	1.76	12.16-20.27	1.29	10.05
60 -270 Days	In-the-Money Groups MONEY>10% Group 1,2 & 3	1.79	6.18-10.03	0.91	7.06
	At-the-money and Slightly out-of-the-money Groups -30% < MONEY < 10 % Group 4 & 5	0.04	0.20 - 0.32	1.21	9.38
	Out-of-the-money Groups Money < -30 % Group 6 & 7	1.89	17.14 - 23.28	1.56	12.09