OFHEO WORKING PAPERS

Working Paper 04-2

Measurement of the Bid-Ask Spread in Equity Option Markets

Michelle A. Danis

January 2004

OFHEO Working Papers are preliminary products circulated to stimulate discussion and critical comment. The analysis and conclusions are those of the authors and do not imply concurrence by other staff of the Office of Federal Housing Enterprise Oversight or its Director. Single copies of the paper will be provided upon request. References to OFHEO Working Papers (other than an acknowledgment by a writer that he or she has had access to such unpublished material) should be cleared with the author to protect the tentative character of these papers.

Measurement of the Bid-Ask Spread in Equity Option Markets*

Michelle A. Danis
Office of Federal Housing Enterprise Oversight
Washington, DC 20552
mdanis@ofheo.gov

Helen Tauchen
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599
tauchen@unc.edu

November 2003

Abstract

This paper develops a joint model of exchange entry and the bid-ask spread in equity option markets. To our knowledge, no other study of spreads in financial markets has incorporated the exchange decision about whether or not to list a security. This allows us to control for the potential endogeneity of the exchanges in a manner consistent with a game theoretic model of entry. Indeed, we find a statistically significant correlation among the unobservables affecting exchange entry and the bid-ask spreads. The correlation between the entry decisions by different exchanges is, however, notably higher than between the bid-ask spread and entry by any one of the exchanges. Further, we find that Ordinary Least Squares regression of the bid-ask spread misstates the effect of specific exchanges. For the 474 markets affected by an alleged conspiracy between the exchanges during the 1990s, we predict a median decline of eight cents in spreads as a result of the additional entry forecast by the model.

*We would like to thank Gary Biglaiser, Jennifer Conrad, David Guilkey, Tom Lutton, and John Stewart for helpful comments. The views expressed in this research are those of the authors and do not represent policies or positions of the Office of Federal Housing Enterprise Oversight or other officers, agencies, or instrumentalities of the United States Government or of the University of North Carolina-Chapel Hill.

1 Introduction

The listing of equity options has a limited history. With few exceptions, options were constrained to trade on a single exchange until a January 1990 SEC rule change allowed all exchanges to list options on any eligible security. The effect of this rule change might not have been fully realized until 1999 for several reasons, the most important of which was the alleged collusive agreement to limit the listing of existing options. Two broad questions emerge from this history. First, how does competition among the exchanges affect the current bid-ask spread on security options? Second, what was the impact of the alleged collusive agreement between the exchanges during the 1990s?

Studies of financial markets typically model the bid-ask spread as dependent upon the competitive structure of the market, under the hypothesis that spreads decline with the number of exchanges offering a security or option. A recognized challenge in estimating such models for the option market is the potential effect of the spread on the attractiveness of offering the option and hence upon the number of exchanges listing it. In this paper, we develop a game theoretic model of the exchanges' listing decisions in order to address this potential endogeneity. Using data from 2003, we estimate a joint model for the Nash equilibrium market outcome and the bid-ask spread.

In brief summary, we find a statistically significant correlation among the unobservables affecting exchange entry and the bid-ask spreads. The correlation between the entry decisions by different exchanges is, however, notably higher than between the bid-ask spread and entry by any one of the exchanges. We then apply the model in order to predict the effect upon the spreads of the alleged collusion during the 1990s. For the 474 securities on which options remained single listed following the SEC's 1990 ruling, we predict the equilibrium entry based on our empirical model. We then use the estimated spread equation to forecast the spreads for the actual and predicted market structures. With the additional entry predicted by the model, we find a median decline in the spreads of eight cents.

The remainder of this paper is organized as follows. Section 2 describes several previous studies that have estimated the bid-ask spread, both in the equity option and other security markets. Section 3 describes the theoretical model used as the basis for the estimation and

Section 4 details the estimation procedure employed in this paper. Section 5 describes the data. Sections 6 and 7 present the results of the estimations as well as a study of the potential impact of the alleged collusive agreement between the exchanges during the 1990s. Section 8 concludes.

2 Previous Literature

There is now an extensive literature on how the competitiveness of financial markets affects the bid-ask spread on equity prices including options. Many of the studies formulate cross-sectional models with the spread as the dependent variable. The competitiveness of the market is generally measured by a binary for whether the securities are multiple listed or by a traditional Herfindahl index, often interacted with volume and other factors that might affect the spread[1].

In their recent theoretical and empirical work on the inventory holding premium, Bollen, Smith and Whaley (2002) provide an excellent overview of the literature on bid-ask spreads in equity markets[2]. Although they do not address option markets specifically, the empirical work on options has followed the same general approach as the work that they review for other equity markets. In addition, they provide new evidence concerning the effect of competition on bid-ask spreads. They estimate Ordinary Least Squares models of the bid-ask spread for NASDAQ-traded stocks during three time periods. In order to measure the competitiveness of the equity markets, they define a modified Herfindahl index equal to $\frac{HI-1/NM}{1-1/NM}$ where HI is the usual Herfindahl index, which is the sum of squared market shares of each market maker, and NM is the number of market makers. As for the usual Herfindahl index, a higher value

[1] Other factors affecting the spread involve order processing costs, inventory holding costs, and adverse selection costs. Order processing costs include all the direct costs of making markets, such as the price of the exchange seat. Inventory holding costs include the costs of providing liquidity, and specifically the opportunity cost of holding the security and compensation for the risk that the price of the security drops. Adverse selection costs stem from the proposition that the market maker's customers may have better information than the market maker and thus undertake trades that are disadvantageous to the market maker.

[2] See especially Table 1.

reflects a more monopolistic market. The estimated coefficient on the modified Herfindahl index is positive in all of their models and generally significant.

A growing number of studies model the bid-ask spread specifically for equity options markets. Neal (1987) was the first. Using Ordinary Least Squares regression, he models the spread as a function of volume, price, volatility, a dummy variable for multiple listing, and an interaction term between volume and the multiple listing dummy. The interaction term is intended to capture possible economies of scale in option listing. Multiple listing in this time period is exogenous, and he finds that the spread is lower for options trading on more than one exchange but only for low-volume options. Danis (1997) updates Neal's study using more recent data and confirms Neal's results.

De Fontnouvelle, Fishe and Harris (2002) examine the bid-ask spread by class for options that became multiple listed after the summer of 1999. They formulate an Ordinary Least Squares model of the spread as a function of a multiple listing dummy, option price, the spread of the underlying stock, the option delta, gamma, volatility, series volume, and volume interacted with the multiple listing dummy. Their regression results imply that spreads fall by about a third following multiple listing, with the effect concentrated primarily on low volume options.

In contrast to most research that has focused on the determinants of the spread and its relation to the competition among exchanges, Wahal (1997) studies the entry and exit of market makers on NASDAQ. He estimates Poisson regression models for the number of market makers with the bid-ask spread as one of the explanatory variables. The coefficients on the spread variables are negative and significant in all specifications. Wahal also models the determinants of market maker entry and exit using ordered probit models for each security. Lagged trading volume, number of trades, volatility, and the bid-ask spread are included as independent variables. The coefficients on the spread variables are positive and significant, indicating that entry is more likely in markets with large spreads. Finally, he estimates models for the change in spread as dependent on binaries for whether market markers have recently exited or entered. The bid-ask spreads are found to decrease following entry and to increase following exit. Further, the magnitudes of the effects are smaller when there are more traders initially in the market.

An inherent difficulty in the use of options data since the mid to late 1990s is that entry is no longer restricted. Hence, the potential profitability, which may depend upon the bid-ask spreads, affects the number of exchanges offering an option. Recognizing this issue, Wang (1999) uses a measure of potential rather than actual competition in her study of bid-ask spreads. To reflect the potential competition she uses a binary for whether or not the option was first traded after 1990 when multiple listing on all options because possible. She estimates an OLS model for the spread as dependent on the "after-1990" binary and other variables, such as volatility, that are often used in such work. As she hypothesized, the estimated coefficient on the binary is negative and significant.

Mayhew (2001) takes a very different approach in that he compares single and multiple listed spreads for CBOE options between 1986 and 1997 using samples matched on the basis of exogenous characteristics such as price, volume, and volatility. Consistent with the rest of the literature, he finds that multiple listed options have smaller spreads. Using the same matched sample methodology, he also compares the spreads on options that were de-listed with the spreads on other options. The spread on options that were de-listed, and thus moved from the multiple to single listing category, increased.[3]

An alternative for addressing the potential endogenity of the market structure is to explicitly model the entry of the agents in a game theoretic setting. Following the work of Bresnahan and Reiss (1991) and Berry (1992), we model the listing of the equity option by an exchange as an entry game. We then apply simulated maximum likelihood methods to estimate the model of entry jointly with an equation for the bid-ask spread, conditional upon entry.

[3]George and Longstaff (1993) address a related issue but do not consider the effect of competition on the bid-ask spread. Specifically, they estimate a two-stage-least-squares model for the bid-ask spread and the trading volume. In order to identify the model, the call price and the usual delta measure of the relation between the option and equity prices are assumed to affect the spread but not the volume. Likewise, the amount by which the option is in- or out-of the money is assumed to affect the volume but not the spread. They find that the spread significantly affects and is affected by the trading volume.

3 Theoretical Model

The exchanges that trade equity options include two major exchanges, the Chicago Board Options Exchange (CBOE) and the American Stock Exchange (AMEX), two historically regional exchanges, the Philadelphia Stock Exchange (PHLX) and the Pacific Stock Exchange (PSE), and an electronic exchange, the International Securities Exchange (ISE).[4] Given the small number of potential competitors, a theoretic model describing competition among them must incorporate the interdependence of entry decisions. We follow Janssen and Rasmusen (2001), who develop a model of competition under uncertainty in which firms compete in prices. The uncertainty stems from the fact that while the number of potential competitors is known, the actual number of firms quoting prices at any given time is not. Using this construct, equilibrium profits decline in the number of firms. This is in contrast to the standard Bertrand models of competition in prices for which two firms are sufficient to achieve the competitive outcome.[5]

Define the formal game by $\Gamma = [J, S, \pi]$. The set of players, denoted J, equals the N potential firms in the market. The players only choice is whether or not to enter, and the strategy set $\{S\}$ of each player is the set {enter, don't enter}. Equilibrium profits for player j depend on its strategy and the strategy of its opponents. The payoffs to player j are

$$\pi_j(s_j, s_{-j}) = \begin{cases} f^j(Z, s_{-j}) & \text{if player } j \text{ chooses } s_j = \text{“enter”} \\ 0 & \text{if player } j \text{ chooses } s_j = \text{“don't enter”} \end{cases}$$

where s_j and s_{-j} denote the strategies of player j and its opponents respectively. The vector Z denotes exogenous factors affecting the profits of a player that enters. As in Janssen and Rasmussen, we assume that a switch by any opponent from “don't enter” to “enter” lowers f^j.

At a Nash equilibrium, each player chooses the best strategy given the strategies of the other players. Formally, a Nash equilibrium strategy combination s has the property that

[4] In the empirical work to follow, we group PHLX, PSE, and ISE together in a category labelled “Regional exchanges”.

[5] The result is well known for the simple Bertrand model. In addition, it is well-established for the bid-ask spreads in Bertrand pricing models of security markets.

$$\pi_j\left(s_j, s_{-j}\right) \geq \pi_j\left(s_j', s_{-j}\right)\ \forall\ s_j' \in S_j.$$

One of the implications of this definition is that a player is out of the market only if entry, given the equilibrium structure, results in negative payoffs to the player. For example, consider the conditions required for the equilibrium to be a monopoly outcome with only player j entering. In order for this to be a Nash equilibrium, the monopoly payoff for player j must be positive; otherwise it would exit. In addition, any other player that entered and created a duopoly outcome along with player j would have a negative payoff.

As is typical for such models, the equilibrium is not unique. The following simple example illustrates clearly why there may be multiple equilibria. The market has only two potential entrants, and payoffs are given by

$$\pi_j(s_j, s_{-j}) = \begin{cases} \delta \cdot D_{-j} + \beta_j & \text{if player } j \text{ chooses } s_j = \text{“enter”} \\ 0 & \text{if player } j \text{ chooses } s_j = \text{“don't enter”} \end{cases}$$

where D_{-j} is a dummy variable equal to one if the other firm has entered the market. The parameter δ is negative, the parameter β_j may be positive or negative. The Nash equilibrium is a monopoly for player j if

$$\pi_j\left(s_j, s_{-j}\right) = \beta_j > 0 \text{ and } \pi_{-j}\left(s_j, s_{-j}\right) = \delta + \beta_{-j} \leq 0$$

If β_1 and β_2 are both between 0 and δ, then there are two equilibria, one with player 1 as a monopolist and the other with player 2 as a monopolist. In order to have a unique Nash equilibrium, additional assumptions are necessary. The two common approaches in the empirical literature are to assume that the market outcome is the Nash equilibrium with the highest profits or to assume that entry occurs sequentially according to the observed order of entry. In our work, we have adopted the first approach.[6] The exchanges that enter and list an option are then assumed to play a Bertrand price game which determines the bid-ask spread. As mentioned above, two exchanges are sufficient to achieve the competitive outcome in the usual model. With the introduction of uncertainty, Janssen and Rasmusen's work offers an explanation for why additional entry could affect the spread.

[6]See Danis (2003) for a more detailed explanation of the identification problem.

4 Estimation Procedure

We rely upon the game theoretic model specified above in order to estimate a joint model for entry and the bid-ask spreads. The empirical specification of latent profits is

$$\pi^*_{iA} = \alpha_A + X_i\beta_A + \theta_{C-A}C_i + \theta_{R-A}R_i + \varepsilon_{iA}$$
$$\pi^*_{iC} = \alpha_C + X_i\beta_C + \theta_{A-C}A_i + \theta_{R-C}R_i + \varepsilon_{iC}$$
$$\pi^*_{iR} = \alpha_R + X_i\beta_R + \theta_{A-R}A_i + \theta_{C-R}C_i + \varepsilon_{iR}$$

where i indexes the security and j indexes the exchange. The subscripts A, C, and R index the AMEX, CBOE, and Regional exchanges. The constants α_{ij} for $j \in \{A, C, R\}$ capture exchange-specific fixed effects. The variable X_i is a vector of market level characteristics, such as the volume of trade and the volatility of security i, which affect the demand for options on the security. Finally, A_i, C_i, and R_i are indicator variables for whether AMEX, CBOE, and Regional respectively offer options on security i. The coefficients on these three dummy variables reflect the extent to which the profitability of an exchange is affected by the entry of others. For example, θ_{C-A} is the effect of CBOE's presence on the profitability of AMEX.

The spread on option i is

$$S_i = \alpha_S + Z_i\beta_S + \theta_{SA}A_i + \theta_{SC}C_i + \theta_{SR}R_i + \varepsilon_{iS}$$

where the spread S_i equals the ask price less the bid price and Z_i is a vector of characteristics of security i and the option on it. The error terms ε_{iA}, ε_{iC}, ε_{iR}, and ε_{iS} are assumed to be jointly normally distributed with mean zero and the variance matrix denoted

$$\begin{bmatrix} 1 & \sigma_{AC} & \sigma_{AR} & \sigma_{AS} \\ \sigma_{AC} & 1 & \sigma_{CR} & \sigma_{CS} \\ \sigma_{AR} & \sigma_{CR} & 1 & \sigma_{RS} \\ \sigma_{AS} & \sigma_{CS} & \sigma_{RS} & \sigma_{SS} \end{bmatrix}.$$

The likelihood function combines the estimation of the latent profit equations and, conditional on entry having occurred, the estimation of the bid-ask spread. For each security, there are eight possible market configurations: (1) no exchanges trading the option, (2)-(4) only AMEX, CBOE, or the regionals, (5)-(7) three combinations with two exchanges trading the option, and (8) all exchanges trading the option. We assume that the error terms in the latent profit equations are known by the exchanges but are unobserved by the econometrician. Thus the error terms affect the market outcome but are not observable in estimation. The resulting likelihood function is similar to a multinomial probit with the addition of the spread equation.

The computation of the likelihood function is simplest for no entry. The probability of this outcome is the probability that even as a monopolist, each of the exchanges has negative profits. The condition may be stated equivalently as $\varepsilon_{ij} < -\pi_{ij}^M$ where π_{ij}^M is the deterministic part of the above latent profit expression for j=A,C,R with only exchange j as a monopoly in trading the option. Let Φ^M denote the joint marginal distribution function for ε_{iA}, ε_{iC}, and ε_{iR} given the joint normal distribution of the four error terms. Since the error terms are normally distributed, the joint marginal is easily determined. The probability of no entry is thus $\Phi^M(-\pi_{iA}^M, -\pi_{iC}^M, -\pi_{iR}^M)$, which we denote as P_i^0.

The contribution to the likelihood function for options that are offered is more involved since the spread is estimated jointly with the probability of entry. Let ϕ^S denote the marginal density function for the error term of the spread equation and $\Phi^C(.|\varepsilon_{iS})$ the distribution of ε_{iA}, ε_{iC}, and ε_{iR}, conditional on ε_{iS} given the joint normal distribution of the four error terms. Consider the market configuration in which "only AMEX" enters. The likelihood contribution for such an observation is $\phi^S(S_i - \alpha_S - Z_i\beta_S - \theta_{SA}C_i - \theta_{SR}R_i)$ times the probability of the error term combinations for which "only AMEX' is the equilibrium outcome. The probability of "only AMEX" is based on the conditional distribution function for ε_{iA}, ε_{iC}, and ε_{iR}.

The definition of the market equilibrium becomes relevant in determining the combinations of the error terms ε_{iA}, ε_{iC}, and ε_{iR} for which only AMEX enters. As described in the previous section, "only AMEX" is an equilibrium if (i) no duopoly combination allows both entering exchanges to earn positive latent profits, (ii) AMEX has positive monopoly latent profits, and (iii) AMEX's monopoly latent profits are greater than those of the other two exchanges. Converting these conditions into the usual inequalities involving the error terms and the deterministic portions of the latent profit functions is straightforward but somewhat tedious.[7] The probability of the market outcome "only AMEX" is the probability of the resulting combination of inequalities being satisfied under the conditional distribution function for the three latent-profit error terms, and is denoted $P^{iA}_{|\varepsilon_{iS}}$. The probabilities of the remaining six market configurations are determined analogously and are denoted in a similar fashion.

The likelihood function is thus

$$L = \prod_{i=1}^{I} [P_i^0]^{(1-A_i)(1-C_i)(1-R_i)} \left[P^{iA}_{|\varepsilon_{iS}}\phi^S(\varepsilon_{iS})\right]^{A_i(1-C_i)(1-R_i)} \left[P^{iC}_{|\varepsilon_{iS}}\phi^S(\varepsilon_{iS})\right]^{(1-A_i)C_i(1-R_i)} \left[P^{iR}_{|\varepsilon_{iS}}\phi^S(\varepsilon_{iS})\right]^{(1-A_i)(1-C_i)R_i} \left[P^{iAC}_{|\varepsilon_{iS}}\phi^S(\varepsilon_{iS})\right]^{A_iC_i(1-R_i)} \left[P^{iAR}_{|\varepsilon_{iS}}\phi^S(\varepsilon_{iS})\right]^{A_i(1-C_i)R_i} \left[P^{iCR}_{|\varepsilon_{iS}}\phi^S(\varepsilon_{iS})\right]^{(1-A_i)C_iR_i} \left[P^{iACR}_{|\varepsilon_{iS}}\phi^S(\varepsilon_{iS})\right]^{A_iC_iR_i}.$$

The parameters of the latent profit and spread equations are estimated along with the elements of the error term variance matrix specified above.[8]

Computing the probabilities such as $P^{iA}_{|\varepsilon_{iS}}$ involves integrating a trivariate normal probability density function over several non-rectangular regions. Given the complexity of the conditions for most of the market outcomes, the computations become intractable. As is typical in applications based on the Nash equilibrium model of entry, we estimate the likelihood function by Simulated Maximum Likelihood (SML) with a smoothing adjustment.[9] Berry (1992) applied this method in estimating the entry behavior of airlines. Later, Reiss (1996) showed that a simulated maximum-likelihood estimator works well for such models. Subsequently, Toivanen and Waterson (2000) use simulated maximum likelihood is estimat-

[7]For example, condition (ii) gives rise to the simplest of the inequalities, namely that $\varepsilon_{iA} > -\pi_{iA}$. The other conditions require more algebraic manipulation but are not difficult.

[8]Starting values were obtained by first estimating simple probit models for the latent profit equations and an OLS model for the spreads. These were used as the initial starting values in estimation of the full model. The starting values were then perturbed by adding to each a random number drawn from the uniform distribution to each. The hill-climbing routine (GQOPT) iterated fifteen times, and the ending parameter estimates and log likelihood value were noted. This step was repeated 15-30 times. The ending parameter estimates associated with the best likelihood were then used as the initial starting values for a full estimation.

[9]For more information on simulated methods with discrete choice data, see Train (2003).

ing a model of entry for two multi-plant fast food chains (McDonalds and Burger King), and Mazzeo (2000, 2002) uses the technique in estimating the entry behavior and product quality choices of motel chains. Even with many fewer possible market outcomes, Mazzeo (2000) finds that the "complexity of the limits of integration make direct computation of the probability of the possible configurations infeasible."

In brief summary, the SML method for computing the probabilities of each of the market configurations is as follows. For each evaluation of the likelihood function, the error term in the spread equation for the option on security i is computed using the parameter values for that iteration. Given this computed spread error term along with the estimated parameters of the variance matrix, we determine the mean and the variance matrix for the distribution of the three latent profit function error terms conditional on the error term of the spread equation.

We next use a random number generator to draw a number of sets of the latent profit error terms ε_{iA}, ε_{iC}, and ε_{iR} in accordance with their joint distribution conditional on the error term in the spread equation. For each set of the three latent profit error term draws we compute the resulting latent profits for the three exchanges. We then use the Nash equilibrium concept to determine the equilibrium market configuration implied by the latent profits. The probability of a market configuration such as "only AMEX" , $P^{iA}_{|\varepsilon_{iS}}$, is the fraction of random draws that yield "only AMEX" as the equilibrium. Since this procedure generates a likelihood function with flat regions and steep spikes, we apply Mazzeo's (2002) method for smoothing the likelihood function[10]. The Appendix provides a more complete description.

One additional detail concerns the manner in which the exchange binaries enter the latent profit equations. The construction of the Nash equilibrium for this type model is based on the explicit assumption that monopoly profits are greater than non-monopoly profits, which implies that the coefficients θ_{i-j} for i,j=A,C,R and $i \neq j$ are negative. In estimating only the latent profit equations, the algorithm converged to negative θ_{i-j} values without ever entering problematic regions in which some of the θ_{i-j}'s were positive (Danis, 2003). In the joint estimation of the latent profit and the spread equations, the early iterations sometimes

[10]Stern (1997) discusses the inherent difficulties in estimating such likelihood functions.

entered the regions with positive θ_{i-j}'s . Given the other complications of the estimation, we forced the coefficients to be negative.

5 Data

We created a data set of market-level variables using Compustat and CRSP data sets. From the 2002 Compustat quarterly update file, we gathered data on industry, assets, and whether the underlying exchange was Nasdaq. We used the latest quarter of available data for the assets variable. From CRSP, we gathered data on trading volume during calendar year 2002. We also computed a measure of volatility from September 30 through December 31, 2002, using the return variable located in CRSP. The observations were merged using cusip number.

The resulting data set contained 8,631 observations with options traded on 2,036. Not all of the remaining 6,595 companies were eligible for options trading. Since October 1991, companies have been required to have a minimum of 2,000 shareholders, have a market price per share of $7.50 for the majority of the previous three calendar months, maintain a trading volume of 2.4 million shares over the previous 12 months, and have a public float of 7 million shares in order for options to be traded. Given the difficulty of determining the number of shareholders, we follow Mayhew and Mihov (2000) in not classifying companies as ineligible for options trading on this basis. With the number of shares outstanding used as a measure of public float, 2,234 out of the 6,595 companies on which options are not traded were classified as ineligible. An additional 998 companies are eliminated because their median prices over the three prior months were less than $7.50, and 118 more companies were eliminated because their trading volumes were less than 2.4 million over the period June 1999 to June 2000. There were 426 companies with options traded for which we could not find any data in Compustat. Option trades on these companies represent 8.7% of the total volume of contracts traded.

To obtain information on bid-ask spreads, we downloaded data from the website of the Options Clearing Corporation (OCC). During the week of May 8 through May 15, 2003, we located information on near-term options. Most of the options (96%) expired in May

2003, but for certain underlying equities the nearest term option was either June, July, or September 2003. We downloaded data on the call contract with a strike price closest to the price of the underlying security at the time of the download. We gathered data on the price of the underlying security and several details about the call contract at each exchange including strike price, bid price, ask price, and open interest. Because the information is gathered in real time, there were sometimes slight differences in the spreads across exchanges. In such cases, the spread is the average across the exchanges. We merged the spread data to the market-level variables using the stock ticker symbol.

The result of the merge was a data set with 2,628 observations of which 1,856 had options traded. Table 1 contains summary statistics for the underlying securities in the data set. Companies whose options are traded have larger assets than those with no options. Also, the underlying stocks of companies for which options are offered have higher volatility but similar volume to those for which options are not offered. The breakdown of the market structures in the data set is in Table 2. Compared to the 2000 data analyzed in Danis (2003), there is a trend towards more entry. In 2000, only 12% of option classes were traded on 3 or more exchanges, while in 2003 that number has risen to 31%. As further evidence of the trend towards more entry, the number of underlying stocks eligible for option listing but without options trading has fallen from 1,024 in 2000 to 772 in 2003.

Table 2 also details the mean and median bid-ask spread by market structure. The spread is smaller for those markets in which more exchanges list an option, from an average 0.25 when only one exchange lists the option to 0.15 when all three list it. Without controlling for any exogenous variables, a negative relationship between the spread and entry is clear.

6 Empirical Findings

We briefly summarize the empirical findings and then used the estimated coefficients, as reported in Table 3, to determine how the collusive agreements of the 1990s affected the spreads on option prices. For the spread equation, the estimated coefficients on the binaries for whether AMEX, CBOE, and the Regionals enter are all negative with t-statistics well above 10. The implied effects upon the spread range from a decline of 3.3 cents for AMEX entering to a decline of 4.5 cents for the Regionals. The estimates are similar to those of Mayhew (2001) who finds that multiple-listed options are on average four cents lower. For comparison, the estimates of the spread equation from an OLS model, which treats entry of the exchanges as exogenous, are reported in Table 4. In comparison to the full model, OLS overstates the effect of CBOE entry on the spread and understates the effect of AMEX and Regional.

The estimated coefficients for the latent profit equations are generally significant and of the expected signs. Three of the six coefficients for the effect of a competitor's presence are statistically significant. The entry of AMEX or CBOE lowers the latent profits of the Regionals. Likewise, the latent profits of AMEX are lowered by the presence of the Regionals. The remaining three of these coefficients are smaller in magnitude and statistically insignificant.

The characteristics of the underlying securities significantly affect the latent profits of all three exchanges. The estimated coefficients on both the value and the trading volume of the underlying asset are positive and statistically significant in all three latent profit equations. As expected, the exchanges are more likely to offer options on high trading volume securities and on securities of firms with high asset values. Somewhat curiously, the estimated coefficients on the volatility of the underlying asset are negative for AMEX and Regional but positive for CBOE. As measured by the effect of a one standard deviation change in the variable, trading volume has far larger effects on latent profits than does asset value or volatility. The coefficients on the industry binaries are generally significant, and the differing signs across exchanges indicate some specialization in the industries for which options are offered by the three exchanges.[11]

The estimated parameters that determine the variance structure for the error terms are all significant and lead to the following estimated variance matrix for ε_A, ε_C, ε_R, and ε_S

[11] In an alternative specification with the industry dummies included in the spread equation, the results were very similar and only the estimated coefficient on the service industries was significant.

$$
\begin{bmatrix}
1 & 0.3361 & 0.7343 & 0.0098 \\
0.3361 & 1 & 0.4257 & 0.0032 \\
0.7343 & 0.4257 & 1 & 0.0098 \\
0.0098 & 0.0032 & 0.0098 & 0.0084
\end{bmatrix}.
$$

The errors of the latent profit equations are strongly and positively correlated. The covariances between the latent profits and the bid-ask spread are also positive but are smaller in magnitude. Based on the above variance matrix, the correlations between the error terms in the spread equation and in the latent profits for AMEX, CBOE, and Regional are 0.11, 0.03, and 0.11 respectively. Not surprisingly, the same unobservables that make an option attractive to one exchange also make it attractive to the others.[12] In addition, the same factors associated with higher latent profits and a greater probability of entry are associated with higher spreads.

7 Study of Spreads during 1990s

During the 1990s, the exchanges allegedly colluded in order to limit the listing of options through the "Joint-Exchange Options Plan" devised by the exchanges. Although the 1990 SEC rule change allowed multiple listing, the exchanges allegedly agreed to limit trading to the existing exchanges. The Department of Justice began an investigation of the exchange listing practices in the late 1990s. Without admitting guilt, the exchanges agreed in 2000 to a consent decree prohibiting them from participating in any formal or informal listing agreements. The listing of all options on multiple exchanges began in earnest in 1999.

We examined 474 option classes that were trading on at least one of the exchanges prior to the implementation of the 1990 rule change and were still trading in January 1995. Using

[12] Allowing for the correlations complicates the model but is strongly reflected in the data. Forcing the correlations in the error terms of the latent profit equations to be zero lowers the likelihood significantly and produces misleading results. In a model with only the latent profit equations and the correlations set to zero, the coefficients on the binaries for which exchanges had entered (eg., θ_{A-C}) were forced towards zero and, if allowed to do so, became positive. With no allowance for correlation in the unobservables, the model could explain multiple listing only by having additional exchanges directly increase profitability.

the model parameters from Table 3 and the characteristics of the underlying securities from 1994, we predict the market configurations that would have occurred in the absence of the collusive agreements not to enter. Overwhelmingly, the model predicts market outcomes with more than one exchange offering the option and often with all three. For the 463 options offered on only one exchange in 1994, the model predicts an equilibrium outcome with three exchanges in 235 of the cases. For the 11 duopoly markets in 1994, the model predicts an equilibrium outcome with all three exchanges in 8 of the cases.

The impact on the bid-ask spread of the additional entry is significant. Although we do not have data on the spreads in 1994, we can predict the values using the parameters from Table 3. We calculate the spread given the actual market configuration and compare it to the estimated spread for the predicted market equilibrium configuration. The estimated decline in the spreads ranges from nearly 25 cents for options on companies such as Arco Chemical, Nashua Corporation, and Dover Corporation to about 3 cents for options on companies such as Micron Technology Inc. and Xerox. For option markets with additional entry predicted, the median decline in the spread is 7 to 8 cents (from spreads of approximately 23 cents to spreads of 16 cents). Even if entry had been limited to one rather than two exchanges and the resulting declines in the spread had been cut approximately in half, the impact on consumers would be significant. Given the trading volume during this time period, such declines in the spread would have a significant impact on consumers.

8 Conclusions

Financial market studies have examined the effect of competition on bid-ask spreads by estimating models for the spread as dependent on measures such as the number of exchanges or the Herfandahl index for the exchanges offering the security or options on it. A well-recognized concern is the potential effect of the spread on the attractiveness of entry and the number of firms listing a security. In order to address this issue, we estimate a joint model for the exchanges offering options on a security and the bid-ask spread on the option. In evidence of the endogeneity of the exchanges' decisions to offer the options, we find a consistent and statistically significant correlation among the unobservables affecting exchange entry and the bid-ask spreads. The correlation between the entry decisions by different exchanges is, however, notably higher than between the bid-ask spread and entry by any one of the exchanges. For the 474 security options affected by the alleged conspiracy, we predict a median decline of 8 cents in spreads had the equilibrium level of entry occurred. Given that the average predicted spread in a monopoly market during this time period is about 24 cents, an 8 cent decline is significant.

9 References

Berry, Steven T. (1992), "Estimation of a Model of Entry in the Airline Industry," *Econometrica*, 60(4), 889-917.

Bollen, Nicolas P.B., Tom Smith, and Robert E. Whaley (2002), "Modeling the Bid/Ask Spread: Measuring the Inventory-Holding Premium," *Journal of Financial Economics*, forthcoming.

Bresnahan, Timothy F. and Peter C. Reiss (1991a), "Empirical Models of Discrete Games," *Journal of Econometrics*, 48, 57-81.

Bresnahan, Timothy F. and Peter C. Reiss (1991b), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99(5), 977-1009.

Danis, Michelle A. (1997), "The Effects of Multiple Listing on Bid-Ask Spreads for Equity Options," Master's Thesis, Department of Economics, Virginia Tech.

Danis, Michelle A. (2003), "A Discrete Choice Approach to Measuring Competition in Equity Option Markets," Working Paper.

de Fontnouvelle, Patrick, Raymond P. H. Fishe and Jeffrey H. Harris, (2003), "The Behavior of Bid-Ask Spreads and Volume in Options Markets During the Competition for Listings in 1999," *The Journal of Finance*, forthcoming.

George, Thomas J. and Francis A. Longstaff (1993), "Bid/Ask Spreads and Trading Activity in the S&P 100 Index Options Market," *Journal of Financial and Quantitative Analysis*, 28, 381-397.

Janssen, Maarten and Eric Rasmusen (2001), "Bertrand Competition Under Uncertainty," mimeo.

Mayhew, Stewart and Vassil Mihov (2000), "Another Look at Option Listing Effects," mimeo.

Mayhew, Stewart (2001), "Competition, Market Structure and Bid-Ask Spreads in Stock Option Markets" mimeo.

Mazzeo, Michael J. (2000), "Product Choice and Oligopoly Market Structure," The Center for the Study of Industrial Organization at Northwestern University Working Paper.

Mazzeo, Michael J. (2002), "Product Choice and Oligopoly Market Structure," *Rand Journal*, 33(2), 221-242.

Neal, Robert (1987), "Potential Competition and Actual Competition in Equity Options," The Journal of Finance, 42(3), 511-531.

Reiss, Peter C. (1996), "Empirical Models of Discrete Strategic Choices," *American Economic Review*, 86(2): 421-426.

Stern, Steven (1997), "Simulation-Based Estimation," *Journal of Economic Literature*, 35, 2206-2039.

Toivanen, Otto and Michael Waterson (2000), "Empirical Research on Discrete Choice Game Theory Models of Entry: An Illustration," *European Economic Review*, 44, 985-992.

United States of America v. American Stock Exchange, Chicago Board Options Exchange, Inc., Pacific Exchange, Inc., and Philadelphia Stock Exchange, Inc. (2000), Complaint for Equitable Relief for Violation of 15 U.S.C. § 1.

United States of America v. American Stock Exchange, Chicago Board Options Exchange, Inc., Pacific Exchange, Inc., and Philadelphia Stock Exchange, Inc. (2000), Competitive Impact Statement.

United States of America v. American Stock Exchange, Chicago Board Options Exchange, Inc., Pacific Exchange, Inc., and Philadelphia Stock Exchange, Inc. (2000), Final Judgment.

Wahal, Sunil (1997), "Entry, Exit, Market Makers, and the Bid-Ask Spread," *Review of Financial Studies*, 10(3), 871-901.

Wang, Elizabeth Xiao-Ru (1999), "Competition among Exchanges: Does Multiple Listing Affect Trading Costs on Options Markets?", University of Chicago Working Paper.

10 Appendix – Error Term Distribution

10.1 Variance Matrix of Error Terms

The components of the variance matrix of the error terms are not directly estimated. As is standard, the error structure is defined in such away to ensure that the variance matrix is positive definite without having to impose inequality constraints. Specifically, the error terms are defined on the basis of four independent, normally distributed error terms. For each security i, the error terms ϵ^*_{ij} are distributed $N(0,1)$ for $j = A, C, R, S$. The error terms appearing in the four estimated equations for security i are defined as

$$\epsilon_{iA} \equiv \epsilon^*_{iA}$$

$$\epsilon_{iA} \equiv \frac{\alpha_{21}\epsilon^*_{iA} + \epsilon^*_{iC}}{\sqrt{1+\alpha_{21}^2}}$$

$$\epsilon_{iR} \equiv \frac{\alpha_{31}\epsilon^*_{iA} + \alpha_{32}\epsilon^*_{iC} + \epsilon^*_{iR}}{\sqrt{1+\alpha_{31}^2+\alpha_{32}^2}}$$

$$\epsilon_{iS} \equiv \alpha_{41}\epsilon^*_{iA} + \alpha_{42}\epsilon^*_{iC} + \alpha_{43}\epsilon^*_{iR} + \alpha_{44}\epsilon^*_{iS}$$

Given this structure, the variance of the error terms is

$$\begin{bmatrix} 1 & \sigma_{AC} & \sigma_{AR} & \sigma_{AS} \\ \sigma_{AC} & 1 & \sigma_{CR} & \sigma_{CS} \\ \sigma_{AR} & \sigma_{CR} & 1 & \sigma_{RS} \\ \sigma_{AS} & \sigma_{CS} & \sigma_{RS} & \sigma_S^2 \end{bmatrix} =$$

$$\begin{bmatrix} 1 & \frac{\alpha_{21}}{\sqrt{1+\alpha_{21}^2}} & \frac{\alpha_{31}}{\sqrt{1+\alpha_{31}^2+\alpha_{32}^2}} & \alpha_{41} \\ \frac{\alpha_{21}}{\sqrt{1+\alpha_{21}^2}} & 1 & \frac{\alpha_{21}\alpha_{31}+\alpha_{32}}{\sqrt{1+\alpha_{21}^2}\sqrt{1+\alpha_{31}^2+\alpha_{32}^2}} & \frac{\alpha_{21}\alpha_{41}+\alpha_{42}}{\sqrt{1+\alpha_{21}^2}} \\ \frac{\alpha_{31}}{\sqrt{1+\alpha_{31}^2+\alpha_{32}^2}} & \frac{\alpha_{21}\alpha_{31}+\alpha_{32}}{\sqrt{1+\alpha_{21}^2}\sqrt{1+\alpha_{31}^2+\alpha_{32}^2}} & 1 & \frac{\alpha_{31}\alpha_{41}+\alpha_{32}\alpha_{42}+\alpha_{43}}{\sqrt{1+\alpha_{31}^2+\alpha_{32}^2}} \\ \alpha_{41} & \frac{\alpha_{21}\alpha_{41}+\alpha_{42}}{\sqrt{1+\alpha_{21}^2}} & \frac{\alpha_{31}\alpha_{41}+\alpha_{32}\alpha_{42}+\alpha_{43}}{\sqrt{1+\alpha_{31}^2+\alpha_{32}^2}} & \alpha_{41}^2+\alpha_{42}^2+\alpha_{43}^2+\alpha_{44}^2 \end{bmatrix}.$$

Any positive definite matrix can be parameterized as written in the matrix above, and any such matrix is positive definite. Accordingly, we estimate the α coefficients that appear in the above matrix rather than directly estimate the variances and covariances of the error terms in the latent profit and the spread equations. The variances and covariances are then computed from the estimated $\alpha' s$.

10.2 Conditional Distribution of ε_A, ε_C, ε_R, and ε_S

Since ϵ_A, ϵ_C, ϵ_R, and ϵ_S are jointly normally distributed, ϵ_A, ϵ_C, and ϵ_R are normally distributed conditional on ϵ_S. The conditional distribution in terms of the variances and covariances of the joint normal distribution is

$$\epsilon_A, \epsilon_C, \epsilon_R | \epsilon_S \sim N\left(\begin{bmatrix} \frac{\sigma_{AS}}{\sigma_S^2} \\ \frac{\sigma_{CS}}{\sigma_S^2} \\ \frac{\sigma_{RS}}{\sigma_S^2} \end{bmatrix} \epsilon_S, \begin{bmatrix} 1 - \frac{\sigma_{AS}^2}{\sigma_S^2} & \sigma_{AC} - \frac{\sigma_{AS}\sigma_{CS}}{\sigma_S^2} & \sigma_{AR} - \frac{\sigma_{AS}\sigma_{RS}}{\sigma_S^2} \\ \sigma_{AC} - \frac{\sigma_{AS}\sigma_{CS}}{\sigma_S^2} & 1 - \frac{\sigma_{CS}^2}{\sigma_S^2} & \sigma_{CR} - \frac{\sigma_{CS}\sigma_{RS}}{\sigma_S^2} \\ \sigma_{AR} - \frac{\sigma_{AS}\sigma_{RS}}{\sigma_S^2} & \sigma_{CR} - \frac{\sigma_{CS}\sigma_{RS}}{\sigma_S^2} & 1 - \frac{\sigma_{RS}^2}{\sigma_S^2} \end{bmatrix} \right).$$

Table 1
Median Characteristics of Underlying Companies Eligible for Option Listing
May 2003

	Companies with Options	Companies without Options
Total Assets ($millions)	1,337.4	713.4
Underlying Volatility	0.53	0.38
Trading Volume 2002 (million)	1.18	1.16

Industry Classification of Underlying Companies Eligible for Option Listing

	Companies with Options	Companies without Options
Manufacturing	302	89
Services	337	82
Finance, Insurance, and Real Estate	264	293
Transportation, Communications, Electric, Gas and Sanitary Services Dummy	189	72
Other	821	236

Table 2
Mean and Median Spread by Market Structure

Market Structure	Count	Mean	Median
No exchanges entered	772		
AMEX only	282	$0.25	$0.25
CBOE only	161	0.22	0.25
Regional only	193	0.23	0.25
AMEX-CBOE duopoly	212	0.21	0.22
AMEX-Regional duopoly	122	0.21	0.22
CBOE-Regional duopoly	81	0.19	0.20
All three exchanges entered	805	0.15	0.15
Total	2628		

Table 3

Estimation Results on Listing Decisions and Bid-Ask Spreads

May 2003 Data

Variable	Coefficient	Standard Error
Theta coefficients (enters latent profit equations as the negative of the coefficient squared):		
Effect of AMEX entry on CBOE	0.0360	0.0306
Effect of AMEX entry on Regional	0.1617**	0.0048
Effect of CBOE entry on AMEX	0.0009	0.0190
Effect of CBOE entry on Regional	0.0510**	0.0015
Effect of Regional entry on AMEX	0.8053**	0.0072
Effect of Regional entry on CBOE	0.0181	0.1297
AMEX coefficients:		
Assets (standardized)	0.4714**	0.0128
Volatility (standardized)	-0.0258**	0.0068
Trading Volume (standardized)	5.1220**	0.0388
Manufacturing Dummy	0.1121**	0.0161
Services Dummy	-0.0542*	0.0243
Finance, Insurance, Real Estate Dummy	-0.1886**	0.0222
Transportation, Communications, Electric, Gas and Sanitary Services Dummy	-0.5953**	0.0187
Nasdaq Dummy	0.0430*	0.0178
Constant	1.3710**	0.0118
CBOE coefficients:		
Assets (standardized)	0.6209**	0.0314
Volatility (standardized)	0.1372**	0.0057
Trading Volume (standardized)	6.8560**	0.0437
Manufacturing Dummy	0.2392**	0.0205
Services Dummy	0.0485	0.0287
Finance, Insurance, Real Estate Dummy	0.0938**	0.0189
Transportation, Communications, Electric, Gas and Sanitary Services Dummy	-0.5440**	0.0245
Nasdaq Dummy	0.0967**	0.0153
Constant	1.2530**	0.0148

Table 3 (continued)
Estimation Results on Listing Decisions and Bid-Ask Spreads
May 2003 Data

Variable	Coefficient	Standard Error
REG coefficients:		
Assets (standardized)	0.4936**	0.0302
Volatility (standardized)	-0.0484**	0.0066
Trading Volume (standardized)	6.3230**	0.0516
Manufacturing Dummy	0.0014	0.0182
Services Dummy	-0.1129**	0.0152
Finance, Insurance, Real Estate Dummy	-0.1078**	0.0171
Transportation, Communications, Electric, Gas and Sanitary Services Dummy	-0.3720	0.0149
Nasdaq Dummy	-0.0542**	0.0127
Constant	1.4442**	0.0137
Spread coefficients:		
Effect of AMEX entry	-0.0326**	0.0016
Effect of CBOE entry	-0.0386**	0.0022
Effect of Regional entry	-0.0449**	0.0003
Assets (standardized)	-0.0008	0.0014
Volatility (standardized)	-0.0058**	0.0010
Trading Volume (standardized)	-0.0120**	0.0014
Days to expiration	0.0002*	0.0001
Strike price minus underlying price	-0.0030**	0.0005
Constant	0.2688**	0.0006

Table 3 (continued)
Estimation Results on Listing Decisions and Bid-Ask Spreads
May 2003 Data

Variable	Coefficient	Standard Error
Error coefficients:		
α_{21}	0.3569**	0.0103
α_{31}	1.1266**	0.0131
α_{32}	0.2915**	0.0066
α_{41}	0.0098**	0.0000
α_{42}	-0.0001	0.0002
α_{43}	0.0040**	0.0003
α_{44}	0.0913**	0.0003
Log Likelihood		-1,838.2
Number of observations		2628

*denotes significance at the 5% level **denotes significance at the 1% level

Notes: Smoothing parameter=0.01, 100 draws of the error terms.
Model estimated in Fortran using GQOPTs DFP algorithm with stretching.

Sources: CRSP, Center for Research in Security Prices. Graduate School of Business, The University of Chicago 1999-2000. Used with permission. All rights reserved. www.crsp.uchicago.edu. Standard & Poors, COMPUSTAT (North American) data; The Option Clearing Corporations Market Share History.

Table 4
Bid-Ask Spread Equation
Full Model versus Ordinary Least Squares Results

	Full Model		OLS	
	Coefficient	Standard Error	Coefficient	Standard Error
Effect of AMEX Entry	-0.0326**	0.0016	-0.0230**	0.0055
Effect of CBOE Entry	-0.0386**	0.0022	-0.0466**	0.0051
Effect of Regional Entry	-0.0449**	0.0003	-0.0425**	0.0050
Assets (standardized)	-0.0008	0.0014	-0.0016	0.0020
Volatility (standardized)	-0.0058**	0.0010	-0.0078**	0.0022
Trading Volume (standardized)	-0.0120**	0.0014	-0.0128**	0.0021
Days to Expiration	0.0002*	0.0001	0.0001	0.0002
Strike Price minus Underlying Price	-0.0030**	0.0005	-0.0029**	0.0007
Constant	0.2688**	0.0006	0.2717**	0.0064

**denotes significance at the 1% level, *denotes significance at the 5% level