Estimating two structural spread models for trading blind principal bids

November 2006

Christos Giannikos[&] and Tin Shan Suen[⊕]

Baruch College, CUNY

Abstract

In this paper, we estimate two different structural models for a liquidity provider's spread when trading a basket of stocks simultaneously and immediately using a market mechanism called blind principal bid. The spread is a liquidity demander's cost for immediacy. Two previous studies by Kavajecz and Keim (2005) and Giannikos and Suen (2006) investigate a similar question. Neither study, however, is based on any structural models. Considerable market microstructure research has attempted explicitly to model a dealer's spread for trading a single stock in a quote-driven market. This paper applies two structural spread models, developed by Stoll (1978a, 1978b) and Bollen, Smith and Whaley (2004), in the context of trading a basket of stocks simultaneously. We found that both models perform quite well, explaining a liquidity provider's spread for trading a basket of stocks. Using Bollen, Smith and Whaley's (2004) methodology, we are able to decompose the spread into various fundamental and structural cost components and rank their relative magnitude. The inventory-holding costs are found to be the largest cost component. We empirically deduce an implied trading rate used by a liquidity provider to unload shares in a basket after accepting a basket. For trades within a basket, we also analyze an empirical relation between informed trades and their (dollar) trade size (expressed in percentage of average daily dollar volume). A potential application of this study is to help asset managers develop a benchmark trading cost when trading a basket of stocks using blind principal bids.

[&] *Corresponding author:* Christos Giannikos, Department of Economics and Finance, Baruch College, One Bernard Baruch Way, Box B10-225, New York, NY 10010, USA. Tel: (646) 312-3492, Fax: (646) 312-3451, Email: Christos_Giannikos@baruch.cuny.edu

[⊕] Tin Shan Suen, Ph.D. student, Department of Economics and Finance, Baruch College, One Bernard Baruch Way, Box B10-225, New York, NY 10010, USA. Tel: (646) 312-3501, Fax: (646) 312-3451, Email: Tin-Shan_Suen@baruch.cuny.edu

I. Introduction

The market microstructure models for a stock dealer's[1] spread in a quote-driven market are based on trading a single stock. In this paper, we estimated two structural models for a dealer's spread when a basket of stocks is being traded. These two models were originally developed by Stoll (1978a, 1978b) and Bollen, Smith and Whaley (2004), respectively. Whereas they use trade data on an individual stock, our model estimation is done using data on trading a basket of stocks simultaneously. This is one of the most important features of this study. Our data come from two active equity asset managers who trade stock baskets using a special market mechanism called blind principal bid (BPB).[2] The liquidity providers, typically major sell-side firms, are customarily called blind principal brokers. BPB brokers commit their own capital to provide liquidity by being the counter party for each of the trades within a basket of stocks. (Section III provides the institutional details of BPB, in brief.)

The motivation for this study is based on our observation that the role and function of a BPB broker is very similar to that of a dealer in a quote-driven market, especially with respect to the risk exposure faced by a dealer.[3] Both dealers and BPB brokers facilitate trading by providing immediate liquidity using their own capital. Their capital is exposed

[1] We use the term dealer and market maker interchangeably throughout this paper.

[2] For the remainder of the paper, we use BPB as the abbreviation for blind principal bid.

[3] The term "blind principal bid dealers" (rather than "blind principal bid brokers") may be a better term for this type of liquidity providers, since their roles and functions are similar to a dealer (who bears risk) in a quote-driven market, Strictly speaking, a broker does not face the same risk as a dealer. Nevertheless, we shall continue to use the industry norm, blind principal broker, for the remainder of the paper.

to certain risks while providing a liquidity service (e.g. adverse price movement or the other party is informed). Therefore, they will certainly request compensation for their risk exposure: for dealers, in the form of a spread; for BPB brokers, a fee (or trading cost[4]) paid by an asset manager. Our hypothesis involves using the two dealer's spread models mentioned above, subject to some minor extensions, to model the trading cost paid by users of BPB. Our insight is that a BPB broker's fee and a dealer's spread are conceptually equivalent. One potential application of our result is that BPB users can now set a benchmark cost on how much trading cost they should pay for a BPB broker's liquidity service.

The remainder of this paper is organized as follows. Section II reviews the literature on modeling spread and the various components of a spread. Section III describes the basket trading data used in this study. Section IV and V describe how we estimate Stoll's (1978a, 1978b) model and Bollen, Smith and Whaley's (2004) model, respectively. Section VI concludes the paper.

II. Literature Review

Since Garman (1976) coined the term "market microstructure," there has been an explosive growth of research in this field. This section provides a very brief overview of various research topics relevant to this paper.[5]

[4] Giannikos and Suen (2006) refer to this type of trading cost as a "liquidity risk premium."

[5] Several excellent survey papers and books offer a more detailed and comprehensive treatment; see, for example, Coughenour and Shastri (1999), Madhavan (2000), Stoll (2003), Biais et al. (2005), O'Hara (1997), and Hasbrouck (2006).

One early approach to modeling a dealer's spread was to tackle the problem as an inventory management problem from the dealer perspective. This approach is commonly referred as the inventory model in the market microstructure literature (e.g. O'Hara (1997)). Several studies take this approach. Garman (1976) models the spread as an optimization problem in which the dealer chooses an optimal bid and ask price. The broker sets bid and ask prices only once at the beginning of time. The objective function is to maximize the dealer's expected profit while taking into consideration the problem of running out of cash or inventory and when the arrival of trade orders is stochastic. Stoll (1978a) also solved the inventory management problem as an optimization problem, but models the compensation for a dealer who holds a sub-optimal portfolio while providing liquidity. Holding sub-optimal portfolio means extra risk exposure that generates a dealer's compensation. Stoll (1978a, 1978b) also analyzes various fundamental cost components in a dealer's spread. Amihud and Mendelson (1980) began with framework similar to that of Garman (1976), but allow bid and ask prices to change when inventory changes. Ho and Stoll (1980), Ho and Stoll (1981), and Ho and Stoll (1983) are extensions and enhancement to Stoll (1978a). O'Hara and Oldfield (1986) tried to address some of the restrictions in Ho and Stoll's (1981) model, for example, by employing an infinite time horizon rather than the finite time horizon in Ho and Stoll (1981).

There are several empirical studies as well. Stoll (1978b) is the empirical test for Stoll (1978a). Ho and Macris (1984) use transaction prices and dealer inventory of some

American Stock Exchange options and find some empirical evidence supporting Ho and Stoll (1981). They show that dealers' spread is positively correlated to asset risk and dealers adjust their quotes in response to their inventory position. Hasbrouck and Sofianos (1993) find that dealers have a preferred level inventory, adjusting the bid and ask prices to bring inventory to a preferred level. Hansch et al. (1998) conduct direct tests of Ho and Stoll's (1983) inventory model using London Stock Exchange data. One of their findings is that relative inventory position is significantly related to the ability to execute large trades, which supports the inventory model of dealer's spread. Naik and Yadav (2003) test Ho and Stoll's inventory model and also investigate a particular question where dealer firms manage inventory on a stock-by-stock or portfolio basis. They find that individual dealers manages their own inventory but do not focus on the overall inventory of their firms.

Another approach to model a dealer's spread is from an information asymmetry perspective, based on adverse selection theory. This class of models takes into consideration that a dealer may be disadvantaged when trading with an informed trader. Some of the literature, e.g. O'Hara (1997), refers to this as the information asymmetry model. Bagehot (1971) is thought to be the first study of this information asymmetry. Copeland and Galai (1983) model a dealer's spread as maximum expected profit, balancing losses from trading with informed traders and gains from trading with uninformed (liquidity) traders. Glosten and Milgrom's (1985) model incorporates an additional element not included in Copeland and Galai's model, namely, that informed traders' trades have information content. Even a dealer cannot distinguish an informed from an uninformed trader; a dealer alters his expectation of a stock's true price conditional on the trades he receives. Easley and O'Hara's (1987) model differs from Glosten and Milgrom's (1985) model in one major aspect: they explicitly consider the effect of trade size executed by traders on stock prices, motivated by empirical observation that large trades are executed at worse prices.

Since the spread can be modeled as a function of inventory holding and information asymmetry, researchers have investigated whether one can decompose the spread into these two components. A broader issue is the estimation of various components (e.g. order processing costs, inventory costs, and information asymmetry costs) that contribute to the bid-ask spread. Empirical studies along these lines include Roll (1984), Choi et al. (1988), Glosten and Harris (1988), Stoll (1989), George et al. (1991), Madhavan et al. (1997), and Huang and Stoll (1997). Huang and Stoll manage to generalize all these works. Coughenour and Shastri (1999) provide a concise survey of this topic, wherein most papers are based on time series analysis. There are, however, some studies that use a cross-sectional approach, for example, Stoll (1978b) and Bollen et al. (2004). Since time series data do not exist for a BPB basket, Stoll and Bollen et al.'s cross-sectional approach makes estimation of various trading cost components feasible for the BPB basket. We discuss these two studies in more detail in Section IV and V.

Some literature also investigates the empirical evidence of the various determinants of a dealer's spread. Demsetz (1986), Tinic (1972), Tinic and West (1972, 1974), Benston and Hagerman (1974) and Branch and Freed (1977). Bollen, Smith and Whaley et al. (2004)

provide a concise summary and comparison of these studies. Two studies, Kavajecz and Keim (2005) and Giannikos and Suen (2006), are particularly relevant to this paper. Both investigate the empirical determinants of BPB trading cost.

III. Data

Before describing in detail the BPB data used in this study, it is beneficial to give a brief overview of BPB as a trading mechanism, since some readers may be unfamiliar with this alternative trading venue.[6]

III.a. BPB overview

BPB is a trading mechanism that allows a liquidity demander (e.g. quantitative asset managers[7] or transition managers[8]) to trade a basket of stocks at a predetermined contractual execution price for each stock in a basket. The cost of trading the basket is determined by a bidding (auction) process in which competing liquidity providers (i.e. sell-side BPB brokers) submit their best bid (price), usually quoted as cents per share, for

[6] For a more in-depth description of BPB's institutional details, see Giannikos and Suen (2006), Kavajecz and Keim (2005), Almgren and Chriss (2003), and Kissell and Glantz (2003).

[7] Many quantitative asset managers use an optimizer to re-balance their portfolio. A trade list is being generated by an optimizer and it converts an existing portfolio to an optimal portfolio solved by an optimizer. In other words, these asset managers would like to execute trades within a trade list quickly and simultaneously.

[8] A transition manager takes control of a portfolio from an asset manager who has just been terminated, receiving instructions from a newly hired asset manager. The instructions include a list of securities the new manager would like to have in his portfolio. The transition manager needs to execute the necessary trades (in a short period) to convert the existing portfolio into that requested by the new manager.

trading an entire basket for a manager. A broker whose bid is the minimum usually executes a basket for the manager. The winning BPB bid will be used to calculate the trading cost paid by the liquidity demander to the BPB broker for providing liquidity (and satisfying the manager's desire for immediacy). Our key insight in this paper is that such a trading cost is conceptually very similar a dealer's spread, that is, dealer's spread is the cost of immediacy for a single stock while the trading cost for BPB is the cost of immediacy for a basket of stocks.

Two bidding procedures are typically used for a BPB basket:

1. pre-open bidding
2. post-close bidding

For pre-open bidding, summary characteristics[9] of a basket are sent to competing brokers early in the morning before the market is open (e.g. some time between 7 a.m. and 8 a.m.). Bidding is blind in the sense that individual names in a basket are not provided to competing brokers. The execution price for each stock in a basket is agreed upon by the manager and brokers. In this case, the execution price is a stock's closing price on the

[9] Typical summary characteristics include the dollar value of a basket (buy and sell), the number of shares in a basket (buy and sell), the number of stocks in a basket (buy and sell), sector exposures (buy and sell), and so on. An example list can be found in Giannikos and Suen (2006). Standard reports that describe a basket's characteristics can be sent by a manager to competing brokers.

previous business day. Broker's bid is usually quoted as cents per share.[10] For example, if a broker submits a bid, say, 6 cents per share for a basket with 1,000,000 shares, it means the manager pays $60,000 for trading the entire basket at the agreed-upon execution prices (i.e. the previous day's closing prices). The broker whose bid is the lowest usually wins the basket and executes all the transactions within a basket. At this point, a manager can regard all trades within a basket as completed. Basically, the buy and sell lists within a basket are transferred from a manager to a BPB broker's inventory. Since the execution price is contractual, there is no market impact cost from the manager's perspective. However, the winning (lowest) bid is usually higher than the commission for institutional agency trades (typically about 2 cents per share) because broker needs to commit its own capital in order to provide liquidity to a manager. If prices of stocks in a basket move adversely against a winning broker, he may suffer severe capital loss. Therefore, the broker charges a higher price to compensate for his risk exposure while providing liquidity to the manager. From the managers' perspective, the tradeoff is between a higher explicit winning bid cost with no market impact risk and no unfinished trades and a lower explicit commission cost with unknown market impact and the possibility of unfinished trades.

The post-close bidding procedure is very similar to that of pre-open bidding. Basket summary characteristics are sent out to competing brokers right after the maker closes. The execution price for each stock in a basket is its corresponding same-day closing

[10] We may also express the cost of a blind principal bid as basis points. It is computed as the ratio between the dollar value paid to the winner broker and the total dollar value of a basket.

price. The winning broker is not able to unload most names in a basket until the market reopens on the next business day.

III.b. Historical BPB Data

The following is a description of the BPB basket data used in this study. We have gathered 196 BPB baskets executed regularly by two active asset managers.[11] These baskets were traded during the period from January 2002 to September 2005. The trading activities are quite evenly distributed throughout the sample period.[12] Manager A uses only pre-open bidding, while Manager B uses both pre-open and post-close bidding. For each basket executed, we gathered the following data items. (Data items specific to Stoll's (1978) model and Bollen, Smith, and Whaley's (2004) model are discussed in the next two sections.)

1. Stock identifier (CUSIP or ticker) for each name in a basket
2. Transaction type for each name (buy or sell) in a basket
3. Number of shares traded for each name in a basket
4. Execution price[13] for each name in a basket

[11] We would like to thank a consulting firm specializing in securities transactions for providing transaction records for one of its managers. Due to issues of confidentiality, the names of the money managers and the winning brokers were excluded from the records before we received the data. We obtained a second set of transaction records from another asset manager. We refer to these two managers, both of whom specialize in quantitative investment strategies, as Managers A and B.

[12] The trading frequency is about once a week.

[13] For pre-open bidding, the execution price is the previous day's closing price. For post-close bidding, the execution price is the same- day closing price. Please refer to Section III for more detail.

5. Lowest (winning) bid (cents per share)

The BPB trading cost (in dollars) paid by the manager for trading a BPB basket can be computed as the total number of shares in a basket times the lowest bid (i.e. winning bid). In some cases, it is convenient to express the cost in basis points. It can be computed by dividing the cost (in dollars) by the total dollar value of a basket. When a BPB trading cost is expressed in basis points, it is conceptually similar to the relative spread or percentage spread (i.e. Spread / Price) in the market microstructure literature. Table I provides descriptive statistics of some basket characteristics.

IV. Estimating Stoll's (1978a, 1978b) model

In this section, we describe how to estimate Stoll's (1978a, 1978b) model using the BPB basket trading data. Stoll (1978a) focuses on developing a theoretical model for the holding cost component in a dealer's spread based on inventory modeling and, to a lesser extent, other cost components. Stoll (1978b) is the empirical counterpart of the structural model developed by Stoll (1978a). Moreover, Stoll (1978b) tried to estimate a structural model in which the spread consists of three cost components (holding cost, order cost, and information cost). As mentioned in the original paper, Stoll sought to "develop a more explicit and rigorous model of the individual dealer's spread." This point is again emphasized by Bollen, Smith and Whaley (2004), who are concerned with the structural form of a spread model because many models are based on economic reasoning rather than formal mathematical modeling. This leads to the criticism of ad hoc model specification and variable selection.

The original model is a spread model for a single stock name, but our description emphasizes the issues that are more relevant in the context of trading a basket of stocks simultaneously. Our model estimation basically follows the Stoll's (1978b) approach. However, we have a slightly different model specification. In the original model, Stoll included a factor for competition (number of dealers). We cannot include this competition factor in our model because we do not have data on the number of BPB brokers that were bidding on each basket.

IV.a. Model overview

In the following we describe the dependent and independent variables in our cross-sectional regression. As in Stoll's original paper, the natural logarithms of the variables are used when conducting the regression. We also borrow the variable symbols from the original paper to facilitate the cross-reference between two papers.

Basket trading cost (expressed in basis points[14]), s_i, is the dependent variable. In the original paper, s_i is the percentage spread. i is the BPB basket identifier (i.e. i = 1, 2, 3,…., 196). All other variables given below are independent variables.

Basket variance, σ^2_i. It is a direct measurement of the risk for a basket.[15] The sign of the estimated coefficient for σ^2_i is expected to be positive. A BPB broker will charge more for a more risky basket.

[14] The BPB trading cost = (cost per share × total number shares traded) / (total BPB basket dollar value traded)

[15] The basket variance is estimated using the MSCI Barra U.S. risk model.

Basket weighted-average volume, V_i, is the weighted average of dollar volume for a basket. For each stock in a basket, we compute its average daily dollar trading volume during the last 10 trading days (before the date of basket bidding). The weight used in calculating V_i is the dollar value of a trade for a stock (within a BPB basket) divided by the total dollar value of a basket. Basket volume is used as a proxy for the holding period in Stoll's model. The greater the volume, the shorter time (i.e. less risk) taken by a BPB broker to unload (or reverse) positions in a basket. The sign of the estimated coefficient for V_i is expected to be negative.

Basket weighted-average turnover, $(V/T)_i$. The turnover for each stock within a BPB basket is defined as the 10-day average dollar volume divided by its market capitalization.[16] The weight used in calculating $(V/T)_i$ is the dollar value of a trade for a stock (within a basket) divided by the total dollar value of a basket. Basket turnover is used as a proxy for adverse selection in Stoll's model. Based on Stoll's original argument, if a trade is driven by liquidity need (i.e. it originated from an uninformed trader), the traded volume tends to be proportional to market capitalization. However, if a trade for a stock originates from an informed trader, volume tends to be out of proportion to the stock's market capitalization. The sign of the estimated coefficient for turnover is expected to be positive.

[16] Market capitalization is calculated based on the latest available closing price for a stock when basket bidding (auction) occurs.

Basket weighted-average stock price, P_i, is a proxy for minimum cost in the original model. For each stock in a BPB basket, we gather the stock's latest closing price (see the footnote for computing stock's market capitalization). As with other variables, the weight used in calculating the basket weighted-average stock price is the dollar value of a trade for a stock (within a basket) divided by the total dollar value of a basket. Stoll argues that there is no prior expectation for the sign of the estimated coefficient.

In the original paper, Stoll also models the effect of dealer competition on the size of a spread. Due to the lack of data on the number of BPB brokers bidding for each historical basket, we do not include this particular independent variable in this study. Table II provides some descriptive statistics for the variables used in the estimation.

IV.b. Model estimation result and analysis

Table III summarized the result of estimating Stoll's (1978b) model. Overall, Stoll's model performs quite well to explain BPB basket trading cost. The R-square of the OLS regression is approximately 74%, which is similar to that of Kavajecz and Keim (2005) and Giannikos and Suen (2006). All estimated coefficients are highly significant with the exception of "basket weighted-average stock price." In the original paper by Stoll (1978b), the stock price is used as a proxy for minimum cost. Since the estimated coefficient for basket weighted-average stock price is not significant, it may indicate that we need another proxy for minimum cost in the context of BPB basket trading. All the signs for the estimated coefficients are consistent with our expectation, as shown in Table II.

A larger data sample that includes data from two different assets managers (rather than one manager, as in Kavajecz and Keim's (2005) study), as well as a more parsimonious model specification than Giannikos and Suen (2006), are not the only improvements in this paper. A key feature that distinguishes this study from the two earlier studies is our use of a theoretical framework for the trading cost of a BPB basket and estimation of the structural trading cost model. The structure of our model is not based on ad hoc economic reasoning, but is well defined.

There is, however, one limitation when applying Stoll's model directly in the context of a BPB basket trading. In Stoll's original model, the inventory holding cost component reflects a single period model and the variance of a stock is assumed to be stationary during the period. This assumption may not be true when a BPB broker tries to unload a BPB basket. Our discussions with BPB brokers reveal that they usually unload stocks within a basket at different speeds. Therefore, the basket variance is unlikely to remain stationary during the period when a BPB broker is unloading stocks from a basket. Our second structural model can address this limitation.

V. Estimating Bollen, Smith, and Whaley's (2004) model

In this section, we describe and estimate a spread model based on the methodology developed by Bollen, Smith and Whaley (2004). As we have emphasized earlier, this spread model provides a theoretically grounded functional form of the relationship between the spread and its determinants. Their original model is for trading a single stock; a straightforward extension allows us to apply their model in the context of trading BPB baskets. An additional benefit is that this model can compute various components of a spread. Thus, we can compare the distribution of various cost components within a BPB basket spread with some other studies (which tend to analyze the cost categorization of a stock spread). A unique feature is that this model is one of the few that does not use time series data; as mentioned in the literature review, many previous studies do utilize time series data to estimate various cost components in a spread. Since there are no time series data available for BPB baskets, Bollen, Smith and Whaley's (2004) methodology becomes attractive in studying trading cost components for BPB baskets. To facilitate discussion of our study, we provide a very brief overview of Bollen, Smith, and Whaley (2004). Followings the model overview, we present our results on the model estimation and analysis.

V.a. Model overview

To minimize potential confusion and facilitate cross-reference between their paper and ours, we follow the terminology and original notation of Bollen, Smith and Whaley (2004). Moreover, the assumptions of their paper are directly applicable: (1) risk-free rate and dividend yield are ignored, and (2) the BPB broker has no existing inventory. Bollen, Smith and Whaley (2004) use the following functional form for a (stock) spread:

$$SPRD_i = f(OPC_i, IHC_i, ASC_i, COMP_i), \qquad (1)$$

where[17]

---

[17] The following are the descriptions for each cost component as given by Bollen, Smith and Whaley (2004). *Order-processing costs* are those directly associated with providing the market making service and include items such as the exchange seat, floor space rent, computer costs, informational service costs, labor

i = identifier for a stock

OPC_i = order-processing costs

IHC_i = inventory-holding costs

ASC_i = adverse selection costs

$COMP_i$ = degree of competition

As discussed in Section IV, Stoll (1978b) uses a similar functional form given by (1). For this paper, we are unable to model the effect of competition on the trading cost of BPB baskets due to a lack of competition data. We have argued that the cost faced by a dealer is very similar to that of a BPB broker; therefore, we adopt the following functional form for the cost of trading a BPB basket:

$$(BPB\ basket\ trading\ Cost)_i = f(OPC_i, IHC_i, ASC_i) \quad (2)$$

where

i = identifier for each BPB basket in our data sample

Value inventory-holding premium using at-the-money option

---

costs, and the opportunity cost of the market maker's time. *Inventory-holding costs* are those a market maker incurs while carrying positions acquired in supplying investors with liquidity. *Adverse selection costs* arise from the fact that market makers, in supplying immediacy, may trade with individuals who are better informed about the expected price movement of the underlying security. *Degree of competition* is likely to affect the level of the market maker's bid/ask spread, particularly in an environment where barriers to entry are being slowly eliminated.

Bollen, Smith and Whaley (2004) argue that a dealer's spread "needs to include a premium to cover expected inventory-holding costs, independent of whether the trade is initiated by an informed or uninformed customer." In the context of a BPB basket, an analogous interpretation is that some trades within a BPB basket are liquidity-driven[18] and some others are information-driven.[19] If a basket's trades are purely liquidity-driven, (e.g. subscription or redemption of a S&P 500 index fund), the trade can be hedged easily using S&P 500 future contracts. A blind principal bid is unlikely to be used for this type of trading. Bollen, Smith and Whaley (2004) call this the inventory-holding premium (*IHP*). By using a European style at-the-money option to hedge stock price movement, they show that a dealer's expected *IHP* is equal to the following:

$$E(IHP) = S\left[2N\left(0.5\sigma E\left(\sqrt{t}\right)\right) - 1\right], \tag{3}$$

where

S = true stock price

σ = standard deviation of security return

t = time until next offsetting order

$N(\cdot)$ = cumulative unit normal density function

---

[18] One examples of a liquidity-driven trade is trades that trim back aggressive over-weight or under-weight positions (relative to a manager's benchmark) that are hitting the allowable upper or lower bound mandated by a portfolio owner.

[19] Since our data are from active asset managers, some of their trades are by definition supposed to be informed so that they can add value for their clients.

Equation (3) is, in fact, the value of an at-the-money option based on the Black-Scholes formula (with the risk-free rate equal to zero). The strike price is equal to the true stock price. The intuition of Bollen, Smith and Whaley's (2004) argument is that the value of the at-the-money option can be thought as the BPB broker's hedging cost so that he is protected from adverse stock price movements while he is holding a basket in his inventory.[20]

The following is an analogous interpretation of equation (3) in the context of trading BPB basket. S is the latest closing price[21] for each stock in a BPB basket. From a BPB broker's perspective, his profit or loss is calculated based on the closing prices. Therefore, closing price is the appropriate reference price for hedging purpose. σ is the standard deviation of rate of return for each stock in a basket. t is the time taken by a BPB broker to unload (a.k.a. unwind) a stock from a BPB basket. We introduce an extra variable that helps us model t. The new variable, unloading rate, g, is defined as the percentage of average daily (dollar) volume (ADV[22]) a BPB broker would like to trade (to unload a stock in a basket) during one trading day. For example, if average daily volume for a stock j (within a BPB basket) is \$1,000,000 and if g = 25% per day, then a BPB plans to trade \$250,000 worth of a stock each day to unload stock j. Further assume that the dollar trade size for stock j

[20] Please refer to the original paper for a more formal and mathematical argument for valuing the inventory-holding premium.

[21] For pre-open bidding, the latest closing price is the previous business day closing price. For post-close bidding the latest closing price is the same-day closing price.

[22] For the rest of the paper, we use the abbreviation ADV to stand for average daily (dollar) volume.

is \$500,000. Hence, it will take the BPB broker about two days to unloaded stock *j* from a BPB basket. Therefore, we have the following relation between *t* and *g*:

t = (dollar trade size for a stock within a basket / average daily volume) / *g* (4)

So, it takes two days for the BPB broker to unload this particular stock from a BPB basket. For half of all the shares for stock *j*, he needs options that expire in one day. For the other half, he needs options that expire in two days. For simplicity, we use the following definition of *t* in computing *IHP*:

t = 0.5 [(dollar trade size for a stock in a basket / average daily volume) / *g*] (5)

By computing *IHP* stock by stock (rather than *IHP* for a BPB basket as a whole), we do not require the unloading process to be done in a proportionally manner, so that the dollar weight for each stock within a basket remains stationary. As mentioned in the previous section, Stoll (1978b) does have this implicit limitation in the case of trading BPB basket. Therefore, Bollen, Smith and Whaley's (2004) methodology has an advantage in modeling the BPB basket trading cost. In summary, the expected *basket IHP* is the summation of each stock's *IHP* and for each stock its *t* is given by equation (5):

$$E(basketIHP) = \sum S_j \left[2N\left(0.5\sigma_j E\left(\sqrt{t_j}\right)\right) - 1\right], \tag{6}$$

where $j = 1, 2, \dots ,$ number of stocks in a BPB basket.

Bollen, Smith and Whaley (2004) have the following regression model specification for a dealer's spread:

$$SPRD_i = \alpha_0 + \alpha_1 InvTV_i + \alpha_2 MHI_i + \alpha_3 IHP_i + \varepsilon_i, \qquad (7)$$

where

$InvTV_i$ = inverse of trading volume = 1/number of shares traded

MHI_i = modified Herfindahl index

IHP_i = inventory-holding premium

As we mentioned in the preceding section, we do not have the necessary data to model competition among BPB brokers. It is unfortunate that we are forced to omit *MHI* in our BPB basket trading cost model specification. *InvTVi* is used to model order-processing cost. Bollen, Smith and Whaley (2004) argue that order-processing costs are largely fixed; hence, wthe order processing cost per shares goes down when share volume rises. However, the number of shares transacted in BPB baskets is much larger than that transacted by a dealer for a single stock.[23] Therefore, we define *InvTVi* slightly differently and re-define *InvTVi = 1 / sqrt(Total shares in a BPB basket)*. The following is a regression model specification for the BPB basket trading cost:

[23] In Table 3 of Bollen, Smith and Whaley (2004), the mean number of shares traded in a day for a stock is between 250,000 and 600,000 during three different sampling periods. However, the mean number shares traded for a BPB basket is about 11 million shares.

$$(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvTV_i + \alpha_2(Basket\ IHP)_i + \varepsilon_i, \qquad (8)$$

where

i = identifier for each BPB basket in our data sample

$(BPB\ basket\ trading\ cost)_i$ = fee (in dollars) paid by manager to BPB broker,

InvTVi = 1 / sqrt(Total shares in a BPB basket),

$(Basket\ IHP)_i$ = Basket IHP given by equation (6).

However, in the case of a BPB basket, there are two dimensions for order-processing cost. The first dimension is the total number of shares traded in a basket. The second dimension is the number of names in a basket. Another possible alternative regression model specification for BPB basket trading cost is:

$$(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvNumofNames_i + \alpha_2(Basket\ IHP)_i + \varepsilon_i, \qquad (9)$$

where

$InvNumofNames_i$ = 1 / Number of names being traded in a BPB basket

In the next section, we compare the estimation results for these two specifications (equation (8) and equation (9)). Since the number of shares and the number of names are positively correlated,[24] we prefer not to include both proxies in a single model

[24] In our data sample, the correlation between the two is 0.77.

specification. Including both proxies may introduce multicollinearity. Moreover, we prefer a more parsimonious model.

Bollen, Smith and Whaley (2004) argue that α_2 in equation 8 (or α_2 in equation 9 or α_3 in equation 7) should be equal to one.[25] They also prove that the expected *IHP* defined in equation (3) is approximately linear in the square root of t. Therefore, by adjusting the value of *g*, the unloading rate variable, we are able to have the following regression model specification for the BPB basket trading cost:

$$(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvTV_i + (Basket\ IHP(g))_i + \varepsilon_i, \qquad (10)$$

In summary, equation (10) is a regression model specification which uses an at-the-money option to value inventory-holding premium. The coefficient of $(Basket\ IHP(g))_i$ is calibrated to be equal to one. The estimation results for this model specification are given in next section. There is also some special interpretation of α_0. Bollen, Smith and Whaley (2004) argue that the intercept term represents the minimum tick size. In the case of trading a BPB basket, there is no minimum tick size. A BPB broker is free to submit any bid during a basket auction. Therefore, our prior is that α_0 will not be significantly different from zero.

Informed and uninformed trades in a BPB basket

[25] If we regard the *IHP* as hedging cost, there is no obvious reason for a BPB broker to over-hedge or under-hedge.

Bollen, Smith and Whaley (2004) also proposed the following interesting way of interpreting the inventory-holding premium. The major benefit and contribution of this interpretation is that it allows the decomposition of the inventory-holding premium into two components: (1) premium for uninformed trades and (2) premium for informed trades.

$$IHP = IHP_U + p_I(IHP_I - IHP_U) \tag{11}$$

where

IHP_U = inventory-holding premium for uninformed trades,

IHP_I = inventory-holding premium for informed trades,

p_I = probability of an informed trade.

Similar to Bollen, Smith and Whaley's (2004) methodology, we use the following Black-Scholes formula to compute the value of IHP_U and IHP_I when a name in a basket is a buy.[26] Basically the value of IHP_U and IHP_I is equal to the value of two slightly different call options.

$$IHP_{k,j} = S_{k,j} N\left(\frac{\ln(S_{k,j} / X_j)}{\delta_j \sqrt{t_j}} + 0.5\delta_j \sqrt{t_j}\right) - X_j N\left(\frac{\ln(S_{k,j} / X_j)}{\delta_j \sqrt{t_j}} - 0.5\delta_j \sqrt{t_j}\right) \tag{12}$$

[26]Equation (12) is the *IHP* calculation for one share. To calculate the *IHP* for a stock's trade within a basket, one needs to multiply equation (12) by the number of shares traded for that stock in a basket.

where

$j = 1, 2, 3, \ldots$, number of stocks that are purchased in a BPB basket

k = U or I. U stands for uninformed trade and I stands for informed trade

σ_j = standard deviation of security return

t_j = time for unloading a stock from a BPB basket

X_j = latest closing price for stock j + per share cost of BPB basket trading cost[27]

$S_{U,j}$ = latest closing price[28] for stock j

$S_{I,j} = (1 + q) \times X_j$[29] and $q > 0$

$N(\cdot)$ = cumulative unit normal density function

The following is a brief description of the intuition of $IHP_{U,j}$ and $IHP_{I,j}$. First, consider a uninformed buy of stock j ordered by an asset manager (i.e. a BPB broker is shorting stock j and needs to hedge the upward price movement of stock j), which corresponds to $IHP_{U,j}$. A manager buys stock j at price X_j. Since the trade is uninformed, we argue that the stock price will not significantly deviate from $S_{U,j}$. $IHP_{U,j}$ is the value of a slightly out-of-money call option that provides to the BPB broker protection when the price of

[27] During a BPB basket auction, competing BPB brokers usually submit their best bid in term of cents per share. Please refer to Section III for a numeric example. Conceptually, X_i can be viewed as the ask price quoted by a stock dealer in the case of a manager buying a single stock name.

[28] In the original argument in formulating IHP_U, $S_{U,I}$ is the stock's true price. However, in the context of BPB basket trading, a stock's latest closing price is used to calculate a BPB's profit and lost. Therefore, we assume the latest closing price as the stock's true price.

[29] We do not know the real stock price in the case of an informed trade. Like Bollen et al. (2004), we assume that the true price is q percentage above X_i.

stock j goes above X_j. One can think of $IHP_{U,j}$ as a hedging cost. In the case of an uninformed trade, the BPB broker is likely to obtain a profit under the assumption that the price of stock j is very unlikely to rise above X_j. The premium for an informed buy is $IHP_{I,j}$. Similarly, a manager buys stock j at price X_j. However, since the buy is informed, one can imagine the price of stock j jumping to $S_{I,j}$, which is above X_j, right after a BPB broker wins a basket. Therefore, the BPB broker is going to incur a loss on this informed trade. One can imagine the hedging cost ($IHP_{I,j}$) to be an in-the-money call option with strike price = X_j and stock price = $S_{I,j}$.

To calculate the $IHP_{k,j}$ for a stock in a BPB basket whose trade is a sell, $IHP_{k,j}$ is equal to the value of a put option where:

X_j = latest closing price - per share cost of BPB basket trading cost

$S_{I,j} = (1 - q) \times X_j$ and $q > 0$

To calculate the IHP_U or IHP_I for a BPB basket, we sum up stock level IHP_U and IHP_I. Using Bollen, Smith and Whaley's (2004) methodology, we have the following regression model specification to estimate the probability that a trade is informed:

$$(BPB\ basket\ trading\ cost)_i =$$

$$\alpha_0 + \alpha_1 InvTV_i + \alpha_2(Basket\ IHP_{U,i}(g))_i$$
$$+ \alpha_3(Basket\ IHP_{I,i}(g,q) - Basket\ IHP_{U,i}(g))_i + \varepsilon_i, \qquad (13)$$

where

g = unloading rate that we used in estimating the model given by equation (10),

q = a factor that links $S_{I,j}$ and X_j[30],

α_3 = probability of an informed trade.

We chose a value of g such that the estimated coefficient for the *Basket IHP(g)* in equation (8) (or equation (9)) is equal to one. We need to make some empirical assumption about the value of q. It is because we do not know the true price of a stock (it is assumed only the informed know the true price). For the model given by equation (13) to make sense, the estimated value of α_2 must not be significantly different from one and the value of α_3 must not be outside the range between zero and one (since α_3 is the estimated probability of an informed trade).

However, there is one consideration unique in the context of trading a BPB basket. It is very difficult to argue that every stock being traded in a basket is an informed trade. In terms of dollar trade size within a BPB basket, it is common that there are small trades that are uninformed (or at least that contribute relatively lower risk to the BPB brokers). This type of trade size distribution is often due to the fact that the trade list (i.e. a BPB basket) is generated by a portfolio optimizer.[31] For example, if a manager's portfolio is out of bounds for certain industry exposures then the optimizer tries to bring the exposures within a pre-specified lower or upper bound. We argue that IHP_I is equal to

[30] In Bollen, Smith and Whaley's (2004) paper, they use symbol k and we use symbol q in this paper.

[31] A trade list generated by traditional fundamental active manager usually has a lower number of names and the distribution of the dollar trade size is more concentrated.

IHP_U for this type of trade. Otherwise, IHP_I for a BPB basket is overstated. Therefore, we introduce a new parameter, *c*, called uninformed trade cutoff. This parameter is expressed in units of percentage of ADV. If the dollar trade size of a trade expressed as percentage of ADV is less than *c*, the IHP_I is set equal to IHP_U (i.e. these type of trades are assumed to be uninformed). Otherwise, IHP_I is computed based on equation (12).

We will use the following regression model specification for our empirical analysis:

$$(BPB\ basket\ trading\ cost)_i =$$

$$\alpha_0 + \alpha_1 InvTV_i + \alpha_2(Basket\ IHP_{U,i}(g))_i$$

$$+ \alpha_3(Basket\ IHP_{I,i}(g,q,c) - Basket\ IHP_{U,i}(g)\ )_i + \varepsilon_i, \qquad (14)$$

where

c = uninformed traded cutoff

V.b. Model estimation result and analysis

In this section, we discuss various estimation results of the Bollen, Smith and Whaley (2004) model using historical BPB basket trading data.

Choosing the proxy for order-processing cost

The first analysis is of the selection of proxy for order-processing cost. In the model overview, we suggested two possible proxies for order-processing cost:

1. *InvTVi* = 1 / sqrt(Total shares in a BPB basket *i*)

2. $InvNumofNames_i$ = 1 / Number of names being traded in a BPB basket i

To begin with, we arbitrarily set the unloading rate variable, g, to 25% of average daily volume per day[32] to calculate t (time taken to unload a stock from a basket) for each stock in a basket. We need t to compute the expected basket *IHP*. The estimation result for modeling the specification given by equation (8) and equation (9) is summarized in Panel A of Table IV. Panel B is a summary of descriptive statistics for the variables. Panel C is the correlation matrix. When taken separately, both proxy definitions perform quite well. Both have an expected positive sign and their estimated coefficients are significant. The estimation result when including both proxies in a single regression is shown at the bottom of Panel A. In this case, *InvTVi* becomes insignificant while $InvNumofNames_i$ continues to be significant. Therefore, we use $InvNumofNames_i$ as the proxy for the order-processing cost for the remainder of the analysis.[33]

Calibrating the unloading rate (g) and IHP as at-the-money options

The next step in our analysis is to calibrate the value of variable g, the unloading rate, for the regression model specification given by equation (15). This specification is based on equation (9) and the method of calibration is to choose a value of g such that the estimated coefficient α_2 is one (or not significant different from one).

[32] The value of g affects only the estimated coefficient of E(*basket IHP*). The two proxies for order-processing cost are independent of g.

[33] Analysis results using *InvTVi* as order-processing cost proxy are also available from us. We do not report these results here because $InvNumofNames_i$ continues to perform better in all other analysis.

$$(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvNumofNames_i + \alpha_2 (Basket\ IHP(g))_i + \varepsilon_i, \quad (15)$$

Bollen, Smith and Whaley (2004) show that *IHP* varies in an approximately linear fashion as the square root of t varies. As Panel A of Table IV shows, the estimated value of α_2 is 0.9647, which is slightly lower than one. To make α_2 equal to one requires a reduction the expected value of *Basket IHP(g)* by reducing t or increasing g. Equation (5) shows that t and g are inversely related. By adjusting the value of g from 25% of ADV to 26.86% of ADV[34], the estimated value of α_2 is calibrated to one. Table V shows the estimation result of equation (15). There is an empirical interpretation of g. The sample data implies that the BPB broker is unloading shares (of a stock) at a rate about 27% of ADV per day. If a dollar trade size of a stock in a BPB basket is less than 27% of ADV, the BPB broker will finish unloading that particular stock in the first day of trading after winning a BPB basket. Similar, if the trade size is 100% of ADV, BPB broker is going to take more than four days to unload the stock.

As Table V shows, the estimated value of α_2 is 0.9999 and is highly significant. The estimated coefficient for InvNumofNames, α_1, is positive and significant. These observations are consistent with our prior as described earlier. The R^2 of this regression is 74.91% and is comparable to that reported by Bollen, Smith and Whaley (2004).[35] However, we have a significant negative intercept in our model estimation. We mentioned earlier that we expect the intercept not to be significantly different from zero.

[34] Since $26.86 = 25/(0.9647 \times 0.9647)$

[35] R^2 reported by them ranges from about 50% to 80% during three different sample time periods.

There are two possible explanations for a negative intercept. The first is that the *IHP* is overstated when *IHP* is valued as at-the-money options in equation (15). It is because these at-the-money options protect against unfavorable price movement (e.g. an asset manager sold a stock to a BPB broker and the price of the stock subsequently goes down) and a BPB broker can profit from favorable price movement. Bollen, Smith and Whaley (2004) argue that such profit is capped at a certain level due to competition. This explanation is testable if we compute *IHP* as an option collar. Based on option collar approach, the value of *IHP* is the value of an at-the-money option minus the value of a slightly out-of-money option.[36] We report the test result for this explanation in the next section. The second possible explanation is that some of the trades within a basket are crossed with existing inventory carried by a BPB broker. *IHP* for crossed trades should be zero. Testing the second explanation is a more difficult task because data on BPB broker's inventory are usually not publicly available.

Inventory Holding Premium (IHP) as option collars

Following is the test result related to explaining the negative intercept described above. When computing the out-of-money option, we assume the strike price is 0.5% away from the stock price.[37] The model specification is also based on equation (15), but here the

[36] For a buy trade ordered by a manager, *IHP* = value of at-the-money call – value of slightly out-of-money put. In this case, *IHP* is modeled as buying a call and selling a put. Similarly, for a sell trade ordered by a manager, *IHP* = value of at-the-money put – value of slight out-of-money call. In this case, *IHP* is modeled as buying a put and selling a call.

[37] The 0.5% is an arbitrary number. Bollen, Smith and Whaley (2004) also use 0.5% for calculating the value of the out-of-money option.

inventory-holding premium is modeled as an option collar. Table VI shows the estimation result. The estimated coefficients for *InvNumofNames* and *Basket IHP* are both positive and significant. The result continues to be consistent with our prior belief. However, the intercept is still significantly negative. Therefore, the negative intercept may be due to the internal-crossing between stocks in a basket and a BPB broker's existing inventory. There appears, however, to be a weak point in this argument, as the R^2 of the regression is reduced. In summary, there is some weak evidence that the negative intercept might be attributed to internal crossing.

Probability of informed trades and distribution of various trading components

As discussed in the earlier section on an overview of the model, it is possible to estimate the probability that a trade is an informed trade. In this case, the model we estimate is given by equation (16), which is modified from equation (14).

$$\begin{aligned}(BPB\ basket\ trading\ cost)_i = \\ &\alpha_0 + \alpha_1 InvNumofNames_i + \alpha_2(Basket\ IHP_{U,i}(g))_i \\ &\quad + \alpha_3(Basket\ IHP_{I,i}(g,q,c) - Basket\ IHP_{U,i}(g))_i + \varepsilon_i, \qquad (16)\end{aligned}$$

When we estimate the model, we need to make some assumptions about the values of *q* and *c*. We assume *q* takes the following values: 0.5%, 1%, 2%, and 3%, while *c* takes the following values: 3% of ADV, 4% of ADV, and 5% of ADV. Table VII gives the estimation results for twelve different combinations of *q* and *c*. Panels A, B, and C correspond to three different values of *c*. The four rows of the sub-table represent four different values of q. There are twelve sub-tables in total. The estimated result will only have economic (and mathematic) interpretations if both of the following conditions are satisfied:

1. The estimated value of α_2 is not significantly different from one.
2. The estimated value of α_3 is not significantly outside the range between zero and one.

Many of our results are similar to those reported by Bollen, Smith and Whaley (2004). The estimated values of α_1, α_2, and α_3 are positive and significant. As q increases, the probability of informed trades (i.e. the estimated value of α_3) decreases. When q changes, the R^2 of the regression remains quite stable. This implies that the adverse selection component of the BPB basket trading cost appears to be relatively constant. The T-stat for α_2 is always the highest and most significant. This implies that inventory-holding cost is the most significant component of the trading cost.

However, there are some results that differ from those of Bollen, Smith and Whaley (2004). Using their argument, the intercept terms should not be significantly different from zero, because the intercept can be interpreted as minimum tick size. However, there is no such concept of minimum tick size in a BPB basket trading. Our result shows that the intercepts are not significantly different from zero in the usual statistical sense. However, the values of intercept are always negative. A possible explanation may be the possibility of internal crossing between trades in a basket and the existing inventory held by BPB broker. The lack of data on a broker's inventory makes further investigation of this explanation impossible. Due to a lack of data on BPB competition, we are unable to model the effect of competition among BPB brokers on the trading cost of BPB baskets.

Several results are unique in this study. When q is between 0.5% and 3%, BPB brokers begin to assume that a trade is informed when it reaches 3% or 4% of ADV. When q increases, c tends to increase as well. Therefore, when a BPB broker assumes a manager is informed, the broker would expect that informed trades are those with higher percentage of ADV. For a manager with higher skill, the BPB broker expects informed trades are likely to be those with a higher percentage of ADV. Based on Panel A and B, we can compute an average distribution of various cost components for the BPB basket trading cost. The average percentage of trading cost attributed to inventory-holding cost is about 61%, to adverse-selection about 34%, and to order-processing cost is about 23%. We are unable to find any other studies that analyze the distribution of various cost components when trading a basket of stocks. Therefore, it is difficult to compare other results to ours. Alternatively, we may compare our cost distribution with the cost distribution for trading a single stock name. However, the cost distribution for trading a single stock varies significantly among different studies. Stoll (1989) reports order-processing cost as the largest component (47%), followed by adverse-selection cost (43%) and inventory-holding (10%). Stoll's (1989) ranking of various cost components is exactly the opposite of our findings. Bollen, Smith and Whaley (2004) find that inventory-holding cost is the largest component, which corresponds to our results. But they find that adverse-selection cost is the smallest, which differs from our results. Huang and Stoll (1997) report that the biggest cost component is order-processing (62%) followed by inventory-holding (29%). The smallest cost component is adverse-selection (9%). Their ranking is also different from ours.

In an overall sense, inventory-holding cost and adverse-selection cost are more important in the case of trading a BPB basket. The rationale might be that a BPB broker commits relatively more capital than that in market-making for a single stock, and each BPB basket always has some informed trades. However, one might argue that the relative ranking of various cost components can change over time. For example, during a period of higher risk (i.e. in a market with higher cross-sectional dispersion), adverse selection cost may become the biggest cost component. We leave this question for further research.

VI. Conclusion

In this paper, we estimate two structural spread models for trading BPB baskets. We extend and improve upon the work done by both Kavajecz and Keim (2005) and Giannikos and Suen (2006) by providing a formal framework to model the trading cost of BPB baskets. This modeling involves using two spread models developed by Stoll (1978a and 1978b) and Bollen, Smith and Whaley (2004). The main contribution of this paper is the successful application of these models in estimating the cost of immediacy (i.e. spread) for trading a basket of stocks. We are also able to characterize some empirical behavior of BPB brokers. Based on our data sample, the unloading rate used by a BPB broker is about 27% of ADV. The more a broker thinks a manager's trade is informed, the bigger the mis-pricing and the bigger the (dollar) size for that trade in terms of higher percentage of ADV. The largest cost component when trading a BPB basket is found to be inventory-holding cost, followed by adverse-selection cost and order-processing cost. We also find some weak evidence that internal crossing used by the BPB broker who won a basket[38] helps reduce the overall BPB basket trading cost. However, we cannot model this effect of internal crossing formally due to a lack of data. For the same reason, we cannot model the effect of competition on the trading cost of BPB baskets. These are potential future research.

One possible application of this paper is to help asset managers establish a benchmark trading cost when trading BPB baskets. Managers can use this benchmark trading cost to judge the fairness of bids submitted by BPB brokers.

[38] Based on our informal discussion with BPB brokers, the range of crossing can range from 0% and up to 30% of a basket.

References

Almgren, Robert and Neil Chriss, 2003, Bidding Principal, Risk 16, 97-102

Amihud, Yakov and Haim Mendelson, 1980, Dealership Market: Market Making with Inventory, Journal of Financial Economics 8, 31-53

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

Benston, George J. and Robert L. Hagerman, 1974, Determinants of bid-ask spreads in the over-the-counter market, Journal of Financial Economics 1, 353-364

Biais, Bruno, Larry Glosten, and Chester Spatt, 2005, Market microstructure: A survey of microfoundations, empirical results, and policy implications, Journal of Financial Markets 8, 217-264

Bollen, Nicolas P.B., Tom Smith and Robert E. Whaley, 2004, Modeling the bid/ask spread: measuring the inventory-holding premium, Journal of Financial Economics 72, 97-141

Branch, Ben and Walter Freed, 1977, Bid-ask spreads on AMEX and the big board, Journal of Finance 32, 159-163

Choi, J.Y., Dan Salandro, and Kuldeep Shastri, 1988, On the estimation of bid-ask spreads: The theory and evidence, Journal of Financial and Quantitative Analysis 23, 219-230

Copeland, Thomas E. and Dan Galai, 1983, Information effects on the bid-ask spread, Journal of Finance 38, 1475-1469

Coughenour, Jay and Kuldeep Shastri, 1999, Symposium on market microstructure: A review of empirical research, The Financial Review 34, 1-28

Demsetz, Harold, 1968, The cost of transacting, Quarterly Journal of Economics 82, 33-53

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

Garman, Mark, 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 31, 71-100

Giannikos, Christos and Tin-shan Suen, 2006, Pricing determinants of blind principal bidding and liquidity provider behavior, working paper, Baruch College, CUNY

Glosten, Lawrence R. and Lawrence E. Harris, 1988, Estimating the components of the bid-ask spread, Journal of Financial Economics 21, 123-142

Glosten, Lawrence R. and Paul R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 14, 71-100

Hansch, Oliver, Narayan Y. Naik and S. Viswanathan, 1998, Do Inventories Matter in Dealership Markets? Evidence from the London Stock Exchange, Journal of Finance 53, 1623-1656

Hasbrouck, Joel, 2006, Empirical market microstructure: The institutions, economic and econometrics of securities trading, Oxford University Press, USA

Hasbrouck, Joel and George Sofianos, 1993, The trades of market makers: An empirical analysis of NYSE specialists, Journal of Finance 48, 1565-1595

Huang, Roger D. and Hans R. Stoll, 1997, The components of the bid-ask spread: a general approach, Review of Financial Studies 10, 995-1034

Ho, Thomas and Richard G. Macris, 1984, Dealer Bid-Ask Quotes and Transaction Prices: An Empirical Study of Some AMEX Options, Journal of Finance 30, 23-45

Ho, Thomas and Hans R. Stoll, 1980, On Dealer Markets Under Competition, Journal of Finance 35, 259-267

Ho, Thomas and Hans R. Stoll, 1981, Optimal Dealer Pricing Under Transactions and Return Uncertainty, Journal of Financial Economics 9, 47-73

Ho, Thomas and Hans R. Stoll, 1983, The Dynamics of Dealer Markets Under Competition, Journal of Finance 38, 1053-1073

Kavajecz, Kenneth A. and Donald B. Keim, 2005, Packaging liquidity: Blind acuctions and transaction efficiencies, Journal of Quantitative Analysis 40, 465-492

Kissell, Robert and Morton Glantz, 2003, Optimal trading strategies, AMACOM, New York, NY

Madhavan, Ananth, 2000, Market microstructure: A survey, Journal of Financial Markets 3, 205-258

Madhavan, Ananth, Matthew Richardson, and Mark Roomans, 1997, Why do security price change? A transaction level analysis of NYSE stocks, Review of Financial Studies 10, 1035-1064

Naik, Narayan Y. and Pradeep K. Yadav, 2003, Do dealer firms manage inventory on a stock-by-stock or a portfolio basis?, Journal of Financial Economics 69, 325-353

O'Hara, Maureen and George S. Oldfield, 1986, The microeconomics of market making, Journal of financial and quantitative analysis 21, 361-376

O'Hara, Maureen, 1997, Market microstructure theory, Blackwell, Cambridge, MA

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

Stoll, Hans R., 1978a, The supply of dealer services in securities markets, Journal of finance 33, 1133-1151

Stoll, Hans R, 1978b, The pricing of security dealer services: An empirical study of NASDAQ stocks, Journal of Finance 33, 1153-1172

Stoll, Hans, 1989, Inferring the components of the bid-ask spread: Theory and empirical test, Journal of Finance 44, 115-134

Stoll, Hans R, 2003, Market Microstructure, in Constantinides et al. ed.: Handbook of the Economics of Finance (North-Holland, Amsterdam)

Tinic, Seha M. and Richard West, 1972, Competition and the pricing of dealer services in the over-the-counter market, Journal of Financial and Quantitative Analysis 8, 1707-1727

Tinic, Seha M. and Richard West, 1974, Marketability of common stocks in Canada and the U.S.A.: a comparison of agent versus dealer dominated markets, Journal of Finance 29, 729-746

Table I
Descriptive statistics for some blind principal bid (BPB) basket characteristics

The following are the descriptive statistics for some of the BPB basket characteristics. There are 196 baskets used in this study.

Basket Characteristics	Mean	Std. Dev.	Min.	25th	Median	75th	Max.
Number of names in a basket	199	113	41	93	193	294	609
Total trade size ($ million)	333.99	239.56	20.09	146.20	268.09	487.05	1,188.14
Total number of shares (million)	11.26	8.27	0.60	5.05	8.95	16.29	45.06
Names that are buys (%)	45.59	9.34	13.14	40.06	46.24	50.73	100.00
Lowest bid (basis points)	53.57	33.00	8.95	26.90	44.79	73.42	164.75

Table II
Descriptive statistics for the variables in Stoll's (1978a, 1978b) model

The following table lists the variables we used in estimating Stoll's (1978a, 1978b) model. The table also summarizes the expected sign of the estimated coefficient for the independent variables. The sample size is 196. Following Stoll's original approach, the natural logarithm version of all variables is used in the model estimation. The descriptive statistics are based on the natural logarithm version of the variables. s_i = Basket trading cost which is the dependent variable. σ^2_i = Basket variance. V_i = Basket weighted-average (dollar) volume. $(V/T)_i$ = Basket weighted-average turnover. P_i = Basket weighted-average stock price.

Variable	Mean	Std. Dev.	Min.	25th	Median	75th	Max.	Expected Coefficient sign
s_i	-5.42	0.63	-7.02	-5.91	-5.41	-4.92	-4.11	Dep. Var.
σ^2_i	-4.34	0.86	-5.62	-5.02	-4.39	-3.85	-1.78	+
V_i	18.40	0.90	16.04	18.20	18.48	19.04	20.02	-
$(V/T)_i$	-6.45	0.88	-9.73	-7.02	-6.46	-5.80	-4.49	+
P_i	3.37	0.21	2.94	3.58	3.68	3.82	4.17	Ambiguous

Table III
Stoll's (1978a, 1978b) model estimation result

The following table summarizes Stoll's (1978a, 1978b) model estimation result. The sample size (N) is 196. Following Stoll's original approach, a natural logarithm version of all variables is used in the model estimation. s_i = Basket trading cost, which is the dependent variable. σ^2_i = Basket variance. V_i = Basket weighted-average (dollar) volume. $(V/T)_i$ = Basket weighted-average turnover. P_i = Basket weighted-average stock price.

	s_i	σ^2_i	V_i	$(V/T)_i$	P_i	intercept		**N**	R-sq
Coefficient	Dep. Var.	0.1347	-0.1660	0.4528	0.0565	0.9330		196	74.32
T-stat		4.49	-4.35	14.21	0.37	1.90			
p value		0.0000	0.0000	0.0000	0.7140	0.0587			

Table IV
Proxy for order-processing cost in Bollen, Smith and Whaley's (2004) framework.

Table IV has three panels. Panel A summarizes the model estimation result for the three different model specifications given below. Panel B shows the descriptive statistics of the variables. Panel C shows the correlation among variables. The unloading rate, *g*, is set at 25% of ADV when calculating basket IHP.

(1) $(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvTV_i + \alpha_2 (Basket\ IHP)_i + \varepsilon_i,$

(2) $(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvNumofNames_i + \alpha_2 (Basket\ IHP)_i + \varepsilon_i,$

(3) $(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvTV_i + \alpha_2 InvNumofNames_i + \alpha_3 (Basket\ IHP)_i + \varepsilon_i,$

where

i = 1, 2, 3, …, sample size of traded BPB baskets (identifier for each BPB basket),

$(BPB\ basket\ trading\ cost)_i$ = fee (in dollars) paid by manager to BPB broker,

$InvTV_i$ = 1 / sqrt(Total shares (expressed in million of share) in a BPB basket i),

$InvNumofNames_i$ = 1 / Number of names being traded in a BPB basket i,

$(Basket\ IHP)_i = \sum_j S_j[2N(0.5\sigma_j E(\sqrt{t_j})) - 1]$,

j = 1, 2, 3, …, number of stock names in BPB basket i,

S_j = latest closing price for stock j,

σ_j = standard deviation of return for stock j.

t_j = time taken to unload stock j from basket i which is a function of g,

$N(\cdot)$ = cumulative unit normal density function.

Panel A – Regression Results

	(BPB basket trading cost)$_i$	*$InvTV_i$*	*$InvNumofNames_i$*	*(Basket IHP)$_i$*	intercept		**N**	R-sq
Model estimation for specification (1)								
Coefficient	Dep. Var.	840,475.75		0.9789	-429,215.56		196	74.26
T-stat		1.97		20.51	-1.83			
p value		0.0499		0.0000	0.0683			
Model estimation for specification (2)								
Coefficient	Dep. Var.		41,081,072.62	0.9647	-389,989.57		196	74.91
T-stat			3.01	23.55	-2.45			
p value			0.0030	0.0000	0.0152			
Model estimation for specification (3)								
Coefficient	Dep. Var.	-300,690.52	48,528,485.53	0.9528	-308,924.29		196	74.94
T-stat		-0.46	2.29	19.62	-1.30			
p value		0.6451	0.0229	0.0000	0.1947			

Panel B – Descriptive Statistics

	Mean	**St. Dev.**	**Min.**	**25%**	**Median**	**75%**	**Max.**
(BPB basket trading cost)$_i$	1,672,476.90	1,857,896.00	45,777.13	608,587.02	1,217,757.30	2,060,928.90	15,557,582.00
$InvTV_i$	0.3835	0.1931	0.1490	0.2478	0.3835	0.4448	1.2890
$InvNumofNames_i$	0.0075	0.0052	0.0016	0.0034	0.0052	0.0107	0.0244
(Basket IHP)$_i$	1,753,569.90	1,662,649.00	55,917.13	636,577.97	1,366,714.50	2,297,741.80	13,058,043.00

Panel C – Correlation Matrix

	(BPB basket trading cost)$_i$	*$InvTV_i$*	*$InvNumofNames_i$*
$InvTV_i$	-0.43		
$InvNumofNames_i$	-0.16	0.77	
(Basket IHP)$_i$	0.85	-0.56	-0.34

Table V
Calibrating the value of *g*

By setting the value of *g*, the unloading rate, to 26.86% of ADV, the estimated value of α_2 in the following model specification is calibrated to one.

$$(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvNumofNames_i + \alpha_2 (Basket\ IHP(g))_i + \varepsilon_i,$$

where

i = 1, 2, 3, …, sample size of traded BPB baskets (i.e. identifier for each BPB basket),

$(BPB\ basket\ trading\ cost)_i$ = fee (in dollars) paid by manager to BPB broker,

$InvNumofNames_i$ = 1 / Number of names being traded in a BPB basket i,

$(Basket\ IHP(g))_i = \sum_j S_j[2\mathrm{N}(0.5\sigma_j \mathrm{E}(\sqrt{t_j})) - 1]$,

j = 1, 2, 3, …, number of stock names in BPB basket i,

S_j = latest closing price for stock j,

σ_j = standard deviation of return for stock j.

t_j = time taken to unload stock j from basket i which is a function of g,

$N(\cdot)$ = cumulative unit normal density function.

	(BPB basket trading cost)$_i$	*$InvNumofNames_i$*	*(Basket IHP)$_i$*	intercept		**N**	R-sq
Coefficient	Dep. Var.	41,078,664.54	0.9999	-389,918.73		196	74.91
T-stat		3.01	23.55	-2.45			
p value		0.0030	0.0000	0.0152			

Table VI
Model inventory-holding premium as option collars

In this model estimation, we model the inventory-holding premium, $IHP^C(g)$, as an option collar rather than an at-the-money option. When calculating the value of the out-of-money option, we assume the strike price (X) is 0.5% away from the latest closing price (S). The unloading rate, g, continues to be 26.86% of ADV.

$$(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvNumofNames_i + \alpha_2(Basket\ IHP^C(g))_i + \varepsilon_i,$$

where

i = 1, 2, 3, …, sample size of traded BPB baskets (i.e. identifier for each BPB basket),

$(BPB\ basket\ trading\ cost)_i$ = fee (in dollars) paid by manager to BPB broker,

$InvNumofNames_i$ = 1 / Number of names being traded in a BPB basket i,

$(Basket\ IHP^C)_i = \sum_j S_j[2N(0.5\sigma_j E(\sqrt{t_j})) - 1] - \sum_j$ (out of money option with X_j is 0.5% away from S_j),

j = 1, 2, 3, …, number of stock names in BPB basket i,

S_j = latest closing price for stock j,

σ_j = standard deviation of return for stock j,

t_j = time taken to unload stock j from basket i which is a function of g,

$N(\cdot)$ = cumulative unit normal density function,

X_j = strike price for the out of money option for stock j.

	(BPB basket trading cost)$_i$	*$InvNumofNames_i$*	*(Basket IHP^C)$_i$*	intercept		**N**	R-sq
Coefficient	Dep. Var.	82,245,474.84	5.4516	-887152.88		196	45.43
T-stat		3.68	12.27	-2.95			
p value		0.0003	0.0000	0.0036			

Table VII
Estimating the probability of informed trades and distribution of various trading cost components

This table summaries the estimation result of the following regression model specification given by equation (16):

$$(BPB\ basket\ trading\ cost)_i = \alpha_0 + \alpha_1 InvNumofNames_i + \alpha_2(Basket\ IHP_{U,i}(g))_i + \alpha_3(Basket\ IHP_{I,i}(g,q,c) - Basket\ IHP_{U,i}(g)\ )_i + \varepsilon_i,$$

where

$InvNumofNames_i$ = 1 / (Number of names being traded in a BPB basket i),

$IHP_{U,i}(g))_i$ = inventory-holding premium for uninformed trades,

$IHP_{I,i}(g,q,c)$ = inventory-holding premium for informed trades,

g = unloading rate defined in equation (10),

c = uninformed trade cutoff for $IHP_{I,i}$ defined in equation (14),

q is defined in equation (13).

During model estimation, we assume that q takes the values: 0.5%, 1%, 2%, and 3%, and that c takes the values 3% of ADV, 4% of ADV, and 5% of ADV. The table contains three vertical panels (A, B, and C), which correspond to three different values of c. For each panel, there are four estimation results (four sub-tables) that correspond to the four different values of q. Each sub-table has the following six columns:

1. Name of dependent and independent variables. Trade Cost = *(BPB basket trading cost)*$_i$ which is the dependent variable. InvNames = 1 / (Number of names being traded in a BPB basket i). IHPU = $Basket\ IHP_{U,i}(g)$ and IHPI = $Basket\ IHP_{I,i}(g,q,c)$.
2. Estimated values for the regression intercept and coefficient of the independent variables (i.e. α_0, α_1, α_2, and α_3).
3. T-statistics for each estimated values.
4. Mean value for the dependent and independent variables.
5. Mean trading cost contributed by each of the independent variables.
6. Trading cost contributed by each independent variable expressed as a percentage of the mean BPB basket trading cost.

Panel A

	c = uninformed trade cutoff = 3% of ADV					
	Var.	**Est. Coeff.**	**T-stat**	**Mean Val.**	**Cost Contri.**	
***q* = 0.5%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-333,136.75	-1.21	-333,136.75	-333,136.75	-20%
	InvNames	52,368,942.94	2.48	0.0075209	393,860.17	24%
	IHPU	0.8754	5.44	1,138,439.40	996,589.85	60%
	(IHPI-IHPU)	0.7885	2.48	780,207.93	615,193.95	37%
	R-sq = 51.00					
	estimated coefficient for IHPU is not significantly diff from 1 (t=0.77)					
***q* = 1.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-244,106.34	-0.88	-244,106.34	-244,106.34	-15%
	InvNames	47,866,879.42	2.25	0.0075209	360,000.72	22%
	IHPU	0.9656	6.18	1,138,439.40	1,099,277.08	66%
	(IHPI-IHPU)	0.2601	1.89	1,757,715.50	457,181.80	27%
	R-sq = 50.36					
	estimated coefficient for IHPU is not significantly diff from 1 (t=0.22)					
***q* = 2.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-209,995.89	-0.75	-209,995.89	-209,995.89	-13%
	InvNames	46,090,009.22	2.16	0.0075209	346,637.11	21%
	IHPU	0.9987	6.49	1,138,439.40	1,136,959.43	68%
	(IHPI-IHPU)	0.0990	1.67	4,029,030.40	398,874.01	24%
	R-sq = 50.16					
	estimated coefficient for IHPU is not significantly diff from 1 (t=0.01)					
***q* = 3.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-236,170.57	-0.85	-236,170.57	-236,170.57	-14%
	InvNames	47,405,547.96	2.23	0.0075209	356,531.11	21%
	IHPU	0.9759	6.31	1,138,439.40	1,111,003.01	66%
	(IHPI-IHPU)	0.0679	1.84	6,490,985.30	440,737.90	26%
	R-sq = 50.31					
	estimated coefficient for IHPU is not significantly diff from 1 (t=0.16)					

Panel B

	c = uninformed trade cutoff = 4% of ADV					
	Var.	**Est. Coeff.**	**T-stat**	**Mean Val.**	**Cost Contri.**	
***q* = 0.5%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-427,010.19	-1.62	-427,010.19	-427,010.19	-26%
	InvNames	56,661,845.96	2.77	0.0075209	426,146.55	25%
	IHPU	0.7446	4.61	1,138,439.40	847,681.98	51%
	(IHPI-IHPU)	1.1456	3.42	720,777.37	825,722.56	49%
	R-sq = 52.34					
	estimated coefficient for IHPU is not significantly diff from 1 (t=1.58)					
***q* = 1.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-344,589.39	-1.29	-344,589.39	-344,589.39	-21%
	InvNames	52,733,197.83	2.54	0.0075209	396,599.68	24%
	IHPU	0.8466	5.37	1,138,439.40	963,802.80	58%
	(IHPI-IHPU)	0.4055	2.78	1,619,197.90	656,584.75	39%
	R-sq = 51.39					
	estimated coefficient for IHPU is not significantly diff from 1 (t=0.97)					
***q* = 2.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-307,699.38	-1.14	-307,699.38	-307,699.38	-18%
	InvNames	50,894,526.69	2.44	0.0075209	382,771.27	23%
	IHPU	0.8904	5.74	1,138,439.40	1,013,666.44	61%
	(IHPI-IHPU)	0.1570	2.50	3,717,798.70	583,694.40	35%
	R-sq = 51.03					
	estimated coefficient for IHPU is not significantly diff from 1 (t=0.71)					
***q* = 3.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-329,615.91	-1.23	-329,615.91	-329,615.91	-20%
	InvNames	51,920,671.25	2.50	0.0075209	390,488.77	23%
	IHPU	0.8686	5.58	1,138,439.40	988,848.46	59%
	(IHPI-IHPU)	0.1037	2.66	6,004,184.80	622,633.96	37%
	R-sq = 51.23					
	estimated coefficient for IHPU is not significantly diff from 1 (t=0.84)					

Panel C

	c = uninformed trade cutoff = 5% of ADV					
	Var.	**Est. Coeff.**	**T-stat**	**Mean Val.**	**Cost Contri.**	
***q* = 0.5%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-559,354.83	-2.22	-559,354.83	-559,354.83	-33%
	InvNames	61,969,221.73	3.15	0.0075209	466,062.65	28%
	IHPU	0.5687	3.59	1,138,439.40	647,430.49	39%
	(IHPI-IHPU)	1.6767	4.79	666,984.83	1,118,333.46	67%
	R-sq = 54.83					
	estimated coefficient for IHPU is significantly diff from 1 (t=2.72)					
***q* = 1.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-487,923.43	-1.90	-487,923.43	-487,923.43	-29%
	InvNames	58,964,487.98	2.94	0.0075209	443,464.43	27%
	IHPU	0.6799	4.37	1,138,439.40	774,024.95	46%
	(IHPI-IHPU)	0.6312	4.09	1,493,870.50	942,931.06	56%
	R-sq = 53.48					
	estimated coefficient for IHPU is significantly diff from 1 (t=2.05)					
***q* = 2.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-452,102.41	-1.74	-452,102.41	-452,102.41	-27%
	InvNames	57,333,864.96	2.84	0.0075209	431,200.72	26%
	IHPU	0.7326	4.76	1,138,439.40	834,020.70	50%
	(IHPI-IHPU)	0.2504	3.76	3,431,949.50	859,360.15	51%
	R-sq = 52.89					
	estimated coefficient for IHPU is not significantly diff from 1 (t=1.74)					
***q* = 3.0%**	Trade Cost	Dep. Var.		1,672,476.90		
	Intercept	-469,808.59	-1.82	-469,808.59	-469,808.59	-28%
	InvNames	58,044,440.05	2.89	0.0075209	436,544.86	26%
	IHPU	0.7121	4.62	1,138,439.40	810,682.70	48%
	(IHPI-IHPU)	0.1612	3.91	5,553,416.00	895,210.66	54%
	R-sq = 53.16					
	estimated coefficient for IHPU is not significantly diff from 1 (t=1.87)					