Liquidity and the Law of One Price: The Case of the Futures/Cash Basis

by

Richard Roll, Eduardo Schwartz, and Avanidhar Subrahmanyam

August 24, 2005

Abstract

Deviations from no-arbitrage relations should be related to frictions associated with transacting; in particular to market illiquidity, because frictions impede arbitrage. Thus, financial market liquidity may play a key role in moving prices to fair values. At the same time, a wide futures/cash basis may trigger arbitrage trades and thereby affect liquidity. We test these ideas by studying the joint dynamic structure of aggregate NYSE market liquidity and the NYSE Composite index futures basis for a relatively long time-period, over 3000 trading days. We find that liquidity and the basis forecast each other in addition to being contemporaneously correlated. There is evidence of two-way Granger causality between the short-term absolute basis and effective spreads, and quoted and effective spreads Granger-cause longer-term absolute bases. These results are preserved after including a proxy for arbitrage financing costs, the Federal Funds rate, which bears an independent positive and significant relation with the short-term absolute basis. Impulse response functions indicate that shocks to the absolute basis predict future stock market liquidity. Overall, the evidence suggests that stock market liquidity enhances the efficiency of the futures/cash pricing system.

Contacts

	Roll	**Schwartz**	**Subrahmanyam**
Voice:	1-310-825-6118	1-310-825-2873	1-310-825-5355
Fax:	1-310-206-8404	1-310-206-5455	1-310-206-5455
E-mail:	Rroll@anderson.ucla.edu	Eschwart@anderson.ucla.edu	Asubrahm@anderson.ucla.edu
Address:	Anderson School UCLA Los Angeles, CA 90095-1481	Anderson School UCLA Los Angeles, CA 90095-1481	Anderson School UCLA Los Angeles, CA 90095-1481

We are grateful to Nihat Aktas, Yakov Amihud, Eric de Bodt, Laura Frieder, Esa Jokivuolle, Andrew Karolyi, Alexander Kempf, Olaf Korn, Feifei Li, Jun Liu and participants at the UCLA Alumni Association seminar in Tokyo for helpful comments and/or assistance with the preparation of the manuscript.

Liquidity and the Law of One Price: The Case of the Futures/Cash Basis

Abstract

Deviations from no-arbitrage relations should be related to frictions associated with transacting; in particular to market illiquidity, because frictions impede arbitrage. Thus, financial market liquidity may play a key role in moving prices to fair values. At the same time, a wide futures/cash basis may trigger arbitrage trades and thereby affect liquidity.. We test these ideas by studying the joint dynamic structure of aggregate NYSE market liquidity and the NYSE Composite index futures basis for a relatively long time-period, over 3000 trading days. We find that liquidity and the basis forecast each other in addition to being contemporaneously correlated. There is evidence of two-way Granger causality between the short-term absolute basis and effective spreads, and quoted and effective spreads Granger-cause longer-term absolute bases. These results are preserved after including a proxy for arbitrage financing costs, the Federal Funds rate, which bears an independent positive and significant relation with the short-term absolute basis. Impulse response functions indicate that shocks to the absolute basis predict future stock market liquidity. Overall, the evidence suggests that stock market liquidity enhances the efficiency of the futures/cash pricing system.

Introduction

The law of one price is a fundamental building block of modern financial theory. Indeed, several well-known pricing relations in finance depend on the notion that two traded or synthesized instruments with the same future cash flows should trade at the same price. In practice, of course, market microstructure aspects may cause a temporary deviation of prices from their no-arbitrage values. For example, extreme order imbalances in a cash market may create inventory problems for market makers and cause temporary deviations of cash prices from the corresponding no-arbitrage prices implied by derivative markets.

Finance scholars have long recognized that deviations from no-arbitrage relations are related to the frictions associated with transacting, in particular to (il)liquidity indicators such as the bid-ask spread. Thus, financial market liquidity may play a key role in moving prices to an appropriate level, where the futures/cash basis is zero. At the same time, a wide basis may trigger arbitrage trading, which may, in turn, affect liquidity. In this paper, we explore these ideas by examining the joint dynamic structure of the futures/cash basis and stock market liquidity.

Liquidity is increasingly attracting attention from traders, regulators, exchange officials and academics. Recent financial crises suggest that, at times, market conditions can be severe and liquidity can decline or even disappear.[1] Many microstructure studies are devoted to the measurement of and time variation in liquidity. The study of liquidity is often justified to a broader finance audience by arguing that liquidity shocks affect asset prices. Amihud and Mendelson (1986) and Jacoby, Fowler, and Gottesman (2000) provide theoretical arguments and empirical evidence to show how liquidity impacts stock returns in the cross-section, while Jones (2001) and Amihud (2002) show that liquidity predicts expected returns. Pastor and Stambaugh (2003) find that expected stock returns are cross-sectionally related to liquidity risk.

[1] "One after another, LTCM's partners, calling in from Tokyo and London, reported that their markets had dried up. There were no buyers, no sellers. It was all but impossible to maneuver out of large trading bets." - *Wall Street Journal*, November 16, 1998.

Our analysis of the link between the futures/cash basis and liquidity focuses on another potential reason why liquidity is relevant for financial theory. Specifically, we attempt to discern whether aggregate stock market liquidity affects and is affected by the futures/cash basis, and whether liquidity is related to deviations in the law of one price.

There have been many insightful previous studies of the basis[2] and various informal arguments have been offered about the relation between the basis and trading costs.[3] To the best of our knowledge, however, there has not been any previous attempt to dynamically relate the basis to concrete measures of endogenous trading frictions such as stock market liquidity over a comprehensive time-period.[4] Our empirical investigation involves the joint time-series of aggregate liquidity on the NYSE and the futures/cash basis associated with the NYSE Composite index futures contract.. A relatively long sample of over 3000 trading days (1988-2002) allows for the possibility of reliable conclusions.

The motivation for our study derives in part from the earlier literature on the futures/cash basis and liquidity. First, as Kumar and Seppi (1994) point out, arbitrage activities, and hence, the basis, may be affected by liquidity. Further, if the futures/cash gap induced by illiquidity is sufficiently wide, an illiquidity shock may have a lasting effect on the basis as arbitrageurs struggle to close the gap. In the reverse direction, market-wide order imbalances resulting from arbitrage trades in response to a wide basis may have a contemporaneous as well as a persistent impact on liquidity (Stoll, 1978, O'Hara and Oldfield, 1986, Chordia, Roll, and Subrahmanyam, 2002.) Following these ideas, our goal is to explore intertemporal associations between market liquidity and divergence of the futures and cash markets from their no-arbitrage relation.

---

[2] See, for example, Modest and Sundaresan (1983), Kawaller, Koch, and Koch (1987), Chung (1991), Klemkosky and Lee (1991), Subrahmanyam (1991), Bessembinder and Seguin (1992), Chan (1992), Sofianos (1993), Miller, Muthuswamy, and Whaley (1994), Yadav and Pope (1994), Neal (1996), Barclay, Hendershott, and Jones (2003), and Wang (2003).

[3] See for example, Bodie, Kane, and Marcus (1993, pp. 738-740).

[4] In an interesting paper, Bakshi, Cao, and Chen (2000) consider the frequency with which index option prices violate theoretical comparative statics. They do not, however, relate the liquidity of the cash market to these violations. Liquidity is key to arbitraging an index derivative as the cash instrument is a basket of many stocks. Bakshi, Cao, and Chen use data that span about three months; we consider the relation between the futures/cash basis and aggregate stock liquidity using about fifteen years of data.

It may be useful to cast the investigation in terms of a null and alternative hypothesis. Consider two scenarios: First, suppose that order flows in the cash and futures markets are virtually disconnected, and the basis is governed almost solely by quote updating on the part of market makers in the stock and futures markets. Second, assume that arbitrageurs do not affect stock liquidity and their trades are not sensitive to liquidity. In either of these scenarios, the basis would bear little or no relation with liquidity. This forms our null hypothesis. On the other hand, if arbitrageurs have significant price impact, then their trades should be sensitive to liquidity, and their trades should affect liquidity. This forms our alternative hypothesis.

Of course, there is no reason to restrict attention to a contemporaneous link between the basis and liquidity; longer-term linkages and feedback would also be of interest. Foundational work in microstructure by Stoll (1978) and O'Hara and Oldfield (1986) suggests that price pressures created by individual stock order imbalances can have a lasting impact on liquidity. Chordia, Roll, and Subrahmanyam (2002) empirically show that these inventory arguments apply to market-wide imbalances and market returns as well. The futures/cash basis, which is subject to order flow pressures induced by arbitrage trades, could therefore forecast future liquidity. In the reverse direction, suppose that the futures/cash gap widens considerably on a given day. Then the large volume of arbitrage trades required to close the gap may not happen instantly. This is because arbitrageurs may choose to spread their trades out over time to minimize price impact (Kyle, 1985), and such autocorrelated order imbalances would have persistent effects on liquidity.

Our analysis proceeds in two stages. First, we look for long-term relations between the basis and liquidity. To do this, we model the stochastic process of the basis, as in Brennan and Schwartz (1991), and then relate the mean reversion of the process to the average liquidity indicator over each contract's lifetime.

We then proceed to examine short-horizon (daily) time-series relations between the basis and liquidity. We do this by first adjusting the raw time-series of liquidity indicators and the absolute futures/cash basis to account for deterministic time-trends and calendar regularities (cf.

Gallant, Rossi, and Tauchen, 1992). Then, we apply vector autoregression to the adjusted series to uncover the dynamic interplay between liquidity and the absolute basis.

Our liquidity indicators are quoted and effective bid-ask spreads. While other measures of liquidity are potentially available, these measures have been the focus of much recent literature (e.g., Amihud and Mendelson, 1986, Chordia, Roll, and Subrahmanyam, 2000, Hasbrouck and Seppi, 2001), and are readily extractable over a long time period on a daily basis.[5]

Using a daily interval to analyze the joint dynamics of liquidity and the basis is, to some extent, arbitrary. Two justifications: first, dealer inventory considerations loom most significantly over rather short horizons, i.e., daily as opposed to weekly or monthly. Second, the daily horizon is long enough to reduce the influence of non-synchronous trading.[6]

The main results are:

- Over the lifetime of a futures contract, the basis mean-reverts faster when the market is more liquid.
- There is two-way Granger causality between the three-month absolute basis and liquidity, as measured by effective spreads.
- Shocks to absolute bases are significantly informative in predicting the next day's bid-ask spreads.
- Liquidity shocks forecast future shifts in the long-term basis, suggesting that liquidity affects arbitrageurs more in the relatively less active longer-term futures contracts.

In an attempt to identify other drivers of arbitrage costs that could potentially explain the above results, we consider the role of short-term interest rates. We find that the Federal Funds rate has a moderately significant and positive influence on the shorter-term bases. Overall, therefore, the analysis suggests that the efficiency of stock-futures pricing system is sensitive to the costs of

[5] Furthermore, Neal (1996, p. 547) points out that average buy and sell orders by index arbitrageurs average a modest 653 and 1385 shares, respectively, which are within the average quoted depth on the NYSE (Chordia, Roll, and Subrahmanyam, 2001). This is another justification for the use of our liquidity measures. Also note that the alternative liquidity proxies of Amihud (2002) and Pastor and Stambaugh (2003) are not readily interpretable as daily measures.

[6] We revisit this issue in Section V.

financing arbitrage trades. Our results on the bi-directional forecasting relationship between the absolute basis and liquidity, however, are preserved even after accounting for the effect of financing.

The evidence we provide adds to the growing literature on the dynamic determinants of liquidity (viz. Hasbrouck and Seppi, 2001, and Chordia, Roll, and Subrahmanyam, 2001). Our result that the absolute basis provides information about future liquidity suggests that shifts in the basis may change the aggregate cost of capital (viz., Amihud and Mendelson, 1986). Furthermore, our results suggest that liquidity plays a key role in moving futures and cash market prices towards the frictionless ideal of a zero basis.

I. Data

This section provides data sources and explains the methods used for computing the futures/cash basis and the liquidity indicators.

A. Computation of the Basis

Let F be the current index futures price, S be the cash stock market index, r the risk free rate for lending and borrowing over the remaining life of the contract, t the time to contract expiration, and δ the dividend yield over the contract's lifetime. As in MacKinlay and Ramaswamy (1988, MR), the absolute value of the relative index futures basis (henceforth, termed the "absolute basis") can be defined as

$$\frac{\left|Fe^{-(r-\delta)t} - S\right|}{S}.$$

In a frictionless world, the above quantity should be precisely zero. However, it is well known from earlier work (e.g., Brennan and Schwartz, 1991) that the basis exhibits considerable time-series variation.

The basis is constructed empirically with the following components: F is the closing futures price on the NYSE composite index futures contract, while S is the closing value of the NYSE composite index. The riskless rate r is the continuously compounded yield on a Treasury Bill maturing as close to the futures expiration date as possible, and the yield is extrapolated from the last day of the bill's life to the end of the contract. The dividend yield δ is the (continuously compounded) annual yield on the S&P500 index (obtained from CRSP) updated in January of every year.[7] The NYSE Composite Index futures contract expires on the third Friday in March, June, September, or December.[8]

Four different time-series of the bases can be constructed by starting with a contract having X months to maturity and rolling over every third Friday of March, June, September, and December into a successive contracts with the same original time to maturity, where X=3, 6, 9, or 12. We focus on the first three series, because the fourth is inactively traded (an average volume of about two contracts per day) and is not likely subject to arbitrage forces because of its lack of liquidity.[9] Henceforth, the three shorter series are called the three-month, six-month, and nine-month bases, respectively.

Figure 1 plots the three bases over the sample period. The basis in each case appears fairly stationary though the nine-month basis exhibit larger deviations from zero than the other two (notice the differences in scale of the vertical axes.)

---

[7] There inevitably is some error associated with the measurement of both the risk-free rate and the dividend yield. However, the errors should be small (in particular, the dividend yield on a basket such as the NYSE Composite Index should not fluctuate much from day to day). Further, it seems unlikely that any measurement error would be related to liquidity, our variable of interest. Nonetheless, to mitigate possible measurement error problems, we use the measured values of the above variables on the right-hand side of adjustment regressions prior to conducting vector autoregressions. Controlling for the measured values of these variables allows us to capture the component of the measurement error in the basis (the left-hand variable).

[8] All contracts officially expire on Friday. However, within the futures dataset (obtained from www.normanshistoricaldata.com), in some cases there was no trading on Friday and we substituted Thursday as the last day of trading.

[9] There is a slight asynchronicity between the hours of operation of the NYSE and the New York Board of Trade (where the futures contracts trade). Specifically, while both exchanges open at the same time (9:30am), the NYSE closes at 4 p.m., which is fifteen minutes prior to the close of the futures exchange. While this will undoubtedly introduce some measurement error in the basis, it should not affect our inferences because there is no reason to suspect a relation between the error induced by a slightly later closing of the *futures* market and aggregate *stock* market liquidity.

B. Stock Liquidity Data

An aggregate NYSE liquidity index on a daily basis for the period 1988-2002 is constructed as follows:

- To be included, a stock is required be present at the beginning and at the end of the year in both the CRSP and the intraday databases.
- To reduce the impact of firms entering the sample mid-year, if the firm changes exchanges from Nasdaq to NYSE during the year (no firms switched from the NYSE to the Nasdaq during the sample period), it is dropped from the sample for that year.
- Because their trading characteristics might differ from ordinary equities, certificates, ADRs, shares of beneficial interest, units, companies incorporated outside the U.S., Americus Trust components, closed-end funds, preferred stocks, and REITs are expunged.
- To avoid the undue influence of high-priced stocks, which have inordinately large spreads, if the price at any month-end during the year is greater than $999, the stock is deleted from the sample for the year.

Intraday data are purged for one of the following reasons: trades out of sequence, trades recorded before the open or after the closing time, and trades with special settlement conditions (because they might be subject to distinct liquidity considerations). A preliminary investigation reveals that auto-quotes (passive quotes by secondary market dealers) have been eliminated in the ISSM database but not in TAQ. This causes the quoted spread to be artificially inflated in TAQ. Since there is no reliable way to filter out auto-quotes in TAQ, only BBO (best bid or offer)-eligible primary market (NYSE) quotes are used. Quotes established before the opening of the market or after the close are discarded. Negative bid-ask spread quotations, transaction prices, and quoted depths are discarded. Each bid-ask quote included in the sample is matched to a transaction; specifically, following Lee and Ready (1991) any quote less than five seconds prior to the trade is ignored and the first one at least five seconds prior to the trade is retained.

We focus primarily on spreads that are not scaled by price. This is to avoid contaminating inferences by attributing movements in stock prices to movements in liquidity (recall that the stock price already forms the denominator of the absolute basis, which is based on MacKinlay and Ramaswamy, 1988). This procedure is further justifiable because we do not examine the cross-section of liquidity at a point in time, but are interested in the dynamic relation between liquidity and the absolute basis. We have verified, however, that our principal results are robust to using a proportional spreads (i.e., scaled by price).[10]

For each stock the following variables are calculated (these variables have also been the focus of earlier research on long-horizon liquidity; see e.g., Chordia, Roll, and Subrahmanyam, 2001):

- Quoted spread: the difference between the asked and the bid quote.
- Effective spread: twice the absolute distance between the transaction price and the mid-point of the prevailing quote.

An initial scanning of the intraday data reveals a number of anomalous records that appear to be keypunching errors. Filters are applied to the transaction data by deleting records that satisfy the following conditions:

1. Quoted spread>$5
2. Effective spread / Quoted spread > 4.0
3. Proportional effective spread / Proportional quoted spread[11] > 4.0
4. Quoted spread/Mid-point of bid-ask quote >0.4

These filters removed less than 0.02% of all stock transaction records. The spreads are averaged across the day to obtain stock illiquidity measures for each day. To avoid excessive variation in the sample size, stocks are required to have traded for a minimum of 100 days in a year to be included in the sample for that year. Days for which stock return data were not available from CRSP are dropped from the sample. The dates of October 25, 1989 and September 4, 1991 are also dropped from the sample because of transparent reporting errors. Specifically, on the

[10] It may be worth considering the use of futures liquidity data in addition to stock liquidity. Since there are no posted quotes on the futures market, futures liquidity is hard to estimate. Also, as the autocorrelation in futures price changes is small (see footnote 18 to follow), the Roll (1984) measure also is infeasible to implement. Since arbitrageurs have to execute trades in a basket of stocks (i.e., many stocks) simultaneously and incur liquidity costs in each individual security, aggregate cash market illiquidity should be the main impediment to arbitrage.

[11] The proportional spreads in condition 3 are obtained by dividing the unscaled spreads by the mid-point of the prevailing bid-ask quote.

former date, there are no data at all, and on the latter, ISSM recorded only quotes, but no transactions. In addition, data are not available for the period from September 11 to September 14, 2001 (neither for the stock liquidity measures nor for the basis) because of market closures and disruption following the events in New York City. The daily spread measures are averaged, value-weighted, across stocks (with weights proportional to market capitalizations at the end of the previous year) to obtain the aggregate market illiquidity measures.

Figure 2 plots these aggregate quoted and effective spreads. They exhibit a downward trend throughout the sample period. In addition, there are two significant drops corresponding to the reduction in minimum tick size, first from eighths to sixteenths and then to cents. Because of the non-stationary nature of these liquidity constructs we will adjust them for deterministic time trends, calendar regularities, and discrete reductions in the minimum tick size prior to using the adjusted series in vector autoregressions involving the basis variables.

Since deviations from the no-arbitrage relation can arise in either direction, the *absolute* basis is the most sensible object of enquiry. Table 1 presents summary statistics for the absolute bases and liquidity variables. In unreported t-tests, we find that all of the means are statistically significant at the 5% level or less. The mean value of the absolute basis across the three series, however, is only about 0.3%, which suggests that, on average, arbitrage activities are quite effective. Table 1 also reveals that both the mean and volatility of the absolute basis increase with the remaining term of the futures contract. The average value of the quoted spread is about sixteen cents. The quoted spread is more than five cents larger than the effective spread, indicating that many transactions take place within the quoted spread.

II. Preliminary Evidence

A simple measure of the liquidity-basis relation is the correlation between the average absolute basis and the average spread over a contract's lifetime. Such correlations are reported in Panel A of Table 2. The correlation between average effective spreads and the average absolute bases for the contracts corresponding to the three- six- and nine-month absolute bases is 0.293, 0.290, and 0.211, respectively. The corresponding correlations for the quoted spreads are 0.253, 0.260, and 0.201. All correlations are positive and exceed 20%. Further, the correlations for the three and six-month bases are highly significant. This represents preliminary evidence of commonality in the absolute basis and illiquidity.

Our second piece of preliminary evidence also relates the basis with average liquidity over a contract's lifetime, but in a slightly different way. We ask whether the speed of mean reversion in the stochastic process for the basis is faster when the market is more liquid. Consider the setting of Brennan and Schwartz (BS) (1991), who postulate a Brownian Bridge process for the basis:

$$dB(t) = -\frac{\mu B}{\tau}dt + \gamma dz,$$

where *B(t)* denotes the *signed* basis at time t, τ denotes the time to maturity, μ the speed of mean reversion, and γ the instantaneous standard deviation of the process. BS show how to estimate the parameters μ and γ using maximum likelihood techniques; details appear in Appendix A.

Panel B of Table 2 reports correlation coefficient between the estimated speed of mean reversion μ and the average value of the liquidity measure over contract lifetimes. As can be seen, in every instance the reversion parameter's estimate is negatively related to both quoted and effective spreads, and the correlation is highly significant for the six- and nine-month bases. This suggests that the signed basis mean reverts more quickly when the market is more liquid.

Though it supports a relation between the basis and liquidity, the evidence in Table 2 is long run oriented in the sense that employs averages over a contract lifetimes. We also want to consider dynamic and persistent short-term relations between the absolute basis and liquidity, including the possibility that they forecast each other. For these purposes, vector autoregressions (VARs) represent an appropriate technique. In large part, the remainder of the paper explains how VARs apply to our investigation and reports estimates obtained from VARs.

III. Adjustment Regressions and Initial Evidence on the Basis-Liquidity Relation

To perform vector autoregressions, it is desirable to first remove common regularities and trends from the time-series in order to mitigate the possibility of spurious conclusions. Prior research (Chordia, Roll, and Subrahmanyam, 2001) found that spreads exhibit time-trends and calendar regularities. It seems plausible that the absolute basis could also exhibit such phenomena. For example, the increased risk of holding positions over the weekend could cause higher absolute bases towards the end of the week. To alleviate such problems, we adjust the raw time-series for deterministic variations (see Gallant, Rossi, and Tauchen, 1992 for a similar approach to adjusting the series of trading volume). Since little is known about seasonalities or regularities in the stock-futures basis, this analysis is of independent interest.[12] Subsequently, (Section IV), innovations (residuals) from the adjustment regressions are related in a vector autoregression (VAR). To reiterate from Section I.C, the motivation for the VAR approach is that bivariate causality is likely. When the stock market is illiquid, arbitrageurs may have difficulty closing the gap between the two markets. In the reverse direction, a large inventory imbalance caused by arbitrage trades could result in illiquidity.

The following variables are used to adjust the raw basis series: (i) A Friday dummy to account for increased costs of holding positions through the weekend, (ii) 11 calendar month dummies for January through November, (iii) a dummy for days prior to major holidays; Thanksgiving, Christmas Eve, and July 4, (iv) for the three-month basis, four dummies for the four days prior to expiration, to control for any maturity-related effects (v) a time trend and the square of the time trend to remove any long-term trends[13], and (vi) the difference between the risk-free rate and the dividend yield to account for any errors in the measurement of these quantities. For spreads, explanatory variables include day-of-the-week, month-of-the-year, and holiday dummies, linear and quadratic time trends, and two dummies to account for the post sixteenth period (i.e., the period including following the shift in minimum tick size from 1/8 to 1/16 on June 24, 1997) and the post-decimalization period (on and after January 29, 2001.)

---

[12] We discuss the effect of the adjustment procedure on our results in Section V.

[13] The time trends are orthogonalized; i.e., the linear time index t goes from –1 at the beginning of the sample to 1 at the end of the sample, whereas the quadratic term equals $(3\ t^2-1)/2$.

Tables 3 and 4 present the regression coefficients from the adjustment regressions for the absolute basis and spreads, respectively. The absolute basis exhibits strong maturity-date related effects. The maturity effect is presumably due to traders unwinding large positions just prior to expiration. The difference between the risk-free rate and dividend yield is positive and significant for the six and nine-month contracts, but negative and marginally significant for the three-month contract. We expect this variable to correct for the impact of measurement error in the borrowing/lending rate of arbitrageurs as well as the measured dividend yield over the life of the contract.

Calendar time effects also are discernible in that the absolute basis tends to be lower in late summer. This phenomenon is to be understood in conjunction with the finding that trading volume also tends to be abnormally low in August (see Gallant, Rossi, and Tauchen, 1992). The secular regularity of lower trading volume in late summer could result in smaller price pressures, and consequently, smaller deviations of the basis from zero.

Note that the coefficients for month-of–the-year regularities in the adjustment regression are usually more significant and of larger absolute magnitudes for the longer-term bases. Indeed, the August coefficient for the nine-month basis is about double that for the three-month basis, and is also more significant. Further, the explanatory power of the regression is also larger for the longer-term bases. This observation makes intuitive sense; since contracts with longer times to maturity would tend to be less liquid; one would expect more predictable deviations of the basis from zero for such contracts.[14]

Table 4 reveals that quoted spreads are lower on Monday through Wednesday and are also lower in most months relative to January (the omitted month.)[15] Higher spreads in January may be due to imbalances created by reallocations to speculative stocks by institutional money managers after window-dressing in December. Positive order flow from individual investors' buying activity after year-end cash inflows such as bonuses may also contribute to the "January" effect

[14] The average trading volumes for the three-, six- and nine-month bases are respectively 3037, 455, and 14. These numbers are all statistically different from each other at the 5% level, and clearly suggest decreasing contract liquidity with increasing maturity.

[15] The coefficients are negative in all months (relative to January) and are significant except for February and April.

in spreads. The effect of early fall on spreads is large; in absolute terms, the coefficient for the quoted spread in September is about six times as large as the coefficient in April; the corresponding ratio for the effective spread is even larger. These patterns are intriguing and warrant further analysis in future research.

There was a significant reduction in spreads after the minimum tick size was reduced from 1/8 to 1/16 and a further reduction after decimalization. In addition, spreads exhibit a strong secular downtrend. The behavior of the quoted and effective series is similar, and the explanatory power of the spread regressions is high.

The augmented Dickey-Fuller test strongly rejects the existence of a unit root for all five adjusted time-series (three for the absolute basis, and two for spreads) at p-values less than 0.001. Hence, there is no evidence that the adjusted series are non-stationary. For the remainder of the paper, the adjusted series are used in all tests, and references to the original variables actually refer to the adjusted values.

Table 5 presents a correlation matrix for the three adjusted absolute bases (denoted ABAS3, ABAS6, and ABAS9) and the two adjusted spreads, quoted and effective (QSPR and ESPR). All correlations are statistically significant at the 5% level, except that between ABAS9 and QSPR. The three absolute bases series are highly correlated. In addition, correlations of the bases with spreads are all positive, suggesting commonality. Correlations of the absolute bases are higher with effective spreads than with quoted spreads, suggesting that the effective spread, which accounts for transactions executing within the quotes, has a stronger link with the absolute basis.

As a pre-amble to the analysis of whether shocks to the absolute basis have a lasting effect on liquidity and vice versa, Table 6 presents coefficients from OLS regressions of the absolute bases on one lag of the spreads and vice versa. The results indicate that the coefficient of the lagged effective spread is highly significant in explaining all three absolute bases.[16] The lagged quoted spread, however, is significant only for the three-month absolute basis.

In the reverse regression of spreads on lags of the absolute bases, all three absolute bases are significant in explaining the current effective spread and lags of the two shorter-term absolute bases are significant in explaining the quoted spread. In totality, the results of Tables 5 and 6 point to the notion that the absolute bases and liquidity are jointly determined, and that effective spreads (which account for trades executing within the posted quotes) bear a stronger relation to the absolute bases. We now turn to vector autoregressions, which allows for a richer dynamic structure between the bases and liquidity measures.

IV. Vector Autoregressions

The input data for VAR estimation are the adjusted series (i.e., the residuals) from the first-stage regressions described in the previous subsection.[17] The number of lags is determined by the Akaike and Schwarz information criteria. When these two criteria indicate different lag lengths, the lesser lag length is chosen for the sake of parsimony. Typically, the slopes of the information criteria (as a function of lag length) are quite flat for longer lags, so the choice of shorter lag lengths is further justified. Six bivariate VARs are estimated, pairing each of the three absolute

---

[16] An issue is whether the small sample bias discussed by Stambaugh (1999) and Amihud and Mendelson (2004) is a potential concern, since the right-hand side variables in Table 5 are persistent. However, there are more than 3700 observations, suggesting that this is not likely a serious problem. In fact, using the Amihud and Mendelson (2004) correction leaves the Table 5 results virtually unchanged; results are available upon request.

[17] The reader may wonder to what extent bid-ask bounce affects the results. This phenomenon pertains to individual assets and is unlikely to be pronounced for a value-weighted index consisting of over 1500 stocks. Also, index futures contracts are extremely liquid instruments, so bid-ask effects should be small. This is supported empirically by the daily first order autocorrelation in the three-month futures price change series (adjusted for the effects in Table 2), which is only –0.021 (p=0.21). The corresponding autocorrelations for the six and nine-month series are also statistically indistinguishable from zero. Later results indicate that the lagged absolute basis even predicts market depth (with or without controlling for the spread), and there is no reason why depth should be influenced by bid-ask bounce. Also, it is known that liquidity is low on days of heavy selling (Chordia, Roll, and Subrahmanyam, 2001), suggesting the possibility that low liquidity may be accompanied by a artificially high absolute basis due to trades hitting the bid in the index. In this case, however, the basis should tend to narrow the next day as prices bounce back. There is no evidence that the absolute basis decreases after days of high spreads, nor is there any evidence that the basis is significantly wider on days with low (highly negative) returns. Finally, in many cases, the basis has a persistent impact on spreads, and vice versa, a finding unlikely to be related to bid-ask phenomena. It appears that bid-ask bounce is not material here.

basis measures (three-, six-, and nine-months) with the two liquidity measures (quoted and effective spreads). The information criteria imply a lag length of six days for all VARs.

Panel A of Table 7 reports pairwise correlations of innovations (residuals) from each of the six bivariate VARs. They are all positive with magnitudes ranging from about 0.14 to 0.07 and they are all statistically significant at the 5% level. They are slightly higher for effective spreads, again suggesting that effective spreads are more relevant estimates of arbitrage costs. Overall, this panel represents further evidence that illiquidity is positively associated deviations from zero in the futures/cash basis; alternatively, that high stock market liquidity facilitates narrowing of the basis.

Table 7 also reports pairwise Granger-causality tests. For the null hypothesis that variable i does not Granger-cause variable j, we test whether the lagged coefficients of i are jointly zero when j is the dependent variable in the VAR. In Panel B of Table 7, the cell associated with the i'th row variable and the j'th column variable shows the Chi-square statistic associated with this test.

The three-month absolute basis Granger-causes quoted and effective spreads. Reverse causality running from both spreads is found for the nine-month basis and from the effective spread to the three-month basis. The six-month absolute basis Granger-causes the effective spread, but not the quoted spread. The spread measures do not Granger-cause the six-month absolute basis.

The evidence from Panel B of Table 7 is consistent with liquidity concerns being particularly relevant for arbitrageurs in longer-term, relatively less-active contracts.[18] The results also reveal that absolute basis innovations for shorter-term contracts have a stronger effect on liquidity, perhaps because arbitrage activities are more intense in such contracts.[19] Also, as in Table 5, the results depict stronger relations with the effective spread, which accounts for trading costs across both large and small orders.

[18] cf. Footnote 14.

[19] The notion that market frictions can affect the efficacy of arbitrage is suggested in Mitchell, Pulvino, and Stafford (2002), Lamont and Thaler (2003), Hou and Moskowitz (2004), and Sadka and Scherbina (2004). We explore this

However, not all the results in Table 7 lend themselves to easy interpretation; e.g., the effective spread Granger-causes the three- and nine-month absolute bases, but not the six-month one. But Granger causality is based on a single equation. A clearer picture can potentially be provided by impulse response functions (IRFs), which account for the full dynamics of the VAR system.

An IRF traces the impact of a one-time, unit standard deviation, positive shock to one variable (henceforth termed simply a "shock" or "innovation" for expositional convenience) on the current and future values of the endogenous variables. Since the innovations are correlated (as shown in Panel A of Table 7), they need to be orthogonalized, so we use the inverse of the Cholesky decomposition of the residual covariance matrix to orthogonalize the impulses. Results from the IRFs are generally sensitive to the specific ordering of the endogenous variables.[20] However, conclusions about our IRFs turn out to be insensitive to the ordering, and also robust to the computation of generalized impulse responses (Pesaran and Shin, 1988). In the following results, the absolute basis variables are always placed first in the ordering.

Figures 3 through 8 depict responses of the liquidity and basis measures to a unit standard deviation shock in a particular variable traced forward over a period of six days (i.e., the lag length of the VAR). Monte Carlo two-standard-error bands (based on 1000 replications) are provided to gauge the statistical significance of the responses. Period 1 in the IRFs represents the contemporaneous response, whereas subsequent periods represent lagged responses. The vertical axes are scaled to the measurement units of the responding variable. The impulse responses generally decay over time, confirming the unit root tests that indicate stationarity.

It can be seen that shocks to a variable are informative in predicting future values of that same variable in every instance. This confirms that both spreads and absolute bases are persistent. With regard to cross-effects, the IRFs are largely consistent with the Granger causality results. Innovations in the three-month absolute basis have a lasting and significantly positive effect on both spread measures (Figures 3 and 4) thereby indicating that the absolute basis forecasts future stock market liquidity. Specifically, the results are consistent with an interpretation that if the

---

notion as well as the reverse: viz., that exploitation of arbitrage opportunities can impact an endogenous market friction, liquidity.

[20] VAR coefficient estimates (and, hence, the Granger causality tests) are unaffected by the ordering of variables.

basis widens on a particular day, arbitrage forces on subsequent days bring an increase of either buy or sell orders, which strains liquidity. The impulse responses of spreads to absolute bases are more persistent for the three- and six-month contracts than for the nine-month contract, again suggesting more intense arbitrage activity in shorter contracts.[21]

Impulse responses of the three-month and six-month absolute bases to spreads are not significant. Thus, even though the spread Granger-causes the shorter-term absolute bases, after accounting for the joint dynamics, including the persistence of the absolute basis and liquidity variables, shocks to spreads are statistically uninformative in forecasting these bases. Note, however, that for the nine-month case, the spread is indeed informative in predicting shocks to the absolute basis at the second period, (Figures 7 and 8.) This is consistent with our previous interpretation of the Granger causality results, that liquidity affects the future activities of arbitrageurs relatively more in the less liquid, longer-term contracts.

With respect to the economic significance of the IRFs, a one-standard deviation shock from the three-month absolute basis impacts effective spreads and aggregates to an annualized extra trading cost of $2.7 million for a daily round-trip trade of one million shares in the basket of NYSE-listed common stocks. Further, in the case of the nine-month basis measure, for an average stock price of $40 and a trade of one million shares, a one standard deviation shock to spreads results in an annualized cumulative divergence of $1 million between the futures and its cash value.[22] Thus, the economic relevance of both effects appears to be material.

[21] Indirect evidence that arbitrage activity follows days on which the absolute basis widens can be obtained from examining order flow. Harford and Kaul (2005) document a significant index-related component in stock order flows. We do not include order flow in our VAR system because of possible multicollinearity between absolute order flow and the absolute basis. However, using an (imperfect) index of order flow based on Chordia, Roll, and Subrahmanyam (2002), who, in turn, use the Lee and Ready (1991) signing algorithm, we find that the univariate correlation between current net buying activity and the lagged (signed) three- six- and nine-month bases are 0.069, 0.067 and 0.065, respectively, which are all statistically significant at the 1% level. Thus, there is evidence that net buying activity in the cash market follows days on which the futures component of the basis is high relative to the cash component; this is consistent with arbitrageurs reacting to the size of the basis.

[22] This calculation of economic significance is based on the six-day cumulative impulse response of one variable to a one-standard deviation shock in the other. In the first case (the response of liquidity to the absolute basis), taking the total incremental trading cost per million shares traded and multiplying by the number of trading days in a year (250) yields the dollar amount we report. In the second case, the cumulative impulse response over six days is multiplied by the dollar value of the trade, and then by 250, the approximate number of trading days in a year, to arrive at the reported number.

V. Discussion and Robustness Checks

The first potential issue involves our procedure to adjust the bases and spreads for calendar regularities and secular trends. This procedure is intended to remove possible spurious covariation, which could have unintended consequences. It could weaken inferences by capturing common regularities in the absolute basis and spreads that genuinely represent economic causation from one variable to another.[23] The most direct way to investigate this concern is to recompute the vector autoregressions using the unadjusted series. For brevity, we do not report results using the unadjusted series, but they are qualitatively unaltered in every respect, save one, that unlike the result in Panel B of Table 7, the quoted spread Granger-causes the six-month basis. Thus, the adjustment procedure appears to be conservative and has minimal impact on our inferences.

To ascertain robustness to alternative measures of liquidity, we present abridged results using a measure of market depth, which is positively associated with liquidity. This measure is calculated by averaging the posted bid and ask depths accompanying each bid and ask quote, then the mean value over the trading day for each stock, and then value-weighting the means across stocks. We adjust the resulting variable using the variables in Table 4. Panels A and B of Table 8 present the regressions corresponding to Table 6 that replace adjusted depth as a measure of liquidity, while Panel C reports the Granger causality tests involving this measure.

In every instance, the regression coefficients are negative and significant, which is reassuring, since depths are positively related to liquidity (whereas the reverse is true of spreads). The Granger causality results in Panel C of Table 7 indicate that depth and the three-month absolute basis Granger-cause each other, depth Granger-causes the nine-month absolute basis, and the six-month absolute basis Granger-causes market depth. These results are largely consistent with

[23] An example of possible economic causation that the adjustment removes is the following: The mean values of the three-, six-, and nine-month absolute bases decreased, respectively, from 0.191 to 0.150, 0.318 to 0.160, and 0.458 to 0.197 from the pre- to the post-decimal tick size period. All of these decreases are statistically significant at the 1% level. While it is tempting to attribute these changes to a secular increase in liquidity, other events (e.g., technological innovations speeding up execution of arbitrage trades), could also have played a role.

Table 7, suggesting that previous inferences are largely unaltered by using of depth as an alternative measure of liquidity.[24]

Fluctuations in liquidity can arise either from demand-related shifts (e.g., shocks to investors' trading needs) or supply-related ones (e.g., shifts in borrowing terms which affect market makers' access to capital). While a detailed analysis of the causes of daily liquidity fluctuations warrants a separate paper, previous literature suggests that the most important candidates for common determinants of liquidity and the absolute basis are return volatility as well as the signed return (see, for example, Benston and Hagerman, 1974, MacKinlay and Ramaswamy, 1988, and Hasbrouck, 1991). As such, it is of interest to consider the extent to which our results are driven by return and volatility dynamics.

To address this issue, we perform a robustness check by including six lags of return and index volatility in the VAR system.[25] Again, while the estimates from this exercise are not reported for conciseness, our conclusions on commonality between the absolute basis and liquidity are largely unaltered. Conclusions about all three aspects of the VAR, namely, innovation correlations, Granger causality results, and the impulse response functions, remain unchanged. Thus, volatility and returns alone do not capture the common dynamics of the absolute basis and liquidity. This is consistent with the notion that changes in liquidity-related factors other than returns and volatility, such as shifts in market makers' ability to supply liquidity, and shifts in investors' liquidity needs, play an important role in the effectiveness of arbitrage activity.

Arbitrage positions often require financing and thus may depend on short-term interest rates. Further, liquidity may also depend on dealer financing costs (Demsetz, 1968). For example, by reducing the cost of margin trading and decreasing the cost of financing inventory, a decrease in short rates could stimulate trading activity and increase market liquidity. To ascertain whether our results are proxying for this common determinant of trading costs, and to gain a greater

[24] As per the Schwarz criterion, six lags are used in the depth VARs. The impulse response functions are not presented for brevity but are consistent with the Granger causality results. Since spreads and depth are related, there is a possible multicollinearity problem, but the results for depth turn out to be largely insensitive to whether or not spreads are included in the VAR. (Spreads are not included in the reported VARs.)

understanding of the economics of arbitrage, we use two measures of short-term interest rates as exogenous variables in the VAR system. The first measure is the Federal Funds rate, which is related to day-to-day financing of inventory or arbitrage positions.[26] The second is the constant-maturity (six-month) Treasury Bill rate, which is related to the financing of longer-term positions.[27] Prior to their usage in the regressions, these variables are adjusted for weekly and monthly seasonals, holidays, and linear and quadratic trends.

The coefficients on the Federal Funds rate as well as the Granger causality results are presented in Table 9. Inclusion of the short rates does not change the qualitative conclusions on Granger causality relative to Table 7.[28] While the nine-month basis is not related to any of the short rates, we find that the Federal Funds rate is significantly and positively related to the three and six-month absolute bases. This result suggests that arbitrage activity depends on financing costs, which, in turn, depend on short-term interest rates. The liquidity variables are not significantly related to any of the short-term interest rates. Overall, the results in Table 9 suggest that our previously reported results cannot be attributed to an exogenous interest rate influence driving both the basis and liquidity. Doubtless, many phenomena, including news arrivals that cause changes in agents' expectations, and unexpected wealth shocks, drive the dynamics of liquidity, and these drivers warrant separate research.

Our final robustness check involves the effects of possible non-synchronous trading in the constituent stocks of the index. The NYSE Composite index is value-weighted, so the impact of non-synchronous trading at a daily horizon should be fairly small; the first order return autocorrelation for the NYSE Composite Index is a modest 0.045, yet a potential concern is that wider bases may be associated with greater non-synchronous trading (and less liquidity.) In investigate this possibility, we regressed index returns on lagged index returns <u>and</u> the lagged

---

[25] Index return and index volatility are obtained respectively as the residual and absolute value of the residual from a regression of the index return on the past twelve lagged returns and four dummies for days of the week (see Schwert, 1990, Jones, Kaul, and Lipson, 1994, and Chan and Fong, 2000).

[26] Interest rates are obtained from http://www.federalreserve.gov/releases/h15/data.htm.

[27] Using a constant maturity three-month Treasury Bill rate instead of a six-month rate leaves the conclusions qualitatively unaltered.

[28] VARs involving depth are omitted to conserve space, but they are consistent with those for quoted and effective spreads. Also, impulse response functions are qualitatively unaltered from those in Figures 3 through 8.

adjusted absolute basis and found no evidence that serial dependence is greater when the basis is wider.

A related issue is that high absolute return days may be accompanied by high spreads, high non-synchronous trading and a wider basis as the futures price adjusts faster than the cash index to new information; but this possibility conflicts with our forecasting results, which are unaltered after controlling for volatility.

Nonetheless, to further address any possible influence of non-synchronous trading, we adopt the approach of Jokivuolle (1995), who devises a method for estimating the true value of an index when some of its constituents are subject to stale pricing. He shows that the log of the true index value can be represented as the permanent component of a Beveridge and Nelson (1981) decomposition of the series of log index first differences. Jokivuolle uses the Russell 2000 as an empirical example, but this index includes many small stocks so the non-synchronous trading problem is more severe than in our case; this is indicated by his first-order autocorrelation of 27%, compared to ours of about 5%. His technique[29] reduces the autocorrelation of the "true" NYSE composite index returns to 0.006. Details appear in Appendix B.

Using the "true" NYSE composite index to compute the bases, Figure 9 shows the response of effective spreads to the three-month absolute basis, and the response of the nine-month basis to effective spreads. As can be seen, the impulse responses are very similar to their respective counterparts in Figures 4 and 8. Other contracts are not presented for brevity, but they also are qualitatively unaltered when using the "true" index. Non-synchronous trading does not have a material impact on our results.

[29] Jokivuolle (1995) uses an MA(1) process to model the first differences of the log index; we find that the same choice is appropriate here as well, since, as in his paper, the autocorrelations beyond the first lag (up to a week of "meaningful" lags) are quite small (0.03 or less) in magnitude.

VI. Conclusions

Because liquid markets facilitate arbitrage trades, relative pricing in the cash and futures markets should be more efficient when liquidity is high. On the other hand, arbitrage trades themselves could reduce liquidity by creating order imbalances. To study these ideas, we use three futures/cash basis series involving contracts maturing within three, six- and nine-months, along with two measures of liquidity, quoted and effective spreads. We find that the dynamics of the absolute futures/cash basis and liquidity are jointly determined.

The empirical findings include the following: (1) The speed of mean-reversion of the basis (to zero) over a futures contract's lifetime is positively related to liquidity; (2) There is two-way Granger causality between the three-month absolute basis and liquidity, as measured by effective spreads; (3) Impulse response functions reveal that shocks to the absolute bases are significantly informative in predicting future shifts in stock spreads; (4) Contemporaneous innovations to the absolute basis and spreads are positively correlated; (5) Shocks to spreads are more informative in forecasting shifts in the nine-month absolute basis than in the shorter-term bases, suggesting that liquidity affects arbitrageurs more in the relatively less actively traded longer-term futures contracts.

All these results are preserved after controlling for the Federal Funds rate, which bears a modestly significant and positive relationship with the shorter-term bases, thus delineating the additional role of other market frictions such as financing costs in the efficiency of the stock-futures pricing system. Overall, the results are consistent with deviations from no-arbitrage forecasting and, in turn, being forecasted by liquidity. If the degree to which markets satisfy the law of one price is taken to be a measure of the market's efficiency, then the results suggest that liquidity plays a significant role in moving markets towards an efficient outcome. In the reverse direction, the results suggest that deviations from the law of one price on a given day can help in predicting market liquidity on subsequent days, which is consistent with arbitrageurs affecting liquidity by their trading. Hence, the absolute basis may provide information to both retail and institutional investors that can be valuable in the forecasting and control of trading costs.

An empirical extension of our analysis would be to consider whether the ability of futures/cash bases to forecast liquidity and vice versa holds in other markets such as fixed-income and foreign exchange. It would also seem desirable to consider the basis/liquidity relation prior to days of news announcements (when the market might be particularly illiquid because of asymmetric information stemming from news leakage). Finally, it may be worthwhile to consider technological innovations in financial markets (such as the development of electronic communication networks). How such developments affect the basis by way of the liquidity channel would appear to be an interesting area for further research.

References

Amihud, Y., 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31-56.

Amihud, Y., and C. Hurvich, 2004, Predictive regressions: A reduced-bias estimation method, *Journal of Financial and Quantitative Analysis* 39, 813-841.

Amihud, Y., and H. Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223-249.

Ball, Clifford A., and Tarun Chordia, 2001, True spreads and equilibrium prices, *Journal of Finance* 56, 1801-1835.

Bakshi, G., C. Cao, and Z. Chen, 2000, Do call prices and the underlying stock always move in the same direction, *Review of Financial Studies* 13, 549-584.

Barclay, M., T. Hendershott, and C. Jones, 2003, Which witches better? A cross-market comparison of extreme liquidity shocks, working paper, Columbia University.

Benston, G., and R. Hagerman, 1974, Determinants of Bid-Asked Spreads in the Over-the-Counter Market, *Journal of Financial Economics* 1, 353-364.

Bessembinder, H., and P. Seguin, 1992, Futures-trading activity and stock price volatility, *Journal of Finance* 47, 2015-2034.

Beveridge, S., and C. Nelson, 1981, A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle, *Journal of Monetary Economics* 7, 151-174.

Bodie, Z., A. Kane, and A. Marcus, 1993, *Investments* (2nd edition), Irwin Publishers, Boston, MA.

Brennan, M., and E. Schwartz, 1990, Arbitrage in stock index futures, *Journal of Business* 62, S7-S31.

Chan, K., 1992, A further analysis of the lead-lag relation between the cash market and stock index futures market, *Review of Financial Studies* 5, 123-152.

Chan, K., and W. Fong, 2000, Trade size, order imbalance, and the volatility-volume relation, *Journal of Financial Economics* 57, 247-273.

Chordia, T., R. Roll, and A. Subrahmanyam, 2001, Market liquidity and trading activity, *Journal of Finance* 56, 501-530.

Chordia, T., R. Roll, and A. Subrahmanyam,, 2002, Order imbalance, liquidity and market returns, *Journal of Financial Economics* 65, 111-130.

Chung, P., 1991, A transactions data test of stock index futures market efficiency and index arbitrage profitability, *Journal of Finance* 46, 1791-1809.

Demsetz, H., 1968, The cost of transacting, *Quarterly Journal of Economics* 82, 33-53.

Fremault, A., 1991, Stock index futures and index arbitrage in a rational expectations model, *Journal of Business* 64, 523-547.

Gallant, A., P. Rossi, and G. Tauchen, 1992, Stock prices and volume, *Review of Financial Studies* 5, 199-242.

Harford, J., and A. Kaul, 2005, Correlated order flow: Pervasiveness, sources, and pricing effects, *Journal of Financial and Quantitative Analysis* 40, 29-55.

Hasbrouck, J., 1991, Measuring the information content of stock trades, *Journal of Finance* 56, 179-207.

Hasbrouck, J., and D. Seppi, 2001, Common factors in prices, order flows and liquidity, *Journal of Financial Economics* 59, 383-411.

Hou, K., and T. Moskowitz, 2004, Market frictions, price delay, and the cross-section of expected returns, forthcoming, *Review of Financial Studies*.

Jacoby, G., D. Fowler, and A. Gottesman, 2000, The capital asset pricing model and the liquidity effect: A theoretical approach, *Journal of Financial Markets* 3, 69-81.

Jokivuolle, E., 1995, Measuring true stock index value in the presence of infrequent trading, *Journal of Financial and Quantitative Analysis*, 30, 455-464.

Jones, C., 2001, A century of stock market liquidity and trading costs, working paper, Columbia University, New York, NY.

Jones, C., G. Kaul, and M. Lipson, 1994, Transactions, volume, and volatility, *Review of Financial Studies* 7, 631-651.

Kawaller, I., P. Koch, and T. Koch, 1987, The temporal price relationship between S&P500 futures and the S&P 500 Index, *Journal of Finance* 42, 1309-1329.

Kleidon, A., 1992, Arbitrage, nontrading, and stale prices, *Journal of Business* 65, 483-507.

Klemkosky, R., and J. Lee, 1991, The intraday ex post and ex ante profitability of index arbitrage, *Journal of Futures Markets* 11, 291-312.

Kumar, P., and D. Seppi, 1994, Information and index arbitrage, *Journal of Business* 67, 481-509.

Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Lamont, O., and R. Thaler, 2003, Can the market add and subtract? Mispricing in tech stock carve-outs, *Journal of Political Economy* 111, 227-268.

Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data. *Journal of Finance* 46, 733-747.

MacKinlay, C., and K. Ramaswamy, 1988, Index-futures arbitrage and the behavior of stock index futures prices, *Review of Financial Studies* 1, 137-158.

Miller, M., J. Muthuswamy, and R. Whaley, 1994, Mean reversion of Standard & Poor's 500 index basis changes: Arbitrage-induced or statistical illusion?, *Journal of Finance* 49, 479-513.

Mitchell, M., T. Pulvino, and E. Stafford, 2002, Limited arbitrage in equity markets, *Journal of Finance* 57, 551-584.

Modest, D., and M. Sundaresan, 1983, The relationship between spot and futures prices in stock index futures markets: Some preliminary evidence, *Journal of Futures Markets* 3, 15-41.

Neal, R., 1996, Direct tests of index arbitrage models, *Journal of Financial and Quantitative Analysis* 31, 541-562.

O'Hara, M., and G. Oldfield, 1986, The microeconomics of market making, *Journal of Financial and Quantitative Analysis* 21, 361-376.

Pastor, L., and R. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 113, 642-685.

Patell, J., and M. Wolfson, 1984, The intra-day speed of adjustment of stock prices to earnings and dividend announcements, *Journal of Financial Economics* 13, 223-252.

Pesaran, H., and Y. Shin, 1998, Generalised impulse response analysis in linear multivariate models, *Economics Letters* 58, 17-29.

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

Sadka, R., and A. Scherbina, 2004, Analyst disagreement, mispricing, and liquidity, working paper, Northwestern University.

Schwert, W., 1990, Stock market volatility and the crash of 87, *Review of Financial Studies* 3, 77-102.

Schwert, W, 2001, Anomalies and market efficiency, Chapter 17 in George Constantinides, Milton Harris, and René Stulz, eds., *Handbook of the Economics of Finance*, (North-Holland: Amsterdam).

Silber, W., 1981, The economic role of financial futures, in A. Peck (ed.) *Futures Markets: Their Economic Role*, pp. 83—13, Washington, D.C.: American Enterprise for Public Policy Research.

Sofianos, G., 1993, Index arbitrage profitability, *Journal of Derivatives* 1, 6-20.

Stambaugh, R., 1999, Predictive regressions, *Journal of Financial Economics* 54, 375-421.

Stein, J., 1987, Information externalities and welfare-reducing speculation, *Journal of Political Economy* 47, 2015-2034.

Stoll, H., 1978, The pricing of security dealer services: An empirical study of NASDAQ stocks, *Journal of Finance* 33, 1153-1172.

Stoll, H., 2000, Presidential address: Friction, *Journal of Finance* 55, 1479-1514.

Stoll, H., and R. Whaley, 1990, The dynamics of stock and stock index futures returns, *Journal of Financial and Quantitative Analysis* 25, 441-468.

Subrahmanyam, A., 1991, A theory of trading in stock index futures, *Review of Financial Studies* 4, 17-51.

Wang, Q., 2003, A model of stock index security trading: Information, volume and pricing, working paper, Ohio State University.

Yadav, P. and P. Pope, 1990, Stock index futures arbitrage: International evidence, *Journal of Futures Markets* 10, 573-603.

Appendix A

Estimation of the mean reversion parameter in Table 2

The basis is given by:

$$dB(t) = -\frac{\mu B}{\tau} dt + \gamma dz$$

Standard arguments imply that the value of *B(t)* at a point in time can be written as:

$$B(t) = \gamma (T-t)^{\mu} B\left[\frac{(T-t)^{1-2\mu}}{2\mu - 1} \right],$$

where dz is standard Brownian motion.

The log likelihood function of the parameters μ and γ can then be written as

$$\ln L(\mu, \gamma) = -\frac{T}{2} \ln 2\pi - \sum_{t=1}^{T} \ln \sigma_t - \frac{1}{2} \sum_{t=1}^{T} \frac{u_t^2}{\sigma_t^2},$$

where

$$u_t \equiv B(t+1) - \left(1 - \frac{1}{T-1}\right)^{\mu} B(t),$$

and

$$\sigma_t^2 \equiv \frac{\gamma^2 (T-t-1)}{2\mu - 1} \left[1 - \left(1 - \frac{1}{T-t}\right)^{2\mu - 1} \right].$$

The parameters can be estimated by maximizing the above log-likelihood function. The estimate of the mean-reversion parameter μ for each contract can be related to the average level of the liquidity during the contract's lifetime.

Appendix B

Adjustment of the stock index to account for non-synchronous trading

To investigate a possible influence of non-synchronous trading on our results, we adopt a procedure devised by Jokivuolle (1995), who shows that a log index corrected for non-synchronous trading is equivalent to the Beveridge and Nelson (1981) permanent component of a fitted ARIMA process. As an empirical illustration, Jokivuolle uses the Russell 2000 index, which includes many infrequently traded small firms. He notes that autocorrelations in this index' returns (the first differences of the log index levels), are significant only at the first daily lag, which implies that a first-order moving average process is adequate to describe the index returns. Our index, the NYSE composite, also has significant autocorrelation only at the first daily lag, so we adopt his illustrative method.

Specifically, let I_t be the NYSE composite index on day t and define $R_t = \ln(I_t/I_{t-1})$ as the log return. An MA(1) process for R_t implies

$$R_t = \overline{R} + \theta\varepsilon_{t-1} + \varepsilon_t,$$

where $\overline{R}$ is the expected return, θ is a moving average parameter, and the ε's are iid disturbances. Jokivuolle shows that the Beveridge and Nelson corrected log index for day t is given by

$$\ln(I_t) + \theta\varepsilon_t.$$

The parameter θ can be estimated from the first-order autocorrelation coefficient of returns, $\rho = \mathrm{Cov}(R_t,R_{t-1})/\mathrm{Var}(R) = \theta/(\theta^2+1)$. When θ is small, its value is close to the autocorrelation coefficient itself. Nonetheless, we computed an exact estimate by taking the negative root of the equation $\rho(\theta^2+1)+\theta = 0$. Using all daily observations in the sample, we estimate θ to be 0.04531.

To obtain a time series for ε_t, the moving average process must be inverted. Recursive substitution of lagged values; i.e., $\varepsilon_t = R_t - \overline{R} - \theta[R_{t-1} - \overline{R} - \theta(R_{t-2} - \overline{R} - \theta\varepsilon_{t-3})]$ eventually gives the correction term

$$\theta\varepsilon_t = \theta(R_t - \overline{R}) - \theta^2(R_{t-1} - \overline{R}) + \theta^3(R_{t-2} - \overline{R}) - \theta^4(R_{t-3} - \overline{R}) + \ldots$$

$$= \theta R_t - \theta^2 R_{t-1} + \theta^3 R_{t-2} - \theta^4 R_{t-3} + \ldots - \overline{R}\,\theta/\theta+1)$$

Since θ is small, the powered terms decline rapidly and are essentially immaterial after a few lags. We stopped adding terms after four lags as additional terms would have mattered only in the 11th or 12th decimal place. For the final term above, the sample mean return was substituted for $\overline{R}$.

The resulting corrected NYSE index returns have a first-order autocorrelation of 0.00597, with an insignificant t-statistic of 0.366. This contrasts with the observed index, whose returns have a first-order autocorrelation of 0.0452 and a t-statistic of 2.78. The Jokivuolle (1995) procedure does indeed appear to work well in correcting the index for non-synchronous trading. As mentioned in Section V, however, our results are materially unaltered when the corrected index is used in computation of the basis.

Table 1 – Summary Statistics for Deviations from Arbitrage and Liquidity Measures

This table presents summary statistics associated with the estimated absolute basis (in percent relative to the cash index value) for the three-month, six month, and nine-month NYSE Composite index futures contracts, and (in dollars) the NYSE value-weighted quoted and effective spreads; daily data for 1988-2002 inclusive.

	Mean	Median	Standard Deviation
3-month absolute basis	0.186%	0.142%	0.173%
6-month absolute basis	0.298%	0.259%	0.224%
9-month absolute basis	0.425%	0.359%	0.307%
Quoted spread	$0.164	$0.175	$0.053
Effective Spread	$0.110	$0.121	$0.036

Table 2 - Preliminary evidence on the basis-illiquidity relation

Panel A of the table presents correlations between the average absolute NYSE Composite stock-futures basis over a contract's lifetime with average values of the quoted and effective spreads over that time period. Panel B presents the correlation between quoted and effective spreads and each contract's estimated mean reversion parameter of a Brownian Bridge process for the basis. Daily data over the period 1988-2002 are used for the estimation.

Panel A: Correlations between average absolute bases and average liquidity measures across contracts (p-values in parentheses)

	Quoted spread	Effective Spread
3-month absolute basis	0.253 (0.049)	0.293 (0.022)
6-month absolute basis	0.260 (0.043)	0.290 (0.024)
9-month absolute basis	0.201 (0.120)	0.211 (0.102)

Panel B: Correlations between the estimates of the mean reversion parameter and liquidity measures (p-values in parenthesis)

	Quoted spread	Effective Spread
3-month absolute basis	-0.157 (0.229)	-0.142 (0.280)
6-month absolute basis	-0.356 (0.005)	-0.354 (0.006)
9-month absolute basis	-0.283 (0.028)	-0.280 (-0.030)

Table 3 –Absolute Basis Adjustment

In these OLS regressions, the dependent variable is the absolute basis for a NYSE Composite index futures contract. The time-period is 1988-2002 inclusive and the observations are daily. Dummy variables are included for Friday and for months of the year. RemTrm is the number of days until contract expiration, with 1 representing the last trading day. Holiday-1 is a dummy for the trading day prior to Thanksgiving, Christmas, or July 4. Time and Time**2 are orthogonalized linear and quadratic time trends. Rf-Dyld is the difference between the estimated risk-free rate and the dividend yield. All coefficients are multiplied by 100.

	Maximum Maturity of Futures Contract (Months)					
	Three		Six		Nine	
Variable	Coefficient	T	Coefficient	T	Coefficient	T
Friday	0.007	1.02	-0.008	-0.94	-0.009	-0.77
January	-0.039	-2.88	-0.132	-7.95	-0.227	-10.35
February	-0.064	-4.65	-0.155	-9.18	-0.244	-10.92
March	-0.048	-3.63	-0.145	-8.86	-0.247	-11.42
April	-0.017	-1.28	-0.088	-5.29	-0.169	-7.71
May	-0.061	-4.56	-0.129	-7.83	-0.204	-9.37
June	-0.021	-1.57	-0.074	-4.51	-0.120	-5.50
July	-0.029	-2.17	-0.045	-2.74	-0.049	-2.26
August	-0.067	-5.03	-0.094	-5.79	-0.137	-6.34
September	-0.014	-1.08	-0.027	-1.60	-0.034	-1.52
October	0.019	1.45	0.034	2.08	0.051	2.35
November	-0.053	-3.96	-0.047	-2.81	-0.022	-1.02
RemTrm=1	-0.018	-1.17				
RemTrm=2	-0.031	-6.62				
RemTrm=3	0.051	8.51				
RemTrm=4	1.946	10.05				
Holiday-1	-0.098	-4.44	-0.008	-0.40	-0.014	-0.54
Time	-0.065	-2.97	-0.056	-9.60	-0.055	-7.10
Time**2	-0.076	-3.46	0.026	3.47	-0.022	-2.26
Rf-DYld	-0.059	-2.66	4.239	17.59	7.276	22.76
Intercept	0.182	17.27	0.281	21.89	0.376	21.98
Adjusted R^2	0.081		0.147		0.206	

Table 4 – Spread Adjustment

In these OLS regressions, the dependent variables are the daily NYSE value-weighted quoted and effective spreads from 1988-2002, daily observations. Dummy variables are included for days of the week and months of the year. Holiday-1 denotes the trading day prior to Thanksgiving, Christmas, or July 4. Post16 and postdeci denote periods following a reduction in the minimum tick size to 1/16 (on June 24, 1997) and a further reduction to $.01 (on January 29, 2001). Time and time**2 are orthogonalized linear and quadratic time trends. All coefficients are multiplied by 100.

	Quoted spread		Effective spread	
Variable	Coefficient	T	Coefficient	T
Monday	-0.176	-3.16	-0.064	-1.76
Tuesday	-0.186	-3.41	-0.097	-2.70
Wednesday	-0.116	-2.13	-0.080	-2.24
Thursday	-0.072	-1.32	-0.071	-1.99
February	-0.031	-0.36	0.046	0.82
March	-0.223	-2.67	-0.119	-2.17
April	-0.168	-1.98	-0.007	-0.12
May	-0.678	-8.05	-0.294	-5.30
June	-0.906	-10.78	-0.407	-7.37
July	-0.840	-9.94	-0.422	-7.59
August	-0.898	-10.78	-0.423	-7.72
September	-1.086	-12.70	-0.578	-10.27
October	-0.583	-6.98	-0.225	-4.11
November	-0.931	-10.94	-0.535	-9.57
December	-0.850	-10.03	-0.464	-8.33
Holiday-1	0.055	0.56	0.026	0.41
post16th	-2.574	-29.32	-2.181	-37.77
postdeci	-6.459	-64.80	-4.441	-67.73
time	-4.427	-50.32	-2.835	-48.98
(time**2)	0.172	2.43	0.633	13.56
Intercept	18.85	239.2	12.68	244.6
Adjusted R^2	0.960		0.963	

Table 5 – Correlation Matrix for Adjusted Absolute Bases and Liquidity Measures

Correlations are presented between each measure of the adjusted absolute basis (ABASx, where x represents the basis horizon in months) and adjusted quoted and effective spreads (QSPR and ESPR respectively). The adjusted series are residuals from the regressions in Tables 2 and 3. All coefficients are significant at the 5% level except that between ABAS9 and QSPR, which is insignificant.

	ABAS3			
ABAS6	0.725	**ABAS6**		
ABAS9	0.491	0.878	**ABAS9**	
QSPR	0.113	0.046	0.022	**QSPR**
ESPR	0.193	0.173	0.119	**0.621**

Table 6 – OLS Regression coefficients

Regression results are presented for each measure of the adjusted absolute basis (ABASx, where x represents the basis horizon in months) on one lag of adjusted quoted and effective spreads (QSPR and ESPR respectively), and vice versa. The adjusted series are residuals from the regressions in Tables 2 and 3. In Panel A, the coefficients are multiplied by 100.

Panel A: Absolute bases as dependent variables

Independent variable→	Lag(QSPR)		Lag(ESPR)	
Dependent variable ↓	Coefficient	T	Coefficient	T
ABAS3	0.764	3.01	2.379	6.18
ABAS6	0.263	0.83	3.770	7.85
ABAS9	0.014	0.03	3.342	5.24

Panel B: Liquidity measures as dependent variables

Independent variable→	Lag(ABAS3)		Lag(ABAS6)		Lag(ABAS9)	
Dependent variable ↓	Coefficient	T	Coefficient	T	Coefficient	T
QSPR	0.752	7.29	0.232	2.80	0.033	0.52
ESPR	0.843	7.60	0.587	10.93	0.237	5.78

Table 7 - Vector Autoregressions

Vector autoregressions pair each measure of the adjusted absolute basis (ABASx, where x represents the basis horizon in months) with adjusted quoted and effective spreads (QSPR and ESPR respectively). The adjusted series are residuals from the regressions in Tables 2 and 3. The table presents correlations in VAR innovations and chi-square statistics and p-values (in parentheses) of pairwise Granger Causality tests between the endogenous variables.

Panel A: Correlations between VAR innovations

	ABAS3	ABAS6	ABAS9
QSPR	0.111	0.068	0.075
ESPR	0.138	0.080	0.106

Panel B: Granger Causality Tests
Null hypothesis: Row variable does not Granger-cause column variable

	ABAS3	ABAS6	ABAS9	QSPR	ESPR
ABAS3	-	-	-	32.94 (0.00)	45.55 (0.00)
ABAS6	-	-	-	9.892 (0.13)	17.62 (0.01)
ABAS9	-	-	-	8.68 (0.19)	9.10 (0.17)
QSPR	2.944 (0.82)	3.378 (0.76)	17.32 (0.01)	-	-
ESPR	15.74 (0.02)	5.292 (0.51)	28.05 (0.00)	-	-

Table 8– Robustness checks with market depth

Panels A and B present regression results for each measure of the adjusted absolute basis (ABASx, where x represents the basis horizon in months) on one lag of adjusted depth, and vice versa. The adjusted series are residuals from the regressions involving the independent variables in Tables 2 and 3. Panel C presents pairwise Granger Causality tests between the endogenous variables. In Panel A, the coefficients are multiplied by 1000.

Panel A: Absolute bases as dependent variables and lagged depth as independent variable

Independent variable→	Lag(Depth)	
Dependent variable ↓	Coefficient	T
ABAS3	-0.092	-4.96
ABAS6	-0.154	-6.64
ABAS9	-0.190	-6.17

Panel B: Depth as dependent variable and absolute bases as independent variables

Independent variable→	Lag(ABAS3)		Lag(ABAS6)		Lag(ABAS9)	
Dependent variable ↓	Coefficient	T	Coefficient	T	Coefficient	T
DEPTH	-70.05	-4.96	-74.73	-6.64	-52.57	-6.17

Panel C: Granger Causality Tests
Null hypothesis: Row variable does not Granger-cause column variable

	ABAS3	ABAS6	ABAS9	DEPTH
ABAS3	-	-	-	27.98 (0.00)
ABAS6	-	-	-	21.78 (0.00)
ABAS9	-	-	-	8.63 (0.20)
DEPTH	13.60 (0.03)	7.63 (0.27)	20.29 (0.00)	-

Table 9 – Robustness checks with short-term interest rates as an exogenous variable

This table presents results from adding the Federal Funds rate (FFRATE) and the constant maturity Treasury Bill rate (TBILL) as an exogenous variable to the VAR involving absolute bases and liquidity measures. Panel A presents pairwise Granger Causality tests between the endogenous variables. Panel B presents the regression coefficients associated with the two interest rates (the coefficients are multiplied by 1000).

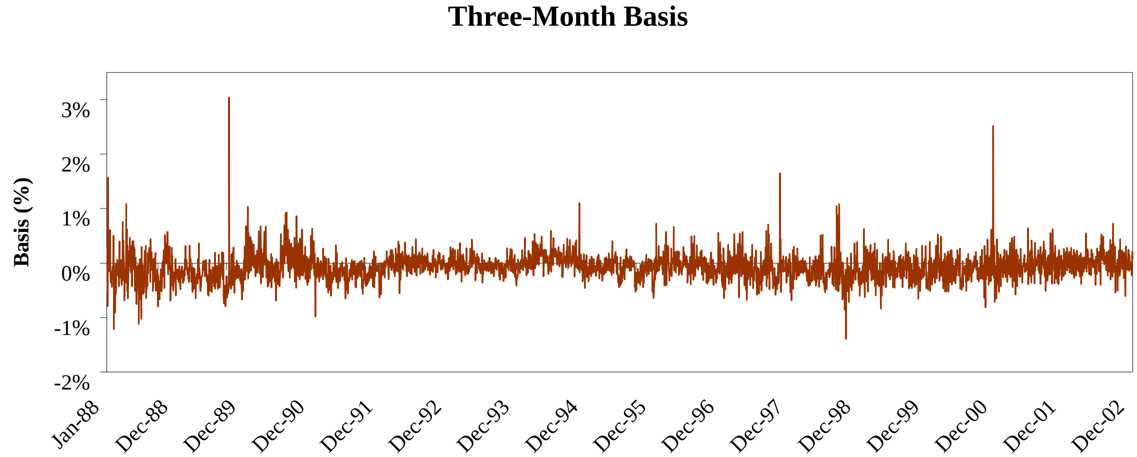
Panel A: Granger causality results
Null hypothesis: Row variable does not Granger-cause column variable

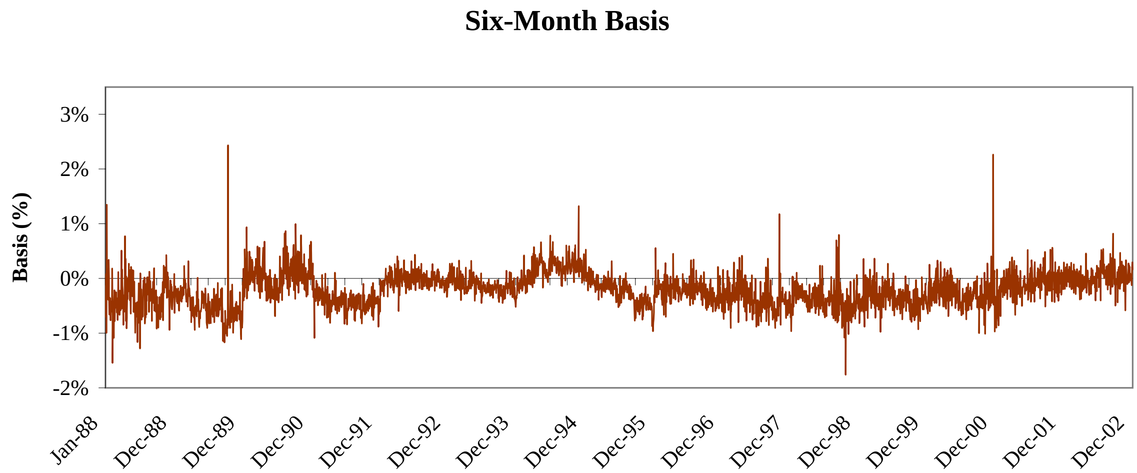
	ABAS3	ABAS6	ABAS9	QSPR	ESPR
ABAS3	-	-	-	32.88 (0.00)	46.12 (0.00)
ABAS6	-	-	-	9.892 (0.13)	18.68 (0.01)
ABAS9	-	-	-	8.77 (0.19)	9.49 (0.15)
QSPR	3.50 (0.74)	3.14 (0.79)	16.78 (0.01)	-	-
ESPR	15.65 (0.02)	5.53 (0.48)	28.07 (0.00)	-	-

Panel B: Regression coefficients for the short-term interest rates

Absolute basis/Quoted spread VARs						
	ABAS3	QSPR	ABAS6	QSPR	ABAS9	QSPR
FFRATE	0.096 (2.36)	-0.165 (-1.20)	0.104 (2.30)	-0.167 (-1.20)	0.053 (1.04)	-0.124 (-1.09)
TBILL	-0.074 (-1.75)	0.164 (1.15)	-0.084 (-1.79)	0.164 (1.13)	-0.035 (-0.67)	0.128 (1.07)
Absolute basis/Effective spread VARs						
	ABAS3	ESPR	ABAS6	ESPR	ABAS9	ESPR
FFRATE	0.082 (2.18)	-0.171 (-1.56)	0.104 (2.30)	-0.194 (-1.75)	0.059 (1.17)	-0.166 (-1.49)
TBILL	-0.061 (-1.57)	0.193 (1.68)	-0.080 (-1.69)	0.209 (1.80)	-0.038 (-0.73)	0.188 (1.62)

Figure 1. NYSE Composite Index Futures Bases Over Time





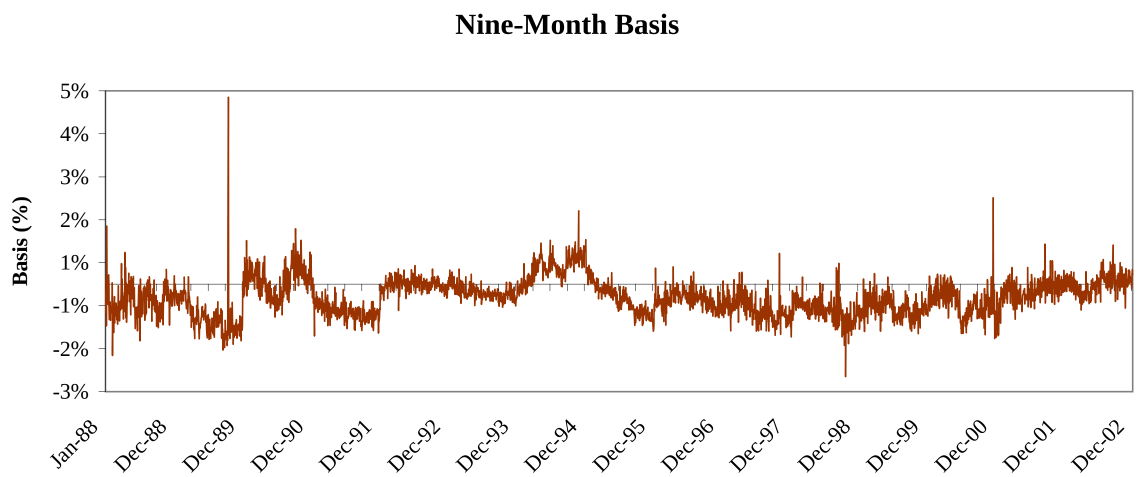


Figure 2. Liquidity Measures Over Time

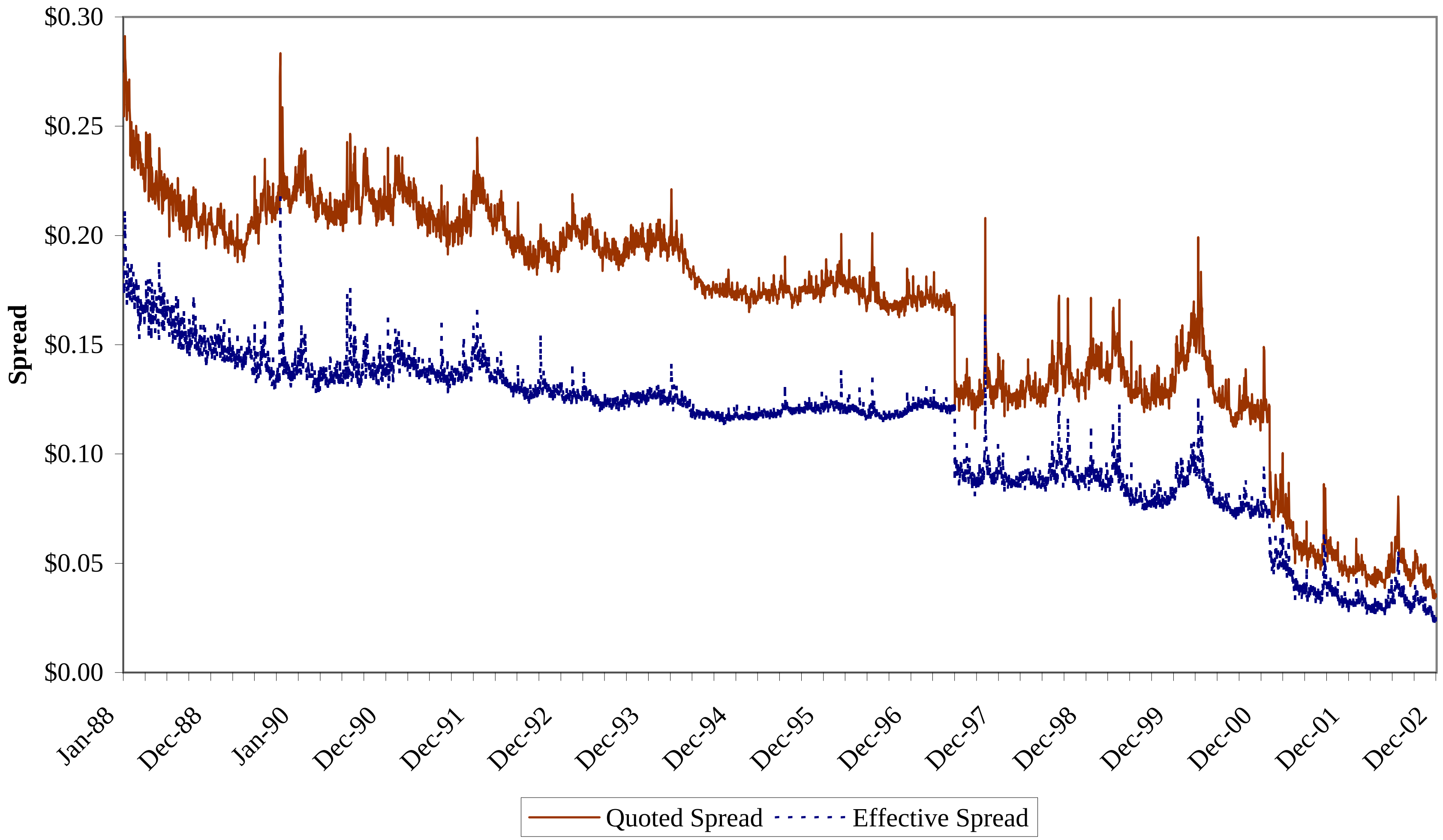


Figure 3 - Impulse response functions for the bivariate VAR with the adjusted three-month basis (ABAS3) and the adjusted quoted spread (QSPR)

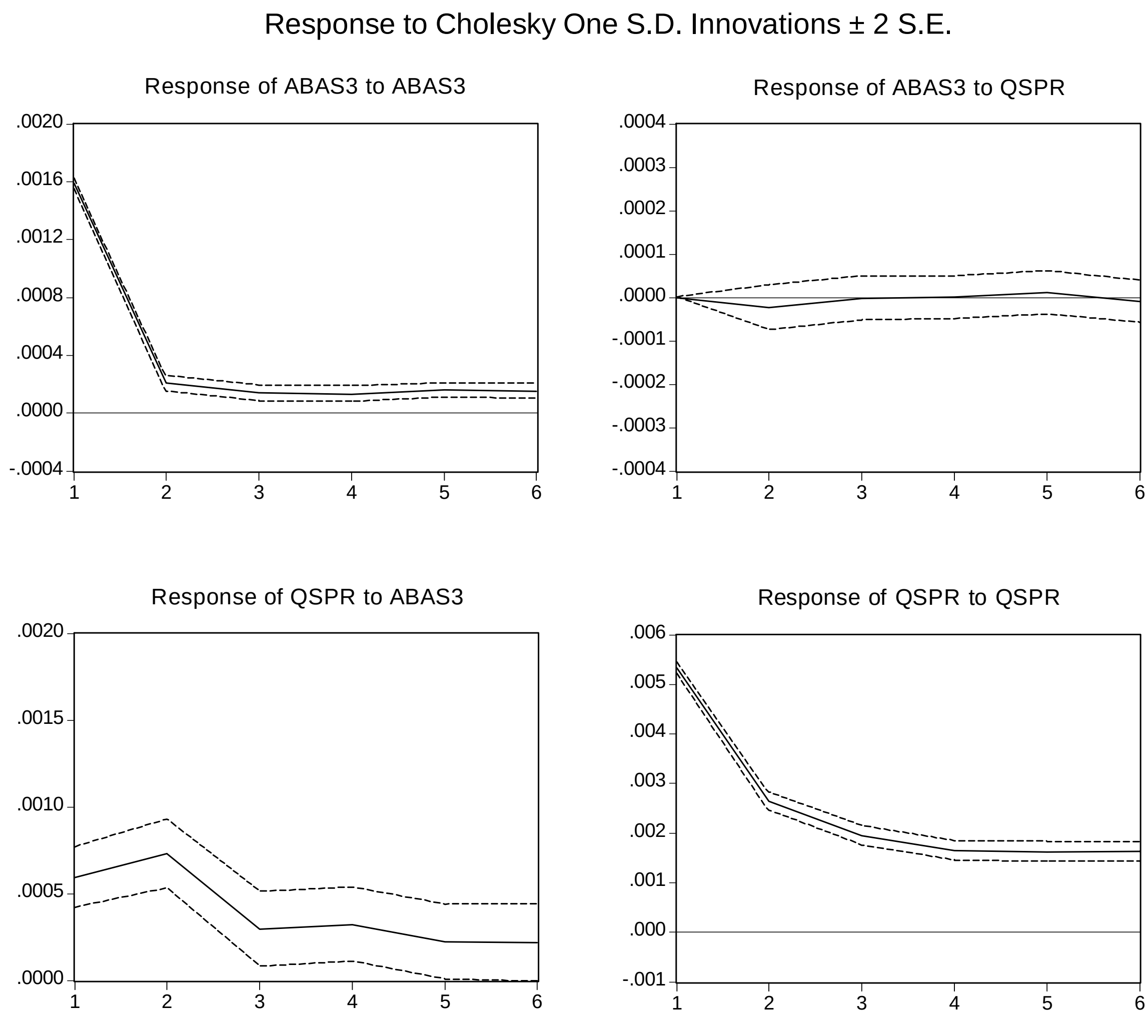


Figure 4 - Impulse response functions for the bivariate VAR with the adjusted three-month basis (ABAS3) and the adjusted effective spread (ESPR)

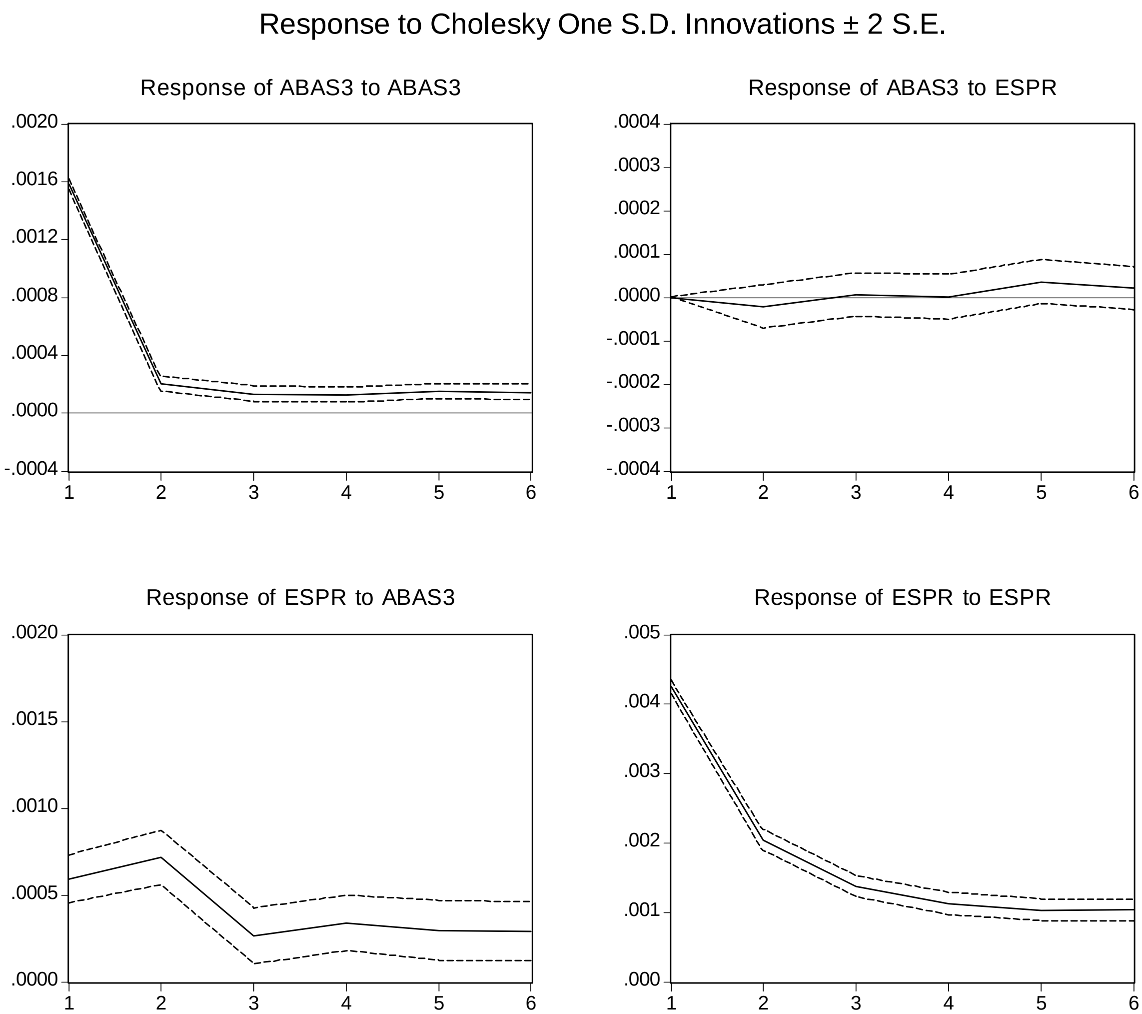


Figure 5 - Impulse response functions for the bivariate VAR with the six-month basis (ABAS6) and the quoted spread (QSPR)

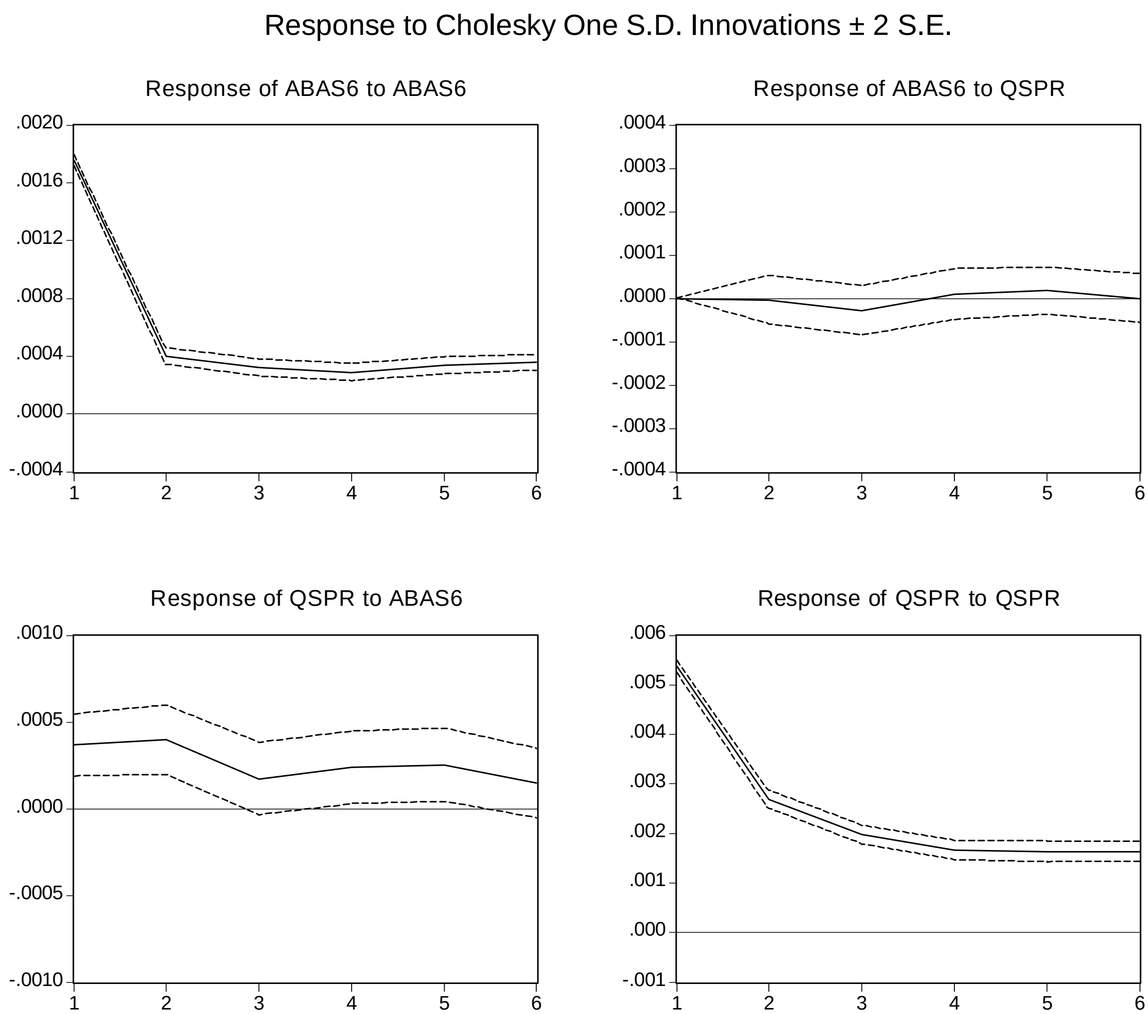


Figure 6 - Impulse response functions for the bivariate VAR with the adjusted six-month basis (ABAS6) and the adjusted effective spread (ESPR)

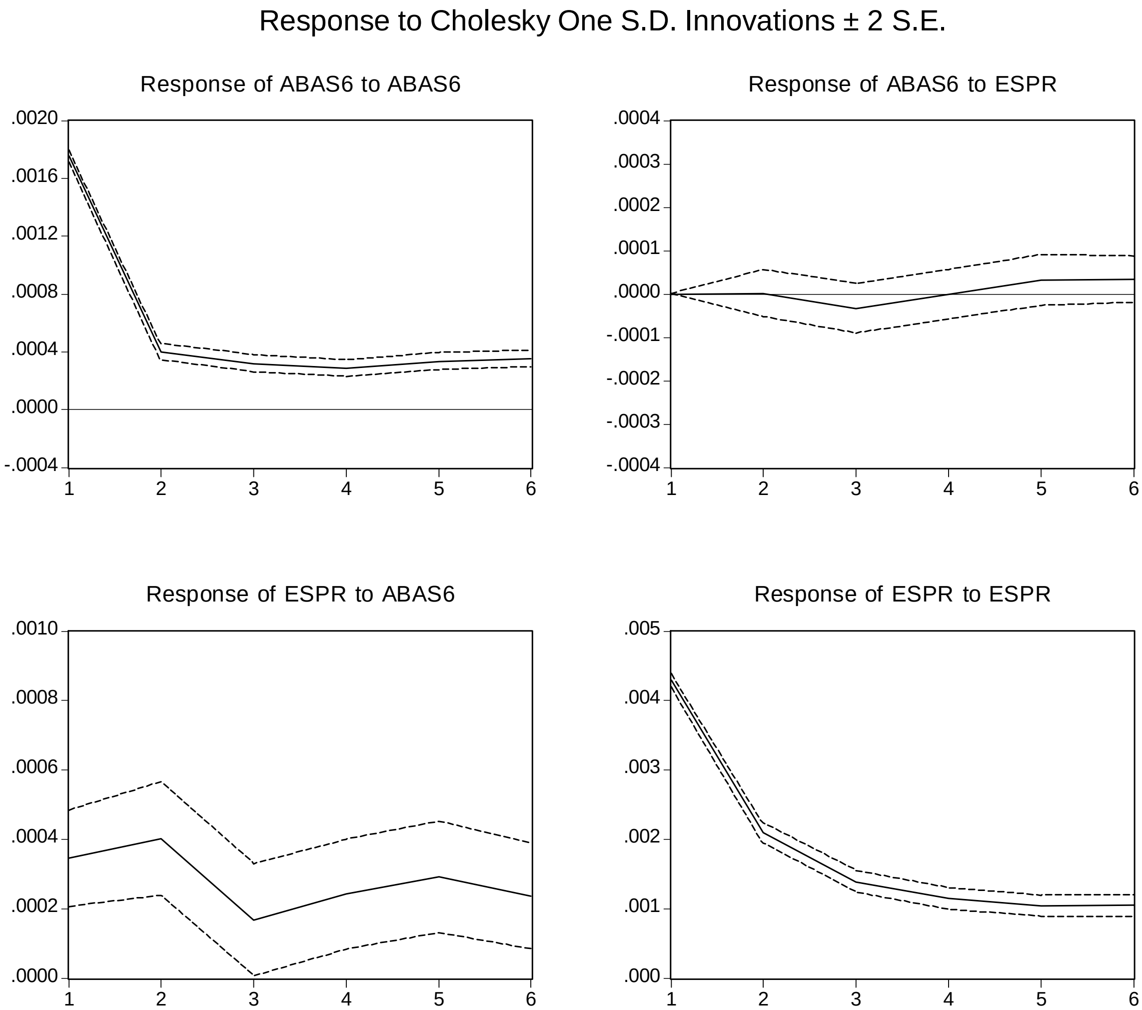


Figure 7 - Impulse response functions for the bivariate VAR with the adjusted nine-month basis (ABAS9) and the adjusted quoted spread (QSPR)

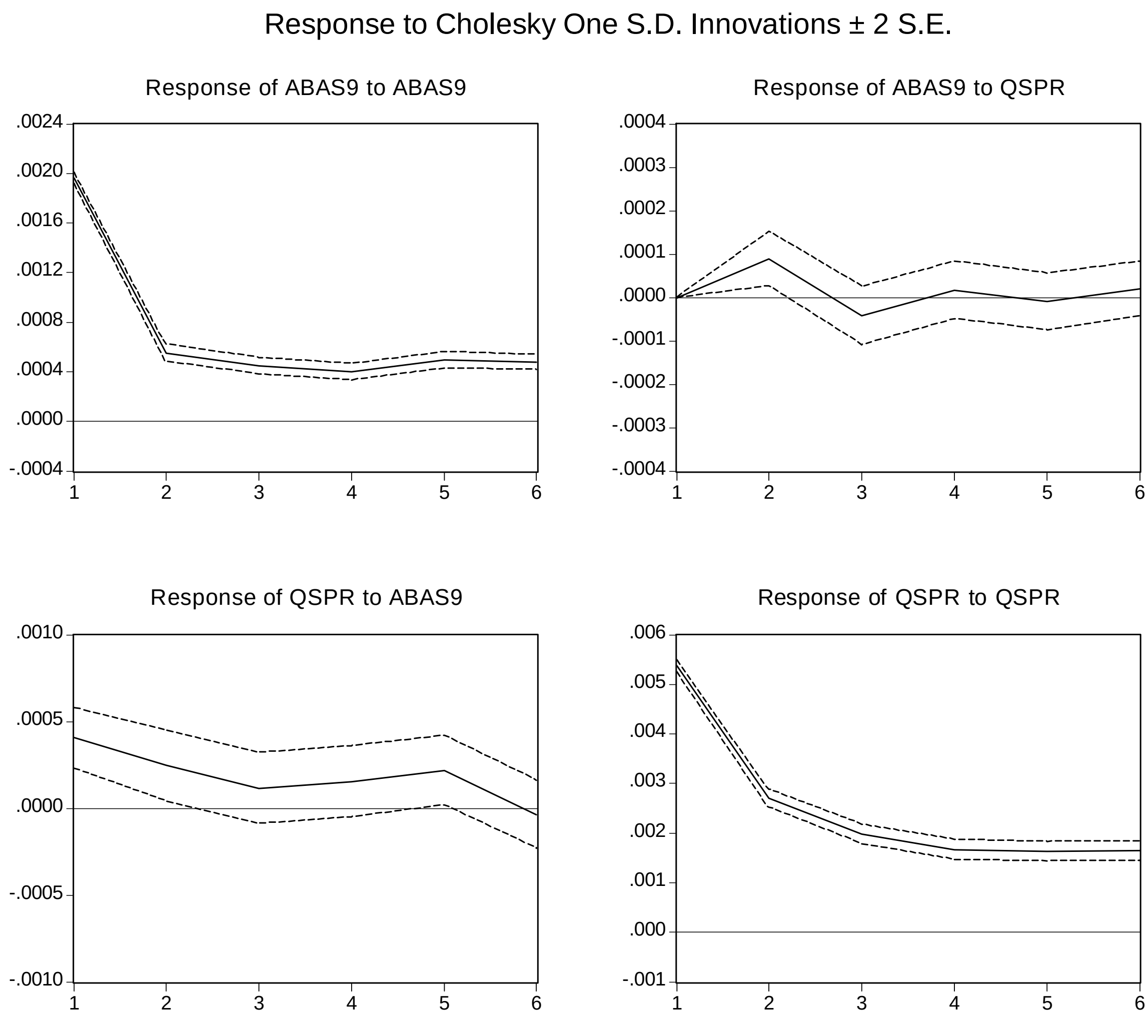


Figure 8 - Impulse response functions for the bivariate VAR with the adjusted nine-month basis (ABAS9) and the adjusted effective spread (ESPR)

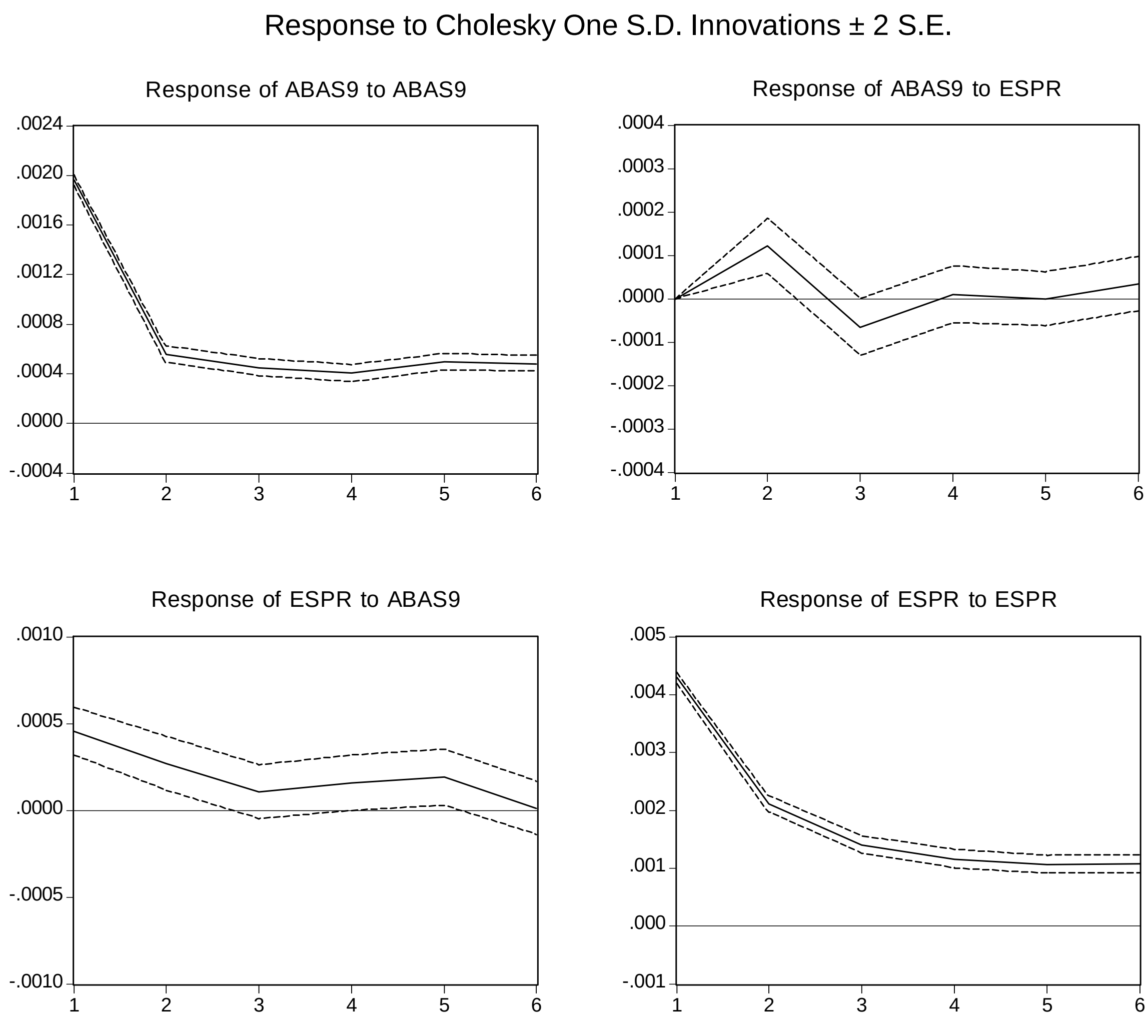


Figure 9 – Illustrative impulse response functions for the basis using the index adjusted for non-synchronous trading (ABAS3T-three month absolute basis, ABAS9T-nine month absolute basis, ESPR-effective spread)