The Short-term Dynamics of Information Risk

Thomas Henker*
Shah A. H. Shah-Idil
Jianxin Wang

COMMENTS ARE WELCOME
PLEASE DO NOT QUOTE WITHOUT PERMISSION

ABSTRACT
Volatility is widely regarded as the manifestation of uncertainty in security prices. The underlying causes of how exactly this uncertainty is incorporated in security prices through the trading process is less well understood. Our paper for the first time documents considerable time variation in adverse selection costs and a high correlation between volatility in security prices and adverse selection costs. As volatility adverse selection costs exhibit long memory properties evident in studies of asset return volatility. Market-wide volatility is more important in explaining the daily variations of a stock's information risk than firm-specific volatility. Information risk increases with the number of trades but decreases with trading volume.

* Thomas Henker, t.henker@unsw.edu.au, +61-2-9385-5854, and Jianxin Wang jx.wang@unsw.edu.au, +61-2-9385-5863, are from the Australian School of Business, Banking and Finance, University of New South Wales, UNSW-Sydney, NSW2052, Australia, fax. +61-2-9385-6347. Shah A.H. Shah-Idil is from the Reserve Bank of Australia, ShahIdilS@rba.gov.au. We thank conference participants at the MMUHFDIC 20808 conference in Perth and the FIRN research day in Brisbane for helpful comments. The usual disclaimer applies.

I. Introduction

Volatility in security prices is the observable evidence for market participants' uncertainty about the appropriateness of prices at which they trade securities throughout the trading day. Parallel to the extensive literature on security price volatility a large body of market microstructure literature exists that concerns itself with the question how uncertainty about informed counterparties faced by the market maker is incorporated through the mechanics of the trading process into security prices [Kyle 1985, Glosten and Milgrom 1985]. A large number of spread decomposition models [for example, Madhavan, Richardson and Roomans 1997, Huang and Stoll 1997, Henker and Wang 2006, Lin Sanger and Booth 1995] identified this informational asymmetry as a cost component that market makers consider when setting bid and ask quotes at which they are willing to trade securities. More recently Easley and O"Hara (2004) show that information asymmetry is a priced and undiversifiable component of security returns. While the dynamic nature of volatility is unquestioned empirical applications of spread decomposition models have thus far been static in nature. Empirical research was limited to cross-sectional and time-of-the day analyses.

In this paper we document that the adverse selection cost component of different securities not only varies over time but also has a high time-series correlation with observed volatility. As uncertainty in financial markets increases so does the spread cost component market makers charge buyers and sellers to compensate for trading against potentially informed counterparties. Since adverse selection costs can be estimated with the help of spread cost

decomposition models and spreads are readily observable, the dynamic nature of adverse selection costs we find provides an estimate of how much informational asymmetry the market maker perceives to exist in the market. We find informational asymmetry estimates to be correlated with realized market and security volatility and to exhibit long memory properties similar to realized volatility.

To accomplish these tasks we first estimate adverse selection costs at a daily interval for three different spread decomposition models and find that daily adverse selection cost estimates indeed vary widely over time. The dynamic nature of the adverse selection cost component is evident for three different spread decomposition models. We then use a modified version of the heterogenous autoregressive realized volatility model (HAR-RV) by Corsi (2004) to model the time-series evolution of adverse selection costs (HAR-α). In a further extension of the modified model we show that various volatility proxy variables as well as realized volatility can be incorporated in the basic HAR-α model to improve forecasts of the time series evolution of adverse selection costs. This model extension forms an empirical link between the adverse selection costs in market microstructure literature and the volatility modeling literature.

According to the theories of the bid-ask spread, a portion of the spread is priced in proportion to the adverse selection risk faced by market-makers. Using intra-day trade and quote data, statistical models have been developed to estimate this component of the spread and thus provide a measure for adverse selection risk. Whilst this approach has yielded promising results in terms of

identifying a clear relationship between adverse selection risk and the price formation process from a *time-of-day*[1] and *cross-sectional* perspective, research to date has set aside the examination related to the *evolution* of adverse selection risk over time. This omission has left many pertinent questions still unanswered. What characterises the behaviour of adverse selection risk at the daily, weekly, monthly or quarterly frequency? How do spread decomposition models compare as measures of information asymmetry beyond the *time-of-day* and *cross-sectional* paradigms? Volatility and volume have been extensively utilised as proxies of informed trading prior to the development of spread decomposition models. How does the temporal evolution in estimated adverse selection risk compare with these measures of trading activity? As is apparent, despite extensive theoretical and empirical attention, the characteristics and determinants of information asymmetry still remain largely unclear.

The importance of understanding the temporal evolution in adverse selection risk is emphasised when we consider empirical studies which have explored the time-of-day and cross-sectional variation in adverse selection costs. For example, the analysis of the behaviour of estimated adverse selection costs from a time-of-day perspective has been able to explain and reconcile various contradictions between observed and predicted intra-day price-movements. Madhavan, Richardson and Roomans (1997) reconcile the theoretically predicted intra-day decline in adverse selection and the contradictory observed U-shape pattern in the bid-ask spread and volatility. They estimate the adverse selection

[1] By *time-of-day* here means studies that document consistent patterns that occur during the period of a trading day and not time-series analysis.

component of the spread, confirming its intra-day decline and attribute the rise in spreads preceding the close to an increase in dealer costs possibly reflecting the risk of holding inventory overnight. From a cross-sectional perspective, Easley, Hvidkjaer and O'Hara (2002) estimate a measure of adverse selection risk for a large sample of NYSE stocks. Using a Fama-French (1992) asset-pricing framework they conclude that stocks faced with a higher likelihood of informed trading have higher rates of return. The magnitude and statistical significance of their results provides strong support for the hypothesis that adverse selection risk affects the determination of prices.

By focusing on the temporal evolution of adverse selection costs, this paper provides an entirely novel perspective from which to view the impact of information on the trade process. Although similar theory drives the time-series variation of adverse selection costs in a similar manner to time-of-day and cross-sectional variation, the nature of variation in these three forms differ. A cross-sectional analysis assumes the constancy of parameter estimates over the sample period, thus results purely reflect the effects of cross-sectional variation in adverse selection costs. Time-of-day analysis, on the other hand, concentrates on the examination of consistent patterns in the behaviour of adverse selection risk over the trading day. With time-series analysis, we are able focus on the observation and analysis of a continuous evolution of adverse selection costs over the entire sample period. To quantify, characterise and model the extent of this variation is the purpose of this paper.

Our objectives are twofold. First we apply three spread decomposition models within a dynamic setting. We choose to estimate adverse selection costs at a daily frequency and emphasise that this study is the first to estimate adverse selection costs within intervals narrow enough to permit time-series analysis. Secondly, we develop an econometric specification to explain the variation in adverse selection costs and relate it to stocks price volatility. This is performed with hypotheses which are formulated in relation to the widely used trade-based measures of volatility and volume. This forms the second major contribution of this paper where we find evidence for an empirical link between the adverse selection costs literature and the volatility-volume literature.

We uncover a number of stylised features of adverse selection costs. One of our most striking results is the finding of long memory in the time-series evolution of adverse selection costs. Not only do adverse selection costs exhibit significant autocorrelation, this autocorrelation is highly persistent and is significant up to the quarterly horizon. We model the long-term dependency of adverse selection costs within the framework of existing long memory volatility models and find results consistent with the long memory volatility literature. This result supports the premise that there is a common variable which produces the long memory property of both volatility and adverse selection costs. Consistent with our hypotheses, both market volatility and firm-specific volatility have positive and significant individual effects on adverse selection costs. It is likely that these separate and uncorrelated volatility components reflect different sources of adverse selection risk priced by the market maker.

Variations in trading volume reflect both a liquidity and information effect. Volume measured by the share volume displays a consistently negative relationship with adverse selection costs. After controlling for liquidity effects, the inclusion of the number of trades appears to proxy for the activity of informed traders with a positive and significant effect. This result is consistent with Jones, Kaul and Lipson (1994) who find that the number of trades explains virtually all of the volatility-volume relation with the size of trade playing a trivial role. Our findings extend this logic and suggest that the size of trades are primarily determined by large uninformed institutions and thus explain liquidity effects. The decomposition of the number of trades into expected and unexpected components displays results consistent with the price impacts of both noise- and information-based trading. The unexpected component of volume is able to significantly explain variations in adverse selection costs even after controlling for volatility. This is particularly notable given that volatility is usually assumed to be a superior proxy for the activitie of informed traders. This supports the notion that the number of trades may contain information regarding the activity of informed traders independent of the variation in volatility. Out of the four specifications of spread decomposition models we consider, the average constant spread version of the Henker and Wang (2006) adaptation of the Huang and Stoll (12997) model exhibits results most consistent with a reliable measure of adverse selection risk. The obtained results from this model are also robust to exogenously and endogenously determined sample structural breaks.

The paper is organised as follows. The next section describes our data and summary statistics. Section III compares the adverse selection costs estimated

from several popular models. Section IV estimates the dynamic evolution of the adverse selection costs and the impact of other information variables. Various robustness tests are performed to ensure the validity of the results. Section V concludes.

II. Data and Summary Statistics

In order to estimate daily asymmetric information measures, we use stocks from the Major Market Index (MMI), which are highly liquid and actively traded stocks. This sample selection is not unlike Huang and Stoll (1994, 1997). We restrict our analysis to stocks with a NYSE listing in order to remove any impact from different trading mechanisms (Affleck-Graves, Hegde and Miller (1994)). This removes Microsoft Corporation (MSFT). We also drop Chevron Corporation (CVX) because of the mergers of the company with Texaco in 2001 and Unocal in 2005. We perform our analysis on the remaining 18 stocks.

The sample consists of intra-day data from each stock beginning the 2nd of July 2001 until the 31st of December 2004. Note that the entirety of our data is placed in the NYSE's post-decimalisation period. Transaction and quotation data for each security is obtained from the NYSE TAQ database. All quotes eligible for the Best-Bid-Or-Offer (BBO) calculation are included. The NYSE opens with a call auction and operates as a continuous auction for the remainder of the day. To avoid confounding the two trading mechanisms, we only consider trades from 9:45 am to 4:00 pm for every trading day.

II.1 Bunching of Intraday Trades

Trade indicator spread decomposition models are dependent on reliable estimation of the autocorrelation coefficient of the trade flow, $(1-2\pi)$, in providing estimates of the adverse selection component of the spread. As Huang and Stoll (1997) and others we bunch consecutive trades at identical prices and unaltered quotes to reduce the effects of order splitting on the spread cost component estimation. We observe that combining trades reduces the sample size by an average of 36% across the 18 stocks. There is no relationship between the frequency at which a stock is traded and the reduction in sample size after bunching.

II.2 Summary Statistics

Table 1 presents the summary statistics for each of the stocks within our sample. The MMI stocks are all large and actively traded. The minimum number of intra-day observations for all stocks are above 250 observations, which is the minimum sample size used by Madhavan *et al* (1997). The dollar spread is positively correlated with price level, with a correlation of 0.953. This highlights the need to control for the effects of share price if we seek to endeavour to make cross-sectional inferences.

II.3 Signing Trades

Two of the adverse selection cost estimators we consider require trade indicator variables as part of their specification. Since the parties originating the order flow are not directly observable, we infer the direction of trades by comparing transaction prices with the prevailing ask and bid quotes at the time of trade. We follow Henker and Wang (2006) and delay quotes for one second

before matching them with transactions. Following Lee and Ready (1991), trade is classified as buyer- (seller-) initiated if it occurs above (below) the quote midpoint. If a trade occurs outside of the spread or at the midpoint, we utilise the tick test by Lee and Ready (1991).

II.4 Daily Realized Volatility

For each stock, we construct daily realized volatility using 5-minute and 30-minute sampling. We also obtain the Chicago Board Options Exchange's Volatility Index (VIX) as a measure for daily market volatility. Table 2 presents the summary statistics for the daily volatility measures. Realized volatility shows positive skewness consistent with the presence of jumps. The summary statistics are comparable to those of earlier studies (Andersen *et al*, 2001).

III. Spread Component Models and Estimates

This section presents three popular models of spread cost components and their estimates. We also draw some empirical comparisons across models.

III.1 Henker and Wang (2006)

The Henker and Wang model (2006) (HW) is a modification of the Huang and Stoll (1997) model which resulted from the discovery of inconsistencies in the original empirical application of the model. The posted spread, S, is decomposed into three components: the adverse selection cost, α, the inventory cost, β, and the order-process cost $(1-\alpha-\beta)$. The estimation is based on the

observed trade indicator, Q, and changes in the midpoint, M, of the posted quotes:

$$\Delta M_t = (\alpha + \beta)\frac{S}{2}Q_{t-1} - \alpha(1 - 2\pi)\frac{S}{2}Q_{t-2} + \varepsilon_t$$

HS and subsequent studies use time-varying spread. HW shows an error in the empirical adaptation and argues for the use of constant spread. We use the Generalized Method of Moments (GMM) procedure to estimate the model. The covariance matrix is estimated via Newey-West (1989), with the bandwidth selected by the automatic bandwidth estimator of Andrews (1991) using the Bartlett Kernel.

As discussed above, the Huang and Stoll model requires that the probability of reversal π>0.5. HS shows that when π<0.5, the estimates of α tend to be negative. Figure 1 draws the relationship between HW α estimates and the daily π. It is clear that as π approaches to 0.5, the estimated α becomes unstable. To determine a cut-off point for π, the data is split into 20 percentiles bounds ranked by π. The percentage of α estimates which fall outside the bounds of 0 and 1 are identified and presented in Figure 1 Panel B. Each column represents the proportion of observations within the percentile bound which occur outside the limits of zero and one. As can be observed, the majority of the negative α estimates occur in the 5th, 10th and 15th percentiles – 98.3% to be specific. Thus, we determine the cut-off point as the largest π estimate of the 15th percentile. This is at a value of 0.5144 and we discard all estimates which have a value of π equal to or less than it. Previous studies such as Clarke and Shastri (2000) and

Van Ness *et al* (2001) have similarly eliminated values of α which lie outside the bound of the spread.

The summary of α estimates from Henker and Wang (2006) is reported in Table 3. The mean and median αs for all stocks are positive. As reported in row 2 of Panel A, the majority of α estimates removed were from MMM and IBM with only 37 percent and 58 percent of the original sample remaining. As a result of this, we exclude these two stocks from further analysis. The average α is 3 bps and the median is 2 bps of the price. The negative αs are not statistically significant.

III.2 Madhavan, Richardson and Roomans (1997)

Madhavan *et al*. (1997) decomposes the spread into two components. The model individually estimates the adverse selection component but combines the separate costs pertaining to order processing and inventory risk into one. Furthermore, the estimates are in dollar terms as opposed to percentages:

$$\Delta P_t = (\phi + \theta) Q_t - (\phi + (1 - 2\pi)\theta) Q_{t-1} + \varepsilon_{t+1} + \xi_t + \xi_{t-1}$$

The estimated adverse selection cost is given by $\alpha = \Phi + \Theta$.

The results from the MRR model are presented in Table 4. The average dollar α stands at 0.278 cents. The percentage α values range from a low of 0.002 percent for General Electric, GE, to a high of 0.012 percent for Eastman Kodak, EK, with an average of 0.007 percent. The mean α estimates are also consistently higher than their corresponding medians implying a similar positive skew to the distribution of α estimates as observed in the HW model. Comparing

the dollar and percentage α estimates between the MRR and HW models, the MRR's average dollar α is less than half of that estimated by HW at 0.799 cents. MRR's average percentage α estimate is also lower compared to the HW value of 0.0206 percent. We perform a paired T-test across the two samples for percentage αs with the results indicating a statistically significant difference in sample means between the two. The null hypothese of equal means was strongly rejected at all conventional levels of significance. This result is consistent with Madhavan *et al* (1997) who argue that adverse selection costs calculated relative to the quoted spread will be relatively overpriced since trades occur both at and within the quoted bid and ask prices.

The recognised stability of the MRR model is very obvious in its capacity to estimate the entire sample of αs with a negligible number of negative values. Interestingly, the MRR model exhibits an elevated level of stability relative to previous studies with only 0.6 percent of estimates being negative. Van Ness *et al* (2001) note that 152 out of 856 estimates or 17.8 percent were extreme. Clarke and Shastri (2000) observe extreme estimates for 13.7 percent of their sample. Since we did not alter the specification of the MRR model for our study, we can only suspect this result to be a product of our different sample period and stock selection. The average standard deviations of the dollar and percentage α estimates are at 0.158 cents and 0.0044 percent, respectively, which are substantially smaller than the results obtained through the HW model. Also, although the results are not presented, estimation through the MRR model provides coefficient estimates with much higher t-statistics. As such, the issue of

decreased sample size is not anticipated to be a considerable problem with the MRR model.

It is important to note however that it is not apparent whether the MRR model's lack of variation can be viewed as beneficial or detrimental to the performance of the model. There is no theoretical guidance on the underlying process we are attempting to estimate. If we draw a comparison between the degree of asymmetric information and volatility, realized volatility, which is a superior volatility measure, shows much greater daily variation than GARCH-estimated volatility; see Wang and Yang (2006). If the true degree of information asymmetry varies significantly from day to day, then one would seek a model that captures such variations.

III.3 Lin, Sanger and Booth (1995)

Lin, Sanger and Booth (1995) decompose the spread into the adverse selection and order processing components as a percentage of the effective spread:

$$M_{t+1} - M_t = \alpha z_t + \varepsilon_{t+1}$$

where z_t is the difference between P_t and M_t with $z_t < 0$ for a sell order and $z_t > 0$ for a buy order. Since all quote revisions in response to trades reflect adverse selection risk, α represents Lin *et al*'s (1995) measure of the percentage adverse selection spread component.

Table 5 provides the LSB estimates. The results indicate a stronger similarity between the LSB and MRR models relative to the comparison between the MRR and HW models in the previous section. The LSB model's average

dollar α 0.41 cents compared to the MRR's 0.27 cents and HW's 0.80 cents. The average percentage α is 0.011 percent, relative to MRR value of 0.007 percent and HW value of 0.021 percent. The lower magnitude of αs estimated by the LSB model relative to HW is may be related to its use of the effective spread. Following the reasoning of MRR, this measure of the spread controls for transactions within the quoted spread and thus is less likely to over-estimate the adverse selection spread component. On the other hand, the LSB model's higher α estimate relative to the MRR model can perhaps be attributed to its assumption of zero inventory costs. It may be the case that a proportion of the LSB model's estimated α may be related to quote revisions due to inventory reasons. More in depth analysis will be required to confirm these statements.

III.4 Correlation of αs across Models

From a cross-model perspective, the three models exhibit significant covariation over the sample period. Table 5 represents a correlation matrix of α stimates from all three models. For this measure it is important that we calculate it separately for every stock to eliminate any effects of cross-sectional covariation. This allows us to observe the temporal covariation of the four estimation methods in isolation.

Referring to the percentage α results, the four models appear to correlate strongly through time with correlation coefficients ranging from a low of 0.434 between the TEC-HW and LSB models to a high of 0.804 between the MRR and LSB models. The superiority of the ACS specification to the TEC is quite apparent as it dominates the TEC in all correlation measures. Similar results occur for the dollar αs. Clarke and Shastri (2000) calculate the cross-sectional

correlation of spread-based measures of information asymmetry for the upper two firm size quartiles of their sample, likely to be representative of ours. Correlation appears to be weaker across the cross-sectional dimension with correlation values ranging from 0.27 to 0.72. This suggests that adverse selection cost estimators may be more responsive to temporal as opposed to cross-sectional variation implying a greater functionality in the former application.

Since the majority of spread decomposition applications involve estimating adverse selection costs over much longer sample periods, we compare α estimates over longer period with our daily estimated αs. We estimate adverse selection costs at half yearly intervals resulting in seven "static" long-run estimates across the sample period for 16 stocks, resulting in 112 (=7x16) estimates for each model. A simple t-statistic is used to test the null hypothesis that the average daily estimates, α-bar, equals the long-run estimates α-hat. For HW, only 4 out of the 112 stock periods (3.57%) reject the null of equality. The numbers for MRR and LSB are 1 and 6 out of 112 respectively. Overall the results indicate that there is not significant difference between the average daily α and the long-run α.

Next we test whether the static and dynamic αs maintain a similar measure of ordinal information asymmetry using Spearman's rank correlation. Table 6 reports Spearman's rank correlation calculated between the α-bar and α-hat across the 16 stocks over seven half-yearly intervals. Most correlation coefficients are above 90% and significantly different from zero (the standard error of the estimated correlation coefficient is given by $\sqrt{(1-\hat{\rho}^2)/(n-2)}$). All MRR

correlations are not statistically different from one. The results show that ranks based on dynamic and static αs are similar but not identical.

IV. Dynamics of Asymmetric Information

IV.1 Temporal Dynamics

Figure 2 presents the time-series plot of the median daily percentage □s from the three models and the VIX as well as the estimated □s against the VIX. Panel A shows that all three information risk measures show similar patterns over time as the market volatility index. The HW measure varies more with the market volatility than MRR and LSB. The higher estimated HW □s occurred in the first half of the sample period when the VIX was high. HW □ converges to MRR and LSB in the second half of the sample when the volatility was much lower. The positive relationship between volatility and information risk is further demonstrated in Panel B where we observe a strong positive relation between the VIX and HW α.

Like volatility, the estimated information risk measures exhibit long-term persistence. Figure 3 shows the autocorrelation function up to 100 lags for the three measures plus the realized volatility. It is constructed by averaging the autocorrelation at each lag across the 16 stocks. The 95% confidence interval is [-0.07, 0.07] for 872 observations. The similarity of the autocorrelation functions across the three measures and volatility is striking. The information risk measures show significant autocorrelation beyond 100 days. The MRR percentage □s display the highest level of persistence. The autocorrelations of

the HW and LSB measures are almost identical as that of the realized volatility after 30 lags. Past studies, e.g. Kirman and Teyssiere (2002), use investor herding to explain the long memory in volatility. The long-run dependence in information risk measures may provide an alternative explanation.

IV.2 A dynamic model for information risk

We now specify a dynamic model for information risk base on its own dynamics and explore its relationship with other information proxies. Given the long memory evident in Figure 3, a natural starting point is the Heterogeneous Autoregressive (HAR) model used by Corsi (2004) to capture the long memory in realized volatility (RV). The baseline HAR-RV model is given by $RV_t^D = \text{const} + \sum_{i=D}^{Q} \beta^i RV_{t-1}^i + \varepsilon_t$ where RV_t^i, i = D, W, *M*, *Q*, is the average realized volatility over the past 1, 5, 21, and 63 trading days respectively including day t. The model is motivated by the observation that investors have heterogeneous investment horizons (see Müller, et al. 1993, 1997) and provides a simple and parsimonious specification to capture the long memory effect of volatility. It has been adopted for modelling of the volatility dynamics in equity, bond, and currencies (e.g. Andersen, Bollerslev, Diebold, 2006, Andersen, Bollerslev, Huang, 2006, Wang and Yang, 2007).

We adopt the HAR specification for the dynamics of the estimated daily α. To understand what affects daily α in addition to its past values, we add daily realized volatility (RV_t), trading volume (S_t), and the number of trades (T_t) of each stock, as well as the market volatility index (VIX_t). Since α should be

positive at any time, we remove days with negative α. We use log transformation of all variables to mitigate the impact from extreme values. Our model for the dynamics of the daily α is given by

$$\ln(\alpha_t^D) = \text{const} + \sum_{i=D}^{Q} \beta^i \ln(\alpha_{t-1}^i) + \lambda^V \ln(RV_t) + \lambda^M \ln(VIX_t) + \lambda^T \ln(T_t) + \lambda^S \ln(S_t) + \varepsilon_t$$

Table 8 reports the estimates of the baseline HAR-α model without the volatility and trade variables for HW and MRR. Consistent with Figure 3, the lagged average α remains highly significant up to quarterly interval. On average, this persistence in α captures 62% of the variations in α. The baseline model is able to remove the long-run persistence at 40 lags or beyond for all stocks in HW model, and for 81% of stocks in MRR model. The difference between HW and MRR is not surprising given the higher persistence in MRR in Figure 3.

Table 9 reports the results for several versions of the extended HAR-α model for HW. Coefficients of past αs are highly significant and are not reported here. Panel A reports four specifications where the volatility and trade variables are added individually. The realized volatility of individual stocks is significant for all stocks and the market volatility is significant for 15 of the 16 stocks (93.8%). The estimated α increases with the number of trades with an average t-stats 3.83, and decreases with trading volume with an average t-stats -3.65. However, only 6 stocks (37.5%) have a t-statistic greater than 1.96 for the number of trades and 7 stocks (43.8%) for trading volume.

Panel B of Table 9 reports joint estimations of volatility and trade variables. Model V shows that trading volume is negative and significant for all

stocks when the stock's realized volatility is included. Model VI shows that when jointly estimated, trading volume and the number of trades are highly significant for 81% and 87.5% of the stocks respectively, much higher than when either is included alone. The strong negative effect from trading volume suggests that volumes are primarily driven by liquidity needs, not private information, lending support to the stealth trading hypothesis of Barclay and Warner (1993). In model VIII, the number of trades is decomposed into two components using a simple AR(10) model with the AR component as the expected number of trades, and the residual as the surprise component. As in previous studies [see Schwert (1990), Bessembinder and Seguin (1993) and Daigler and Wiley (1999)], we do not attempt to find an ideal model for each stock but instead select an arbitrarily long set of autoregressive coefficients.

In model VI, we estimate how much the information risk reflects market-wide uncertainty versus firm-specific uncertainty. We separate a stock's daily volatility into market-driven volatility and firm-specific volatility. A market model $r_t = a + br_{m,t} + e_t$ is estimated for each stock. The market-driven volatility is $\hat{b}\sigma_{m,t}$ where $\sigma_{m,t}$ is approximated by VIX_t. The firm-specific volatility is measured as $|\hat{e}_t|$. The results show that both components are highly significant for all stocks. The coefficient of the market volatility is much larger than that of the firm-specific volatility. The pattern holds for other specifications in Table 9.

V. Conclusion

In this paper we perform a time-series analysis on daily estimates of adverse selection costs and uncover previously undocumented features related to the temporal variation of adverse selection costs. For the first time, spread decomposition models are used to estimate adverse selection costs at a frequency permitting time-series analysis. The resulting econometric specification models adverse selection costs within the framework of recent volatility models and relate its variation to measures of both volatility and volume.

We find of long memory in the time-series evolution of adverse selection costs. Not only do adverse selection costs exhibit significant autocorrelation, this autocorrelation is highly persistent and has significant effects up to the quarterly horizon. We observe adverse selection costs aggregated at the weekly horizons to have the largest effect on current values of adverse selection costs. Consistent with our hypotheses, both market volatility and firm-specific volatility have positive and significant individual effects on adverse selection costs. Volume measured by the share volume displays a consistently negative relationship with adverse selection costs. After controlling the model for liquidity effects, the inclusion of the number of trades appears to proxy for the activity of informed traders. Not unexpectedly we identify strong relations between adverse selection costs, volatility and volume.

Bibliography

Affleck-Graves, J., S. P. Hegde, and R. E. Miller, 1994, "Trading Mechanisms and the Components of the Bid-Ask Spread," *Journal of Finance*, 49, 4, 1471-88.

Andersen, T., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43-76.

Andersen, T. G., T. Bollerslev and F. X. Diebold, 2006, "Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," *Review of Economics and Statistics*, forthcoming.

Andersen, Bollerslev, Huang, 2006, Semiparametric Framework for Modelling and Forecasting Jumps and Volatility in Speculative Prices

Andrews, D. W. K., 1991, "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Barclay, M.J., Warner, J.B., 1993, "Stealth Trading and Volatility: Which Trades Move Prices?," *Journal of Financial Economics* 34(3), December, 281-305

Bessembinder, H. and P. J. Seguin, 1993, "Price Volatility, Trading Volume, and Market Depth: Evidence from Futures Markets," *The Journal of Financial and Quantitative Analysis*, 28, 1, 21-39.

Clarke, J. and K. Shastri, 2000, "On Information Asymmetry Metrics," *University of Pittsburgh*, working paper.

Easley, D., S. Hvidjkaer, and M. O'Hara, 2002, "Is Information Risk a Determinant of Asset Returns?" *Journal of Finance*, 57, 2185–2222.

Corsi, F., 2004, "A Simple Long Memory Model of Realized Volatility," SSRN working paper http://ssrn.com/abstract=626064.

Daigler, R.T., and Wiley, M.K., 1999, "The Impact of Trader Type on the Futures Volatility-Volume Relation," *Journal of Finance* 54(6), December, 2297-2316

Easley, D., S. Hvidjkaer, and M. O'Hara, 2002, "Is Information Risk a Determinant of Asset Returns?" *Journal of Finance*, 57, 2185–2222.

Easley, D., and M. O'Hara, 2004, "Information and the Cost of Capital," *Journal of Finance*, 59, 1553–1583.

Fama, E. F. and K. R. French, 1992, "The Cross Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

Glosten, L. R. and P. R. Milgrom, 1985, "Bid, Ask and Transactions Prices in a Specialist Model with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71–100.

Henker, T., and J. Wang, 2006, "On the Importance of Timing Specifications in Market Microstructure Research," *Journal of Financial Markets*, 9, 2, 162-179.

Huang, R.D., and Stoll, H.R., 2004, "Market Microstructure and Stock Return Predictions," *Review of Financial Studies* 7(1), Spring, 179-213.

Huang, R. D., and H. R. Stoll, 1997, "The Components of the Bid-Ask Spread: A General Approach," *Review of Financial Studies*, 10, 995-1034.

Jones, C. M., G. Kaul and M. L. Lipson, 1994, "Transactions, Volume, and Volatility," *Review of Financial Studies*, 7, 631-651.

Kirman and Teyssiere (2002), Microeconomic Models for Long Memory in the

Volatility of Financial Time Series, *Studies in Nonlinear Dynamics and Econometrics,* Volume 5, Number 4

Kyle, A. S., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53, 6, 1315-36.

Lee, M. C. and M. J. Ready, 1991, "Inferring Trade Direction From Intraday Data," *Journal of Finance*, 46, 2, 733-746.

Leland, H. E. and D. H. Pyle, 1976, "Informational Asymmetries, Financial Structure, and Financial Intermediation," *The Journal of Finance*, 32, 2, 371-387.

Lin, J., G. Sanger, and G. G. Booth, 1995, "Trade Size and Components of the Bid-Ask Spread," *Review of Financial Studies*, 8, 4, 1153-83.

Madhavan, A., M. Richardson, and M. Roomans, 1997, "Why do security prices change? A transaction-level analysis of NYSE stocks," *Review of Financial Studies*, 10, 4, 1035-64.

Müller, U., M. Dacorogna, R. Dav, R. Olsen, O. Pictet and J. von Weizsacker, 1997, "Volatilities of Different Time Resolutions - Analysing the Dynamics of Market Components," *Journal of Empirical Finance*, 4, 213-239.

Myers, S. C. and N. Majluf, 1984, Corporate Financing and Investment Decisions When Firms Have Information Investors Do Not Have," *Journal of Financial Economics*, 13, 187-221.

Myers, S. C. and N. Majluf, 1984, Corporate Financing and Investment Decisions When Firms Have Information Investors Do Not Have," *Journal of Financial Economics*, 13, 187-221.

Newey, W. K. and K. D. West, 1987, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Ross, S., 1977, "The Determination of Financial Structure: The Incentive-Signaling Approach," *Bell Journal of Economics*, 8, 23-40.

Schwert, G.W., 1990, "Stock Volatility and the Crash of '87," *Review of Financial Studies* 3(1), 77-102

Van Ness, B., R. Van Ness and R. Warr, 2000, "How Well Do Adverse Selection Components Measure Adverse Selection?" *Financial Management*, 30, 77-98.

Wang, J., and M. Yang, 2008, Asymmetric Volatility in the Foreign Exchange Markets, *Journal of International Financial Markets, Institutions, and Money*, forthcoming.

Table 1: Descriptive Statistics – Intra-day Data

This table reports, for each stock, the total number of observations in the sample, the average number of observations per day, the minimum number of observations across all days in the sample, mean transaction price and quoted spread and the standard deviation of these variables.

Stock	Total # of Obs.	Mean # of Obs. per day	Min. # of Obs.	Mean Price	Std. Dev. Price	Mean Spread	Std. Dev. Spread
AXP	2,180,199	2,497	647	41.63	7.56	0.0301	0.0320
DD	1,966,783	2,255	610	42.56	2.60	0.0291	0.0294
DIS	2,672,374	3,065	699	21.40	3.55	0.0239	0.0308
DOW	1,503,296	1,724	390	35.85	6.34	0.0284	0.0430
EK	1,273,621	1,459	438	29.60	4.50	0.0276	0.0303
GE	5,827,936	7,187	1,971	30.91	3.74	0.0246	0.0320
GM	2,053,323	2,352	487	43.70	7.86	0.0293	0.0345
IBM	4,323,239	4,952	1,423	88.47	12.35	0.0468	0.0538
IP	1,422,483	1,629	525	39.39	2.98	0.0276	0.0274
JNJ	3,075,982	3,523	1,145	54.97	4.02	0.0346	0.0375
KO	2,197,589	2,517	790	46.39	4.34	0.0261	0.0256
MCD	1,843,651	2,112	612	24.25	4.95	0.0231	0.0244
MMM	2,154,024	2,467	473	105.07	22.65	0.0498	0.0644
MO	2,921,688	3,347	857	46.15	7.29	0.0296	0.0387
MRK	3,171,017	3,632	1,083	48.75	10.84	0.0320	0.0591
PG	2,429,453	2,783	725	83.30	16.58	0.0367	0.0417
WMT	3,375,147	3,866	1,227	54.43	3.43	0.0341	0.0440
XOM	3,467,507	3,972	949	39.97	6.21	0.0256	0.0281

Table 2: Realized Volatility

This table reports summary statistics for the daily realized volatility of each stock and for the market volatility index VIX. RV_05 and RV_30 refer to daily Realized Volatility calculated using 5 and 30 minute returns, respectively.

	AXP	**DD**	**DIS**	**DOW**	**EK**	**GE**	**GM**	**IBM**	**IP**	**JNJ**	**KO**	**MCD**	**MO**	**MMM**	**MRK**	**PG**	**WMT**	**XOM**	**Average**
RV_05																			
Mean	0.0177	0.0163	0.0218	0.0183	0.0198	0.0179	0.0170	0.0155	0.0163	0.0146	0.0139	0.0184	0.0136	0.0150	0.0163	0.0119	0.0153	0.0147	**0.0163**
Med.	0.0155	0.0146	0.0194	0.0164	0.0175	0.0159	0.0148	0.0133	0.0151	0.0129	0.0127	0.0164	0.0120	0.0128	0.0139	0.0110	0.0137	0.0124	**0.0145**
Std. Dev.	0.0107	0.0075	0.0111	0.0089	0.0101	0.0100	0.0078	0.0084	0.0060	0.0076	0.0060	0.0088	0.0066	0.0085	0.0089	0.0053	0.0069	0.0077	**0.0081**
Skew.	1.70	1.65	2.09	2.16	2.69	2.09	1.63	2.35	1.27	2.15	2.60	2.34	3.15	2.79	2.72	3.59	1.98	2.16	**2.28**
Kurt.	4.68	3.97	9.35	7.73	10.63	7.28	3.24	9.44	2.74	7.56	11.88	8.73	18.63	11.07	11.13	31.21	6.09	8.16	**9.64**
RV_30																			
Mean	0.0199	0.0175	0.0192	0.0170	0.0170	0.0148	0.0154	0.0136	0.0131	0.0173	0.0128	0.0142	0.0156	0.0111	0.0146	0.0133	0.0155	0.0199	**0.0175**
Med.	0.0178	0.0153	0.0165	0.0143	0.0151	0.0125	0.0141	0.0117	0.0119	0.0151	0.0111	0.0118	0.0135	0.0099	0.0129	0.0112	0.0135	0.0178	**0.0153**
Std. Dev.	0.0110	0.0100	0.0110	0.0099	0.0086	0.0087	0.0065	0.0076	0.0064	0.0092	0.0069	0.0096	0.0090	0.0056	0.0070	0.0071	0.0084	0.0110	**0.0100**
Skew.	2.18	2.80	2.47	1.99	1.59	2.20	1.40	2.09	2.46	2.18	2.98	3.09	2.51	3.42	1.95	2.04	2.24	2.18	**2.80**
Kurt.	10.57	13.07	9.42	6.75	3.79	8.35	3.54	6.86	11.30	7.06	16.95	14.03	9.71	27.50	6.06	5.67	9.25	10.57	**13.07**
	Mean	*Med.*	*Stdev.*	*Skew.*	*Kurt.*														
VIX	22.24	20.29	7.07	0.877	3.016														

Table 3: Henker and Wang (2006)

This table reports HW results after observations with π less than 0.514 are discarded. 1 day sample periods are used in a block regression method. All statistics are calculated across time and individually per stock. Panel A describes the length of the time series, the percentage of sample remaining after discarding atheoretical estimates, the mean number of observations per estimate, the average quote reversal probability (π) and its standard deviation. Panel C reports the results for the ACS specification. Dollar α quotes are in cents and percentage α quotes are in basis points. The final column calculates an equally weighted average of each statistic across the sample of stocks. #Calculated average does not include IBM and MMM.

	AXP	DD	DIS	DOW	EK	GE	GM	IBM	IP	JNJ	KO	MCD	MMM	MO	MRK	PG	WMT	XOM	Average#
No. of Estimates	725	714	870	761	733	870	792	509	655	818	836	860	320	855	809	577	783	851	**782**
% remaining	83.1%	82.0%	99.9%	87.4%	84.1%	99.8%	90.8%	58.4%	75.1%	93.8%	95.9%	98.6%	36.7%	98.1%	92.8%	66.2%	89.8%	97.6%	**89.7%**
Mean No. of Obs.	2463	2332	3055	1752	1531	7160	2351	4714	1684	3462	2523	2110	2965	3273	3679	2888	3814	3954	**3002**
Mean π	0.585	0.578	0.634	0.588	0.605	0.668	0.600	0.545	0.574	0.606	0.604	0.630	0.554	0.600	0.606	0.561	0.589	0.643	**0.604**
Std. Dev. π	0.043	0.039	0.052	0.045	0.058	0.075	0.048	0.023	0.041	0.054	0.046	0.052	0.024	0.040	0.066	0.041	0.044	0.076	**0.051**
Mean $ α (¢)	1.309	1.456	0.659	1.254	1.230	0.399	1.205	2.187	1.353	1.052	1.174	0.665	1.747	0.961	1.351	1.895	1.376	0.985	**1.145**
Median $ α (¢)	0.937	1.031	0.536	0.861	0.731	0.232	0.837	1.755	1.033	0.667	0.831	0.455	1.499	0.799	0.815	1.517	0.997	0.512	**0.799**
Std. Dev. $ α	1.145	1.253	0.432	1.263	1.234	0.345	1.047	1.526	1.084	0.974	1.161	0.652	1.355	0.635	1.505	1.494	1.140	1.105	**1.029**
Mean % α (bps)	3.872	3.424	3.372	3.802	3.998	1.253	2.645	2.368	3.487	1.884	2.481	2.986	2.199	2.764	2.612	4.251	2.540	2.610	**2.999**
Med. % α (bps)	2.476	2.389	2.660	2.399	2.509	0.769	1.916	1.952	2.541	1.201	1.740	2.144	1.843	2.300	1.531	3.162	1.836	1.322	**2.056**
Std. Dev. % α	3.915	2.991	2.496	4.167	3.832	1.092	2.121	1.601	2.950	1.761	2.377	2.912	1.958	1.883	2.995	3.647	2.153	3.123	**2.776**
% Neg. Obs.	0.1%	0.1%	0.0%	0.5%	0.4%	0.0%	0.3%	0.0%	0.9%	0.2%	0.0%	0.1%	0.0%	0.0%	0.0%	0.3%	0.0%	0.0%	**0.2%**
% Neg. & Sig. Obs.	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	**0.0%**

Table 4: Madhavan, Richardson and Roomans (1997)

This table reports the dynamic estimation results for the Madhavan, Richardson and Roomans (1997) model for both 1 day (block regression) and 5 day (1 day rolling window) sample periods. All statistics are calculated across time and individually per stock. Panel A describes, for 1 day sample results, the length of the time series; the mean number of observations per estimate; the mean, median and standard deviation of dollar and percentage α estimates; and the percentage of negative and significantly negative (5 percent level) α estimates. Panel C reports the results for the 5 day sample results. Dollar α quotes are in cents and percentage α quotes are in basis points. The final column calculates an equally weighted average of each statistic across the sample of stocks.

	AXP	DD	DIS	DOW	EK	GE	GM	IP	JNJ	KO	MCD	MO	MRK	PG	WMT	XOM	Average
No. of Estimates	872	871	871	871	872	872	872	872	872	872	872	872	872	872	872	872	**872**
Mean No. of Obs.	2500	2258	3068	1726	1461	7194	2355	1631	3528	2520	2114	3351	3636	2786	3871	3976	**2998**
Mean $ α (¢)	0.368	0.445	0.223	0.402	0.429	0.065	0.367	0.494	0.221	0.349	0.259	0.280	0.258	0.502	0.242	0.220	**0.320**
Median $ α (¢)	0.332	0.373	0.204	0.332	0.368	0.063	0.294	0.418	0.208	0.291	0.223	0.255	0.226	0.450	0.221	0.193	**0.278**
Std. Dev. $ α	0.141	0.216	0.088	0.200	0.235	0.044	0.198	0.225	0.115	0.185	0.130	0.126	0.180	0.236	0.107	0.102	**0.158**
Mean % α (bps)	1.067	1.053	1.111	1.212	1.388	0.204	0.812	1.269	0.396	0.734	1.150	0.790	0.500	1.159	0.446	0.567	**0.866**
Med. % α (bps)	0.877	0.870	0.980	0.944	1.246	0.203	0.674	1.076	0.365	0.624	1.022	0.727	0.425	0.973	0.404	0.488	**0.744**
Std. Dev. % α	0.539	0.529	0.502	0.704	0.697	0.134	0.404	0.611	0.204	0.362	0.552	0.338	0.343	0.649	0.204	0.283	**0.441**
% Neg. Obs.	0.0%	0.0%	0.6%	0.1%	0.1%	4.8%	0.0%	0.0%	0.5%	0.0%	0.3%	0.3%	1.7%	0.0%	0.1%	0.2%	**0.55%**
% Neg. & Sig. Obs.	0.0%	0.0%	0.0%	0.1%	0.0%	1.3%	0.0%	0.0%	0.1%	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	**0.11%**

Table 5: Lin, Sanger and Booth (1995)

This table reports the dynamic estimation results for the Lin, Sanger and Booth (1995) model for both 1 day (block regression) and 5 day (1 day rolling window) sample periods. All statistics are calculated across time and individually per stock. Panel A describes, for 1 day sample results, the length of the time series; the mean number of observations per estimate; the mean, median and standard deviation of dollar and percentage α estimates; and the percentage of negative and significantly negative (5 percent level) α estimates. Panel C reports the results for the 5 day sample results. Dollar α quotes are in cents and percentage α quotes are in basis points. The final column calculates an equally weighted average of each statistic across the sample of stocks.

	AXP	DD	DIS	DOW	EK	GE	GM	IP	JNJ	KO	MCD	MO	MRK	PG	WMT	XOM	Average
No. of Estimates	872	871	871	871	872	872	872	872	872	872	872	872	872	872	872	872	**872**
Mean No. of Obs.	2500	2258	3068	1726	1461	7194	2355	1631	3528	2520	2114	3351	3636	2786	3871	3976	**2998**
Mean $ α (¢)	0.493	0.607	0.342	0.561	0.626	0.137	0.486	0.646	0.334	0.448	0.386	0.379	0.346	0.644	0.334	0.289	**0.441**
Median $ α (¢)	0.490	0.535	0.338	0.497	0.556	0.133	0.410	0.601	0.313	0.417	0.365	0.364	0.314	0.593	0.320	0.266	**0.407**
Std. Dev. $ α	0.166	0.252	0.103	0.279	0.306	0.057	0.260	0.242	0.146	0.198	0.180	0.160	0.194	0.297	0.139	0.118	**0.194**
Mean % α (bps)	1.409	1.431	1.684	1.683	2.029	0.429	1.065	1.649	0.601	0.948	1.720	1.075	0.675	1.484	0.615	0.740	**1.202**
Med. % α (bps)	1.272	1.264	1.628	1.411	1.876	0.418	0.949	1.521	0.556	0.902	1.624	1.040	0.608	1.257	0.586	0.675	**1.099**
Std. Dev. % α	0.619	0.618	0.566	0.949	0.862	0.167	0.494	0.653	0.259	0.394	0.749	0.444	0.345	0.833	0.262	0.319	**0.533**
% Neg. Obs.	0.00%	0.00%	0.00%	0.11%	0.00%	0.11%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**0.01%**
% Neg. & Sig. Obs.	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	**0.00%**

Table 6: Average Cross-Model Contemporaneous Correlations

***, ** and * refer to 1, 5 and 10 percent significance levels, respectively.

		HW	LSB	MRR
HW	$\alpha^{\%}$	1		
	$\alpha^{\$}$			
LSB	$\alpha^{\%}$	0.569***	1	
	$\alpha^{\$}$	0.546***		
MRR	$\alpha^{\%}$	0.685***	0.804***	1
	$\alpha^{\$}$	0.668***	0.792***	

Table 7: Rank Correlations between Static and Dynamic αs

This table reports Spearman's rank correlations between average dynamic α estimates and static α estimates across the sample of stocks. Results are calculated from percentage α's. The asterisks **and *** indicate 5 and 10 percent significance levels against the null of ρ=1.

	HW	LSB	MRR
Jul 01 – Dec 01	0.932**	0.862***	0.988
Jan 02 – Jun 02	0.938**	0.950**	0.965
Jul 02 – Dec 02	0.679***	0.903**	0.979
Jan 03 – Jun 03	0.977	0.885***	0.979
Jul 03 – Dec 03	0.924**	0.950**	0.968
Jan 04 – Jun 04	0.988	0.924**	0.991
Jul 04 – Dec 04	0.991**	0.915**	0.985

Table 8: The Baseline HAR-α Model

This table reports the estimation results of the following HAR$-\alpha$ model:

$$log(\alpha_t) - c_0 + \sum_{i=D}^{Q} \beta^i log\left(\alpha_{t-1}^i\right) + \varepsilon_t$$

Mean coefficient estimates are presented with t-statistics and the proportion of underlying individual coefficient estimates significant at the 5 percent level in parentheses. The average R-squared is presented with the 5th and 95th percentile values in parentheses. The last row presents the percentage of the sample with statistically insignificant Q^{40} values. ***, ** and * refer to 1, 5 and 10 percent significance levels, respectively. Critical values correspond to Student's t distribution with 15 degrees of freedom.

Explanatory Variables	**HW**	**MRR**
Lagged daily α	0.135*** (6.80, 62.5%)	0.273*** (14.54, 93.8%)
Lagged weekly α	0.338*** (8.90, 93.8%)	0.332*** (11.97, 100%)
Lagged monthly α	0.317*** (8.18, 75.0%)	0.187*** (6.09, 62.5%)
Lagged quarterly α	0.170*** (6.09, 56.3%)	0.157*** (6.51, 50.0%)
Constant	-0.364*** (-6.54, 25.0%)	-0.513*** (-3.71, 43.8%)
Mean R-squared	0.620 (0.428, 0.839)	0.621 (0.308, 0.824)
Mean No. Obs.	691	803
Q^{40}	100%	81.25%

Table 9: Model extensions – Henker and Wang (2006) model

This table reports the estimation results of the model extensions on HW estimated α's. Panel A corresponds to four individual regressions of α on each individual trade measure (Realized Volatility, VIX, share volume and the number of trades) with the HAR$-\alpha$ specification. Mean coefficients are reported with t-stats and the percentage of estimates significant at the 5 percent level in parentheses. Only for Panel A, figures in brackets represent mean adjusted R^2 values. Panel B represents Specifications V – X where various hypotheses are tested. The final column represents mean adjusted R^2 values with the 5th and 95th percentile values in parentheses. The final specification is represented in the following equation:

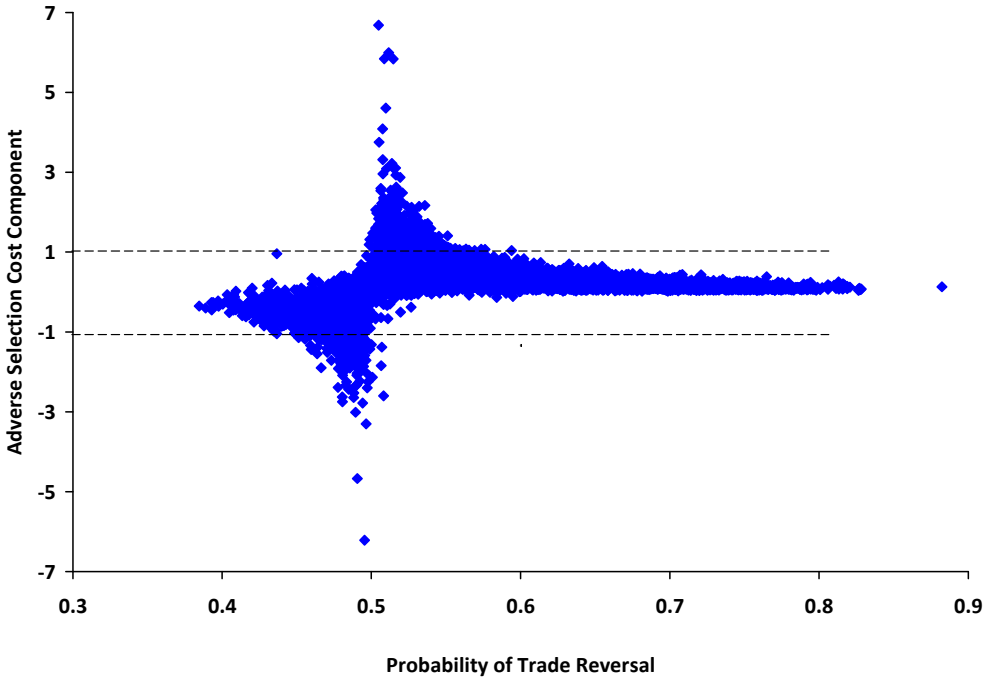
$$log(\alpha_t) = c_0 + \sum_{i=D}^{Q} \beta^i log(\alpha_{t-1}^i) + \theta_1 log(\sigma_{m,t}) + \theta_2 log(\sigma_{s,t}) + \lambda_1 log\left(S_{Volume_t}\right) + \phi_1 log\left(N_{Trades_t}^{Exp}\right) + \phi_2 log\left(N_{Trades_t}^{Unexp}\right) + \xi_t$$

***, ** and * refer to 1, 5 and 10 percent significance levels, respectively. Critical values correspond to Student's t distribution with 15 degrees of freedom.

Spec.	Realized Volatility	VIX	Residual Volatility	# of Trades	Share Volume	Adj. R^2
PANEL A						
I – IV	0.409*** [0.649] (9.11, 100%)	0.604*** [0.632] (8.19, 93.8%)		0.144*** [0.623] (3.83, 37.5%)	-0.072*** [0.622] (-3.65, 43.8%)	N.A.
PANEL B						
V	0.640*** (15.15, 100%)				-0.342*** (-14.52, 100%)	0.668 (0.467, 0.855)
VI		1.014*** (13.32, 100%)	0.595*** (14.38, 100%)		-0.346**** (-14.26, 100%)	0.674 (0.486, 0.856)
VII				0.355*** (10.23, 87.5%)	-0.237*** (-9.45, 81.3%)	0.631 (0.438, 0.844)
VIII		1.01*** (13.26, 100%)	0.584*** (13.88, 100%)	0.148*** (3.19, 25.0%)	-0.406*** (-13.00, 100%)	0.677 (0.487, 0.856)
IX					-0.278*** (-10.04, 87.5%)	0.634 (0.445, 0.846)
X		1.02*** (13.21, 100%)	0.558*** (13.75, 100%)		-0.430*** (-12.60, 100%)	0.678 (0.493, 0.857)

Figure 1: Adverse Selection Component and
Probability of Trade Reversal

Panel A: Estimated αs and the probability of trade reversal



Panel B: Bad αs and the probability of trade reversal

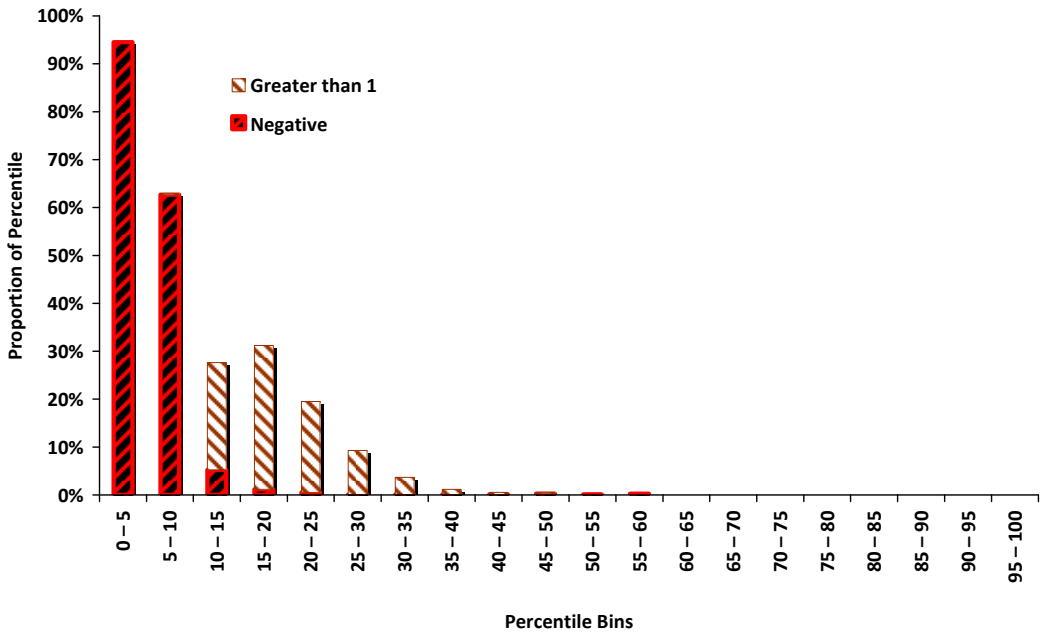
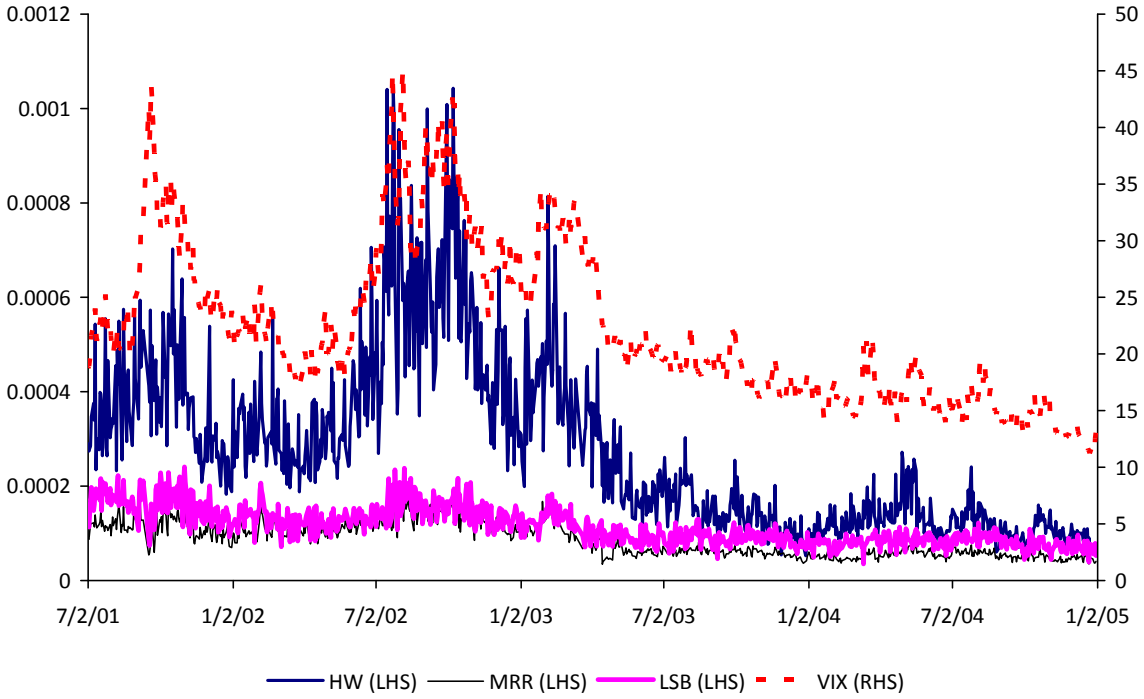


Figure 2: Daily *A*s and VIX

Panel A: Time-series plot



Panel B: Adverse selection risk and market volatility

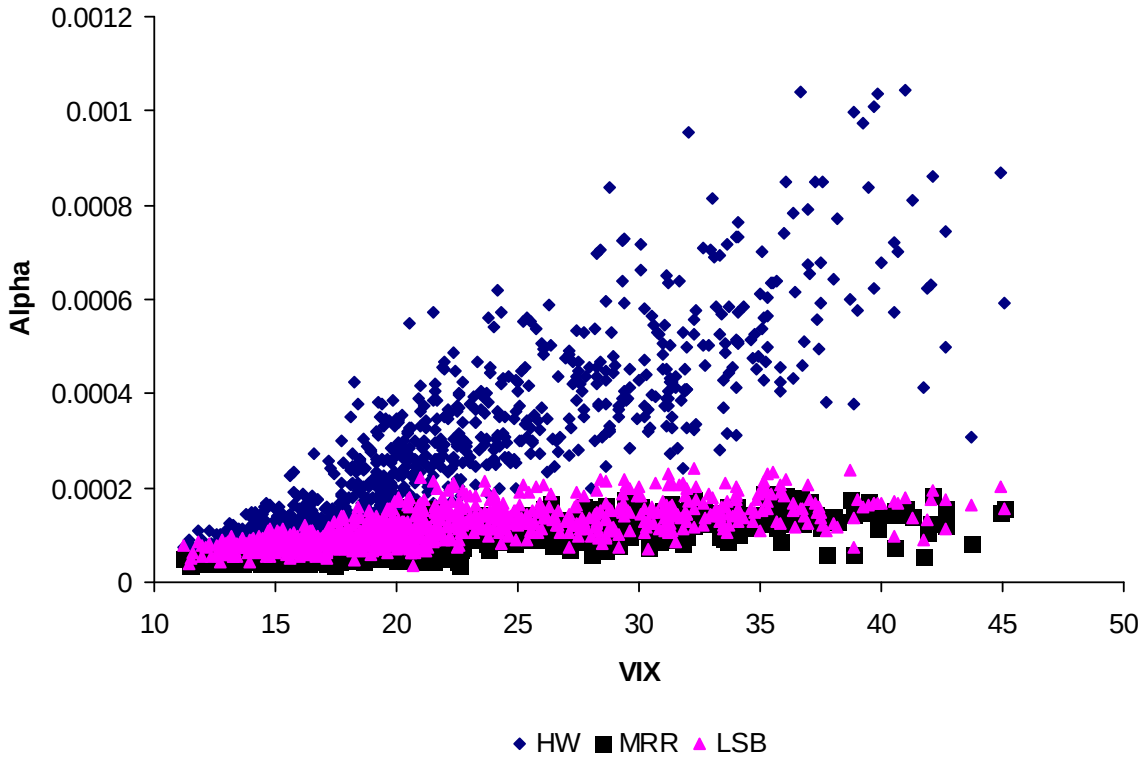


Figure 3: Long Memory in As

