DOI: 10.1111/j.1475-679X.2009.00344.x
Journal of Accounting Research
Vol. 47 No. 5 December 2009
Printed in U.S.A.

Measuring Investors' Opinion Divergence

JON A. GARFINKEL*

Received 20 May 2008; accepted 26 June 2009

ABSTRACT

Numerous proxies for divergence of investors' opinions have been suggested in the empirical accounting and finance literatures. I offer a new proxy constructed from proprietary limit order and market order data. This allows me to capture *additional* information on investors' *private* valuations. Proxies from the extant literature, based on publicly available data, do not contain such information. Given my new measure, I ask which of the extant proxies correlates best with it. In my regression analysis, unexplained volume is the best proxy for opinion divergence. Conditioning on various firm-specific and order-specific characteristics generally does not change this conclusion. The main exception is the sample of firms without IBES forecast dispersion data, for which bid-ask spread is the best proxy for opinion divergence. Factor analysis also suggests that unexplained volume is the preferred proxy for opinion divergence.

1. *Introduction*

Accounting and finance scholarship has long been interested in the effects of heterogeneous investor expectations. In the accounting literature, such varying opinions affect the inferences from the vast "information content

*Henry B. Tippie College of Business, University of Iowa. I thank Matt Billett, Shane Corwin, Bart Danielson, Kristine Hankins, Todd Houge, Rick Mergenthaler, George Neumann, Tom Rietz, Gene Savin, Ashish Tiwari, and Anand Vijh for helpful discussions. I thank Yi Jiang and Tom Barkley for research assistance. I thank Paul Bennett and the New York Stock Exchange for allowing me to use the order data in this study. I am indebted to Marc Lipson for help through both data provision and explanations. I am especially grateful to an anonymous referee for many helpful suggestions, as well as the Editor (Ray Ball). All remaining errors are my own.

research" agenda, in particular the market reaction to earnings announcements.[1] In the finance literature, divergent investor opinions have implications for asset prices and returns. Thus, accurate measurement of divergent investor opinions is an important research goal.

Unfortunately, the numerous proxies for investor opinion divergence found in the literature are just that—proxies.[2] To date, there is little research into the best way to measure private opinions that are almost inherently unobservable. This paper recognizes that part of the difficulty in constructing such a measure is the requisite use of publicly available data, none of which *directly* conveys investors' *private* valuations of assets.

I offer a partial solution to the problem. I use proprietary data on investors' orders in individual New York Stock Exchange (NYSE) stocks to construct a new measure of opinion divergence. My new measure uses investors' limit and market orders as indications of their private valuations. I then assess which extant proxy based on publicly available data aligns best with the new measure.[3] My work is similar in spirit to the construct validity analysis of Core and Guay [2002].

The proxies I examine are: two measures of unexplained volume, bid-ask spread, stock return volatility, and two measures of analyst forecast dispersion.[4] My results have implications for the optimal choice of public-data proxy for investor opinion divergence.

Unexplained volume appears to be the best proxy for investor opinion divergence in most cases. For the main sample, with data available on the six main proxies, regressions with the unexplained volume proxy "change in turnover" indicate greater explanatory power than regressions with any other proxy. Moreover, the next best proxy is an alternative measure of

---

[1] See Lev and Ohlson [1982] for a nice review.

[2] The accounting literature focuses primarily on analyst forecast variation and unexplained volume proxies for investor opinion divergence. See Ajinkya, Atiase, and Gift [1991], Abarbanell, Lanen, and Verrecchia [1995], Byard [2002], Diether, Malloy, and Scherbina [2002], Doukas, Kim, and Pantzalis [2006], Zhang [2006a,b], Alexandridis, Antoniou, and Petmezas [2007], Scherbina and Sadka [2007], Lerman, Livnat, and Mendenhall [2008], Barron, Stanford, and Yu [2009] and Chatterjee, John, and Yan [2009] for analyst-based proxies. Beaver [1968], Bamber [1987], Kandel and Pearson [1995], Ajinkya, Atiase, and Gift [2004], and Garfinkel and Sokobin [2006] suggest that unexplained volume may proxy for divergent investor opinions. The finance literature has suggested viewing bid-ask spread (Bagehot [1971], Houge et al. [2001], and Handa, Schwartz, and Tiwari [2003]), and stock return volatility (Frankel and Froot [1990], Harris and Raviv [1993], Shalen [1993], Wang [1998], Daigler and Wiley [1999], and Boehme, Danielson, and Sorescu [2005]) as proxies for diverging investor opinions.

[3] One interpretation of this approach is that my new measure is a better proxy for opinion divergence than extant proxies. Alternatively, one could argue that any reasonable extant proxy must at least correlate with my new measure (in the correct direction) to be considered acceptable.

[4] Given research decomposing spread into asymmetric information (AI) and other components, my robustness checks examine the relation between my new measure and the "AI" and other components of spread, as well as the related metric of PIN (probability of informed trading—see Easley, Hvidkjaer, and O'Hara [2002]).

unexplained volume—"standardized unexplained volume." The other four main proxies for opinion divergence indicate significantly weaker (and in some cases wrong direction) explanatory power in their regressions.

I conduct conditional analyses to offer other perspectives. I find that for firms without IBES analyst forecast dispersion data, bid-ask spread is the best proxy for opinion divergence. This is consistent with work by Hong, Lim, and Stein [2000], which suggests that analyst coverage changes the information environment for a stock. When I categorize by the level of other firm-specific characteristics such as firm size, firm volume, and firm stock price, I conclude that unexplained volume is a better proxy for opinion divergence than other proxies. At least for the main sample (with main proxies available, particularly forecast dispersion), variation in firm-level characteristics does not appear to influence the optimality of unexplained volume as a proxy for investor opinion divergence. On the other hand, when I categorize by the type of order submitted (program-trade or nonprogram-trade) or the relative magnitude of the probability of informed trading (PIN), there is evidence that under certain circumstances bid-ask spread is as good a proxy as unexplained volume for opinion divergence.

I also address the reasonableness of proxies for investor opinion divergence using factor analysis. I extract the first common factor from the six typical proxies and examine its correlation with my new measure of opinion divergence. Not only is it positive and significant, but it appears to be a stronger correlate than unexplained volume. I also examine the first common factor's relationship with each proxy (raw correlations with, and weights needed to construct, the first common factor). The unexplained volume proxies are the only two with positive correlations and weights, again suggesting the strength of them as proxies for investor opinion divergence.

Finally, I exploit research suggesting significant news events may actually engender diverging investor opinions (see e.g. theories by Holthausen and Verrecchia [1990] and Kim and Verrecchia [1994]).[5] Specifically, for my sample of firms with earnings announcements during the data availability window (January–March 2002), I construct my new measure of opinion divergence over the three-day window surrounding the announcement date. I again find that unexplained volume appears to be the best proxy for investor opinion divergence.

Overall, I draw the following conclusions. For researchers choosing among public-data proxies for opinion divergence, my results strongly suggest that unexplained volume is a good choice. In regression tests, my new construct for opinion divergence is most highly correlated with two measures of unexplained volume. Moreover, these tests also reveal that some of the typical proxies for opinion divergence—forecast dispersion and stock return volatility—are *negatively* related to my construct. This suggests that

---

[5] Evidence from trading activity following unanticipated dividend announcements is somewhat consistent with these theories (Graham, Koski, and Loewenstein [2006]). Suggestive evidence is also presented in Fleming and Remolona [1999].

prior work treating (for example) analyst forecast dispersion as an indicator of dispersed *investor* opinions may contain biased inferences.[6,7] Finally, there may be benefits to combining the usual proxies for opinion divergence. The first principal component drawn from all six typical proxies for opinion divergence is actually a better explainer than the best single proxy (unexplained volume). Nevertheless, it's notable that this factor is significantly related to only the unexplained volume proxies.

The remainder of this research is organized as follows. Section 2 describes my measure of investor opinion divergence and why it's likely to be functional. This includes its theoretical appeal and its construction. Section 2 also provides evidence that my construct indeed measures investor opinion divergence. In section 3, I introduce the six typical proxies for opinion divergence. Each proxy is discussed in terms of its theoretical and empirical foundation as a measure of investor opinion divergence and its construction. Section 4 presents results and section 5 concludes.

2. *Measuring Investors' Opinion Divergence*

Data on investors' orders offers potentially significant advantages in the construction of a measure of investor opinion divergence. First, limit orders contain requested prices, offering an improved view of investors' private valuations. Second, market orders may also contain information about the trader's reservation price, in the context of theoretical and empirical work on order submission strategies.[8]

2.1 PRECEDENT

Theoretical justification for using orders as indications of investors' interests is found in several papers. Foucalt [1999], Harris and Raviv [1993], Handa, Schwarz, and Tiwari [2003], and Hollifield et al. [2006] all suggest this line of thinking. The work by Handa, Schwarz, and Tiwari [2003] and Hollifield et al. [2006] is particularly germane as they both allow limit and market orders where the optimal order submission strategy is a function of investor valuations.

Handa et al.'s model assumes investors with differing valuations for the same security. Uninformed traders have the ability to submit either market or limit orders. They choose the type of order to submit, and conditional on a limit order, they choose the price to request. These choices are based on their desire to obtain the best possible price in the transaction and how this trades off against nonexecution risk. They also face adverse selection risk from trading against informed traders.

Handa et al.'s main theoretical result is that the optimal order submission strategy and the optimal price requested (conditional on a limit order) are

[6] A notable exception is that among higher priced stocks, forecast dispersion is positively correlated with my new measure of investor opinion divergence.

[7] The papers listed in footnote 1 may reach different conclusions if they use an alternative proxy to analysts' forecast dispersion as an indicator of investor opinion divergence.

[8] See references later.

directly related to investors' reservation prices. This is exactly what I seek to capture in my measure of investor opinion divergence. They also test their theory's implications for the behavior of the bid-ask spread, and find empirical support for it. The close link between order type, order price, and investors' reservation prices strongly supports my use of order data to proxy investors' varying opinions.

Hollifield et al.'s work estimates the gains from trade in limit order markets. In order to produce such estimates, they need to empirically link traders' order submissions to their valuations. Their empirical work is directly built upon their theoretical model. In the model, there are two important results with implications for my work. First, different investor valuations link with different limit order requested prices. This supports my use of limit order prices in the construction of my new measure of investor opinion divergence. Second, in their model, market orders and limit orders (under their optimal order submission strategies) imply different investor private valuations. Thus, it is important that I also include market orders in the calculation of my measure of investor opinion divergence, and that my treatment of them implies different valuations from limit orders.

Other empirical work supporting the use of order data for my exercise is found in Ahn, Bae, and Chan [2001]. Finally, Harris and Hasbrouck [1996] show that limit orders are a viable order strategy for individual investors compared to market orders, while Kaniel and Liu [2006] show that even informed traders appear to prefer them. Thus, access to limit order data seems important when evaluating traders' strategies.

2.2 OPINION DIVERGENCE MEASURE (DIVOP) CONSTRUCTION

2.2.1. Data. I begin with a (somewhat) random sample of NYSE-listed firms: those with ticker symbols beginning with the letters A to D as of January 2002. For these firms, I obtain all orders for every trading day during the January 2002–March 2002 window. The order data includes information on order submission date and time, the number of shares submitted, whether the order is a buy or a sell, the order type, account type (for example, index arbitrage or member trade) and in the case of limit orders, the limit order price. Because I am interested in measuring investor opinions, I do not include index arbitrages in the analysis.

The choice of time period (January–March 2002) deserves discussion. The adoption of RegFD in August 2000 suggests a regime shift in the way analysts form their forecasts (see e.g. Agrawal, Chadha, and Chen [2006]. Thus, data from the post-RegFD period is useful if I wish to be prescriptive and aid researchers in their choice of measurement of investor opinion divergence. On the other hand, market mechanisms changed in January 2002 (NYSE Open-book) to allow transparency of limit orders. In other words, limit orders may now reflect both opinions and preferences about how much to reveal regarding those opinions. Distinguishing the two seems difficult at best according to Bloomfield, O'Hara, and Saar [2005]. However, Handa, Schwarz, and Tiwari [2003] suggest otherwise, particularly if the rule for determining how much of the limit order book is displayed is mechanical.

Empirically, Boehmer, Saar, and Yu [2005] document differences in order submission strategy and spread behavior after the implementation of Openbook. Clearly there are trade-offs to using order data from the chosen time period. However, because a principal goal of this study is to guide future research using investor opinion divergence proxies, data drawn from a period that more closely resembles the current regime seems preferable.

2.2.2. Measuring Divergence of Investors' Opinions. My new measure for investor opinion divergence is the simple daily standard deviation (across orders) of the distance between each order's requested price (explicit for limit orders and implicit for market orders) and the most recent trade price preceding that order. If the most recent trade price was \$50 and a limit buy order for \$48 arrives, the limit order investor's opinion apparently diverges from the opinion embodied in the most recent trade. On the other hand, if a market order arrives, and the most recent trade price was \$50, I treat the order as if the implicit requested price equals \$50. In other words, a trader's willingness to accept the market price suggests that his opinion about asset value does not differ markedly from the opinion embodied in the most recent trade.[9] My specific calculations follow.

$$\%\text{Distance} = \frac{\text{OrderPrice} - \text{PriorTradePrice}}{\text{PriorTradePrice}}. \tag{1}$$

Opinion divergence (DIVOP) is then the standard deviation (across all orders on the day) of %Distance:

$$\text{DIVOP} = \left[\sum_{i=1}^{N} \frac{(\%\text{Distance}_i - \overline{\%\text{Distance}})^2}{N-1}\right]_{\text{day}}^{\frac{1}{2}}, \tag{2}$$

where N is the number of orders during the trading day. I require a minimum of 10 orders to calculate DIVOP. Finally, if %Distance is larger than 25%, suggesting an outlier, I ignore that order in my calculation of DIVOP.

2.3 DESCRIPTIVE STATISTICS

I offer two perspectives on the reasonableness of my construct for opinion divergence. The first is a firm-wide perspective. To assess whether the firms in my sample are "representative" of the general NYSE-listed firm universe,

[9] There are several important considerations here. First, the order must be reasonably close temporally to the most recent prior trade. I restrict my sample to those orders that arrive within one minute of the most recent prior trade. Second, a market order may not receive an execution price equal to the most recent prior trade price because of the bid-ask spread. However, on the NYSE, Petersen and Failkowski [1994] note that price improvement occurs with great frequency. Given a tendency for bid and ask prices to surround the most recent transaction price, price improvement will tend to move the transaction price on the current market order closer to the transaction price of the most recent prior trade. Third, my main conclusions are robust to focusing strictly on limit orders. Finally, my assumption is consistent with Hollifield et al.'s [2006] work that implies a difference in investors' private valuations when they submit market as opposed to limit orders.

TABLE 1

Descriptive Statistics

Panel A: Ex ante firm characteristics

Variable	N	Mean	Median	Std Dev	Min	Max
Total Assets	499	18,622	2,313.6[b]	77,081.5	0.6650	839,298
Market Value of Equity	555	6,422.8	1,158.7[c]	18,154.6	0.1865	207,665
Book Leverage	499	0.3113	0.2992	0.2192	0	1.6162
Tobin's Q	482	1.6603	1.2553	1.2505	0.4211	16.689
Net Income/Total Assets	499	0.0397	0.0285	0.1640	−1.3158	3.1059

Panel B: All NYSE-listed firms' ex ante firm characteristics

Variable	N	Mean	Median	Std Dev	Min	Max
Total Assets	2,180	16,077.5	1,841.1	68,107.1	0.1810	1,051,450
Market Value of Equity	2,433	6,220.8	873.3	21,188.3	0.1865	397,832
Book Leverage	2,173	0.3066	0.2934	0.2189	0	1.6707
Tobin's Q	2,024	1.7534	1.2558	2.3155	0.3660	59.3899
Net Income/Total Assets	2,177	0.0544	0.0306	0.5974	−8.0772	21.7891

Panel C: Descriptive statistics for DIVOP

Variable		N	Mean	Median	Std Dev	Min	Max
Full Sample	Percentile↓	13,017	0.0088	0.0072	0.0063	0.0008	0.0770
Percentile of "number	≤10	1,304	0.0127	0.0089	0.0111	0.0008	0.0770
of orders per day"	11–25	1,953	0.0096	0.0077	0.0069	0.0013	0.0612
distribution. Days with	25–75	6,504	0.0083	0.0072	0.0050	0.0014	0.0624
the fewest (most) orders	75–90	1,954	0.0080	0.0070	0.0051	0.0017	0.0639
are in the ≤10	≥90	1,302	0.0076	0.0068	0.0037	0.0021	0.0414
(≥90) percentile.							

Table presents descriptive statistics for the sample firms. The sample is all firms with ticker symbols beginning with the letters A–D, as of January 1, 2002. Panel A reports descriptive statistics that are firm-specific ex ante fiscal year-end values. Panel B reports the same, but for the entire universe of NYSE listed firms. The data's fiscal year-end must occur during calendar year 2001. Variables: total assets; market value of equity, equal to stock price times shares outstanding; book leverage, equal to the sum of long-term debt and debt in current liabilities, all divided by total assets; Tobin's Q, equal to total assets minus book equity plus market value of equity, all divided by total assets; net income divided by total assets. All "levels" are in $millions. [a,b,c] indicates significantly different medians between Panels A and B at the 10%, 5%, and 1%, respectively. Panel C presents descriptive statistics for the measure of investor opinion divergence (DIVOP).

I turn to Compustat. For each firm in my sample, I construct ex ante fiscal year-end measures of its total assets, market value of equity, book leverage, Tobin's Q, and net income to assets ratio. I do the same for the full universe of NYSE-listed firms. Univariate statistics on these five variables are presented in table 1. Panel A presents statistics for my sample, panel B for the universe.

My sample appears to be reasonably similar in terms of most basic firm characteristics. The median values of leverage and Q are within 1% of each other across the two panels, while net income scaled by assets only differs across the groups by 6.9%. These differences are not significant at conventional levels. Nor are the means of any "tabled" firm characteristics reliably different statistically. The lone area where my sample firms appear to be somewhat different from the NYSE universe is in firm size measures. In the median, my firms are larger in terms of assets (25.7% difference in medians) and in terms of market value of equity (32.7%). Both differences are significant at conventional levels (5% and 1%, respectively). Below, I control for firm size in my regressions.

My second perspective is at the firm/day level. I have a distribution of DIVOP measures, one for each firm/day in the sample. Panel C reports measures of central tendency for the full sample and for a set of five subsamples, categorized by how many orders were used in the calculation of the daily DIVOP value. The groups are: (minimum–10th percentile); (10th–25th percentile); (25th–75th percentile); (75th–90th percentile); (90th percentile–maximum).

Relatively more orders associates with a lower mean and median value of opinion divergence. Moreover, more orders also associates with a lower volatility (across observations) in DIVOP. Both of these results suggest care should be taken when studying the relationship between DIVOP and proxies. In particular, it appears that estimates of DIVOP are noisier as the number of orders used in the calculation shrinks. This seems especially the case for observations where the number of orders used to calculate DIVOP is low, that is in the 10th percentile of the distribution or lower. Below, my first tests are conducted for various cutoffs of minimum number of orders to reflect this.

2.4 AN EMPIRICAL ASSESSMENT OF DIVOP

I evaluate the validity of DIVOP in the context of theoretical work by Holthausen and Verrecchia [1990] and Kim and Verrecchia [1994]. An implication of their work is that opinions may diverge more as the information content of a news announcement gets larger. For the firms in my sample with earnings events in the January–March 2002 window, I measure the information content of each earnings announcement as the absolute value of the three-day abnormal return to the event (using standard market-model methodology). I then ask whether larger (absolute) three-day abnormal returns (i.e. greater information content to the earnings announcement) associate with larger values of DIVOP.

The results are as follows.[10,11] There is a significant positive relation between DIVOP and the information content of the news. The coefficient on the magnitude of the three-day abnormal announcement return (0.0159) is significantly positive with a *t*-statistic of 2.36. For robustness, I conduct a second test using the (in-sample) quintile ranking of the three-day abnormal announcement return. The coefficient on the ranked version of information content (0.0004) is also significant ($t = 1.87$). In the context of Holthausen and Verrecchia [1990] and Kim and Verrecchia [1994], DIVOP appears to capture variability in investor opinions, which is likely to be larger when the information content of an event is larger.

[10] They are not tabled (for brevity), but are available from the author upon request.

[11] I also include two control variables (described below). They are the inverse of the firm's stock price and a quintile ranking of the firm's market value of equity (based on where it falls in the NYSE's distribution).

3. *Proxies for Investor Opinion Divergence*

The following describes six proxies for investor opinion divergence common to the extant literature. Each proxy is discussed in terms of its construction, its potential shortcomings and the studies that employ it.

3.1 UNEXPLAINED VOLUME

3.1.1. Precedent. Numerous papers treat high trading volume as an indicator of divergent investor opinions. Bamber [1987] and Bamber, Barron, and Stober [1997, 1999] find that total trading volume is higher around earnings events that are more likely associated with more divergent investor opinions. Ajinkya, Atiase, and Gift [2004] document a positive correlation between trading activity and analysts' forecast dispersion. Kandel and Pearson [1995] find that earnings events that generate no price change, suggesting little reason to trade for information reasons, still cause abnormally large trading volume. They interpret this result as evidence that volume reflects diverging opinions about the value implications of earnings news. Fleming and Remolona [1999] find that trading volume surges while price volatility and spreads remain wide, as investors in Treasury securities trade to reconcile differential interpretations of macroeconomic information releases.

However, trading volume may proxy for more than just opinion divergence. For example, Benston and Hagerman [1974], Branch and Freed [1977], and Petersen and Fialkowski [1994] all use volume to proxy for liquidity. In other words, high volume may simply be due to the fact that a stock always exhibits large volume. Further complicating the interpretation of (total) volume, Tkac [1999] shows that individual stock volume is positively correlated with market volume. She argues that a combination firm-specific/market adjustment is the best way to isolate abnormal trading activity. I control for both firm-specific and market-wide trading activity in my calculations below.

There are a couple of reasons why volume (even the abnormal component of it) may not be a perfect proxy for investor opinion divergence. First, volume is measured on the basis of executed trades. In some cases, investor valuations may cause orders without executions. In these cases, volume is measured based on an attenuated sample of private valuations. Second, this attenuation bias is exacerbated by the use of transaction prices. If an investor's order is not executed, it is arguably because he was unwilling to accept the market price. Thus, execution prices probably do not accurately reflect all investors' private valuations.

3.1.2. Measuring Unexplained Volume. I calculate two measures of unexplained volume. The first is change in turnover and the second is standardized unexplained volume. All data for the calculations come from CRSP.

I calculate the change in turnover on a daily basis. My methodology follows Tkac [1999], Gebhardt, Lee, and Swaminathan [2001], Ajinkya, Atiase, and Gift [2004], and Garfinkel and Sokobin [2006]. My techniques mimic those in Garfinkel and Sokobin [2006]. I begin with a firm's daily turnover calculated as the firm's volume on that day divided by its shares outstanding (i.e. the percentage of outstanding shares traded on any particular day). To control for the correlation between firm-specific and market-wide trading (Tkac [1999]), I subtract market-wide turnover calculated the same way, but across all NYSE/AMEX stocks. The difference is daily market-adjusted turnover (labeled $\text{MATO}_{i,t}$) for firm i on day t.

Given a $\text{MATO}_{i,t}$ for firm i on day t, I recognize that stocks with relatively higher values of it may reasonably be the same stocks with relatively higher turnover overall (i.e. more liquid stocks). In other words, $\text{MATO}_{i,t}$ may also include liquidity trading.

I therefore subtract trading activity over a control period, from the above measure of market-adjusted turnover ($\text{MATO}_{i,t}$). Specifically, I subtract market-adjusted turnover (MATO_i) averaged over the month of December 2001. I label the resulting change in market adjusted turnover ΔTO.[12]

$$\Delta\text{TO} = \left\{\left[\left(\frac{\text{Vol}_{i,t}}{\text{Shs}_{i,t}}\right)_{\text{firm}} - \left(\frac{\text{Vol}_t}{\text{Shs}_t}\right)_{\text{mkt}}\right]\right\} - \sum_{\text{Dec2001}}\left[\left(\frac{\text{Vol}_{i,t}}{\text{Shs}_{i,t}}\right)_{\text{firm}} - \left(\frac{\text{Vol}_t}{\text{Shs}_t}\right)_{\text{mkt}}\right]\Big/ 20. \quad (3)$$

An alternative approach to recognizing firm-specific effects in volume is posited by Crabbe and Post [1994] and Garfinkel and Sokobin [2006]. They note that the arrival of new information about a stock can lead to more volume (as discussed in Holthausen and Verrecchia [1990]) via an "informedness effect."[13] In other words, new information changes investors' average valuation and encourages trade. Simply netting out control period volume to control for information's effect on volume (i.e. not DIVOP's effect), as in equation (3), assumes this effect is similar on the event day and during the control period. If price moves capture information, this is akin to assuming similar price moves during the "measurement day" and control window.

My alternative measure of unexpected volume is designed to control for both the liquidity effect and informedness effect in volume.[14] Similar to Crabbe and Post [1994] and specifically following Garfinkel and Sokobin [2006], I estimate unexplained volume using a methodology that mirrors the market model approach to estimating abnormal returns. Specifically, I construct a measure of "standardized unexplained volume" (SUV), calculated as a standardized prediction error from a regression of trading volume on the absolute value of returns for the tth firm i.

$$\text{SUV}_{i,t} = \frac{\text{UV}_{i,t}}{\text{S}_{\text{i,t}}}, \quad (4)$$

$$\text{UV}_{i,t} = \text{Volume}_{i,t} - E[\text{Volume}_{i,t}], \quad (5)$$

$$\text{E}[\text{Volume}_{i,t}] = \hat{\alpha}_i + \hat{\beta}_1 \cdot |\text{R}_{i,t}|^+ + \hat{\beta}_2 \cdot |\text{R}_{i,t}|^-. \quad (6)$$

---

[12] There are 20 trading days in December 2001.

[13] The evidence in Karpoff [1987] is broadly consistent with this.

[14] The residual is designed to capture opinion divergence.

The plus and minus superscripts on the absolute valued returns indicate when returns were positive or negative. This treatment is designed to recognize the observed empirical regularity that volume and absolute value of return are differentially sensitive to each other when returns are positive versus negative (e.g. Karpoff [1987]). Finally, $S_{i,t}$ is the standard deviation of the residuals from the regression, calculated over the model's estimation period (November/December 2001).[15]

3.2 BID-ASK SPREAD

3.2.1. Precedent. Bid-ask spread has long been considered a proxy for the costs of asymmetric information (AI), (see Bagehot [1971]). The general intuition is that market makers do not know when they are trading against informed individuals, whose opinions about firm value are likely to be closer to "truth" than theirs. In response, the market maker protects herself by charging a spread. The underpinning that different traders have different opinions and that this engenders a spread motivates the use of spread as a proxy in the extant literature. Houge et al. [2001] apply this thinking around IPOs. Handa, Schwarz, and Tiwari [2003] model a link between spread and opinion divergence and document evidence consistent with it.

However, spreads do not reflect solely this AI cost. They also compensate market makers for order processing (OP) and inventory costs. George, Kaul, and Nimalendran [1991] empirically separate these components. They show that OP costs represent the predominant component of bid-ask spreads, AI is the next largest component, and inventory holding costs are essentially zero. In my robustness checks, I present results that link investor opinion divergence with the OP and AI components of bid-ask spread.

Closely related work by Easley, Hvidkjaer, and O'Hara [2002] measures the probability of information-based trading by investors (PIN). They show that it carries significant explanatory power for returns. They interpret PIN as a measure of information risk, which is similar to the "AI" component of spreads. Thus, there is an important link between the probability of informed trading (PIN) and bid-ask spreads. In my robustness checks, I investigate whether PIN is related to diverging investor opinions, and also whether conditioning on the level of PIN influences the relationship between DIVOP and typical proxies.

Bid-ask spreads may also suffer from an attenuation bias. While market makers ostensibly see all submitted orders, their duty is to ensure the orderly operation of the market by standing ready to buy or sell on demand. Because limit orders do not require immediate execution, bid-ask spreads may reflect market order requests, rather more than limit order requests. If limit order likelihood is a function of opinion divergence, then spreads may not fully reflect such divergence in opinions.

[15] I use two months of data for the estimation of parameters because there are only 20 trading days in December 2001.

3.2.2. Measuring Bid-Ask Spread. Data for calculating my bid-ask spread measure comes from NYSE's TAQ data. I obtain individual quotes from trading hours on an intra-day basis. This allows me to calculate individual daily measures of bid-ask spread for any particular firm. If the data are available, each firm has 60 separate daily values of average daily bid-ask spread.

My spread-based proxy for opinion divergence is the daily percentage bid-ask spread.

$$\%\text{Spread} = \left[\frac{\displaystyle\sum_{i=1}^{N} \frac{(\text{Ask}_i - \text{Bid}_i)}{(\text{Ask}_i + \text{Bid}_i)/2}}{N} \right]_{\text{day}}, \tag{7}$$

where i indexes the quotes. Thus, %Spread is simply bid-ask spread scaled by the mid-point of the two quotes that define the spread, averaged across all quotes during the day.

3.3 STOCK RETURN VOLATILITY

3.3.1. Precedent. Theoretical work links divergent investor opinions and stock return volatility. Shalen [1993] provides a model in which the two are positively correlated. She also shows that volume is positively related to divergent opinions and therefore volume and volatility should be positively related. Similar implications emerge from the model by Wang [1998]. Shalen's [1993] model also suggests conditions when stock return volatility may be a stronger proxy for opinion divergence—when there is greater risk-weighted hedging demand. One of my conditioning tests below attempts to exploit such thinking.

Supportive empirical evidence is found in Frankel and Froot [1990]. Using foreign exchange survey data, they find a positive relation between divergent opinions and price volatility. They also find that divergent opinions and volume are positively related. The positive influence of divergent opinions on both volume and volatility is consistent with both Shalen [1993] and Wang [1998]. Other consistent empirical evidence is found in Daigler and Wiley [1999], Boehme, Danielsen, and Sorescu [2006], and Chen and Cheng [2003].

There are at least two potential factors that may mitigate a possible relation between stock return volatility and divergent opinions. Stock return volatility may proxy for risk (Goyal and Santa-Clara [2003]), and this would seem to be a first order consideration. Moreover, stock return volatility is (like volume) measured on the basis of executed trades. As noted earlier, this may create an attenuation bias.

3.3.2. Measuring Stock Return Volatility. I obtain data to construct an intra-day measure of return volatility from TAQ. Again, if sufficient data are available, each firm has 60 separate daily values of daily stock return volatility. I calculate a simple measure of daily stock return volatility using all transactions during normal trading hours:

$$\sigma_{returns} = \left[\sum_{k=1}^{K} \frac{(\text{Ret}_k - \overline{\text{Ret}})^2}{K-1} \right]_{\text{day}}^{\frac{1}{2}}, \tag{8}$$

where k indexes the transactions during the day. The returns used in (8) are transaction to transaction returns.

3.4 ANALYSTS' FORECAST DISPERSION

3.4.1. Precedent. Numerous papers use dispersion in analysts' forecasts as a proxy for investor opinion divergence.[16] The presumption is that analysts express their unbiased opinion in their earnings forecasts, and that investors' opinions follow analysts'. If forecasts vary, and they are unbiased analyst opinions, then forecast dispersion is a reasonable proxy for analyst opinion divergence. If investors' opinions follow analysts', then forecast dispersion translates into divergent investor opinions.

Brown, Foster, and Noreen [1985], McNichols and O'Brien [1997], Lin and McNichols [1998], Dechow, Hutton, and Sloan [2000], and others, cast doubt on the employ of analysts' forecast dispersion to proxy opinion divergence by noting problems with analysts' forecasts.[17] Briefly, there are concerns with the following. IBES data contains stale forecasts, IBES data may be biased, forecast distributions may be truncated and there is a potential bias in the favorableness of the coverage.[18]

The stale forecast concern with using analysts' forecast dispersion to proxy divergent investor opinions (noted by McNichols and O'Brien [1997]) is as follows. Analysts may simply choose to drop coverage in the face of news that implies a reduction in expected future earnings. Thus, a "stale" forecast is no longer an accurate representation of the analyst's view. More generally, any news that arrives between the issuance of the "stale" forecast and the current period can change analysts' views and including the old forecast based on the old information can lead to a measure of forecast dispersion that is the *combination* of opinion divergence and an information effect.[19]

[16] See footnote 1.

[17] See Healy and Palepu [2001] for a nice review of analysts' information intermediation role in capital markets.

[18] I discuss the stale forecast problem in more detail later. The IBES data bias problem is multi-faceted. Ljungqvist et al. [2009] document evidence of backfilling by IBES, though their analysis is restricted to recommendations, rather than forecasts. Diether, Malloy, and Scherbina [2002] highlight the influence of rounding errors on IBES forecast dispersion measures. Forecast distributions may be truncated because analysts do not continuously update forecasts even as the information set changes, nor do they always *immediately* report changed forecasts to IBES. Finally, there is potential upward bias in analysts' expressed opinions because their wealth may actually benefit from this, through generated investment banking business and higher ex post compensation.

[19] See also Guttman [2005].

Since I only seek to measure the former, I must be careful to avoid "overly" stale forecasts in my analyst variability proxy (see below).[20]

3.4.2. Measuring Analysts' Forecast Dispersion. I obtain analyst forecast data from IBES. I use the detail tape to collect individual analysts' forecasts. My main tests use forecast dispersion proxies constructed from forecasts submitted during the month of my opinion divergence measure's calculation. I do so in recognition of the concerns raised by McNichols and O'Brien [1997] as well as Guttman [2005].

$$\sigma_{\mathrm{frcst}} = \left\{ \left[\sum_{k=1}^{K} \frac{(\mathrm{Frcst}_k - \overline{\mathrm{Frcst}})^2}{K-1} \right]^{\frac{1}{2}}_{\mathrm{month}} \right\} \Big/ |\overline{\mathrm{Frcst}}|, \tag{9}$$

where Frcst_k is the kth analyst's forecast of annual earnings per share and $|\overline{\mathrm{Frcst}}|$ is the absolute value of the mean analyst's forecast.

One potential concern with forecast dispersion calculated as in (9) is that mean forecasts near zero generate very large σ_{frcst} measures. To address this concern, I also calculate a second measure of forecast dispersion by scaling the standard deviation of forecasts by the firm's stock price (averaged over the month). This alternative is labeled $\sigma_{\mathrm{frcst\_2}}$.

4. *Results*

4.1 MAIN RESULTS ON OPINION DIVERGENCE PROXY EFFECTIVENESS

I regress my new measure of investor opinion divergence (DIVOP) on each proxy and two control variables to ascertain whether the proxy carries explanatory power for variation in DIVOP. My two control variables are:

MVEQ Quintile: Ranking from 1 (smallest) to 5 (largest) of the firm's size. Rankings are based on which quintile of market value of equity the firm belongs to based on December 2001 market value of equity. Quintile cutoffs are based on NYSE-listed firm market values of equity.

P_inverse: The reciprocal of the firm's stock price at the close of the trading day. By construction (see equation (1)), DIVOP should be inversely correlated with stock price.

Table 2, panel A presents OLS regression results illustrating the link between opinion divergence and the standard proxies. The dependent variable in each regression is my new measure for opinion divergence based on limit and market orders (DIVOP). Each model is a separate regression of DIVOP on one of the common proxies and the two (above) controls. The sample varies with criteria on the minimum number of orders I require to calculate DIVOP. I use cutoffs of at least 10, 20, 30, 50, and 100 orders because of the documented noise in DIVOP estimates at low numbers of orders

[20] It's noteworthy that despite my attempts to control for the stale forecasts concern, forecast dispersion generally appears to be a weak proxy for investor opinion divergence.

TABLE 2

The Relation between Opinion Divergence and Proxies

Panel A: Categorization by minimum number of orders required to calculate DIVOP

Number of Orders	Proxy Variable: ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$	N
≥10	0.1043[c]	0.0005[c]	0.0536[c]	*−0.0039*	*−0.000002*	*−0.0001*[c]	13,017
	(0.2815)	(0.2757)	(0.2704)	*(0.2679)*	*(0.2677)*	*(0.2683)*	
≥20	0.1051[c]	0.0005[c]	0.0486[c]	*−0.0056*[b]	*−0.000003*	*−0.0001*[c]	12,903
	(0.2948)	(0.2888)	(0.2822)	*(0.2801)*	*(0.2798)*	*(0.2804)*	
≥30	0.1052[c]	0.0005[c]	0.0500[c]	*−0.0038*	*−0.000003*	*−0.0001*[c]	12,816
	(0.2911)	(0.2846)	(0.2778)	*(0.2753)*	*(0.2752)*	*(0.2757)*	
≥50	0.1044[c]	0.0005[c]	0.0474[c]	*−0.0019*	*−0.000003*	*−0.0001*[c]	12,643
	(0.2907)	(0.2848)	(0.2764)	*(0.2740)*	*(0.2739)*	*(0.2745)*	
≥100	0.1027[c]	0.0005	0.0435[c]	*−0.0008*	*−0.000001*	*−0.0001*[c]	12,273
	(0.2843)	(0.2785)	(0.2689)	*(0.2666)*	*(0.2666)*	*(0.2673)*	

Panel B: Differences in relative explanatory power (Vuong 1989)

Rankings of explanatory power of proxies based on adjusted R^2:

Number Orders	(1)	>	(2)	>	(3)	>	(4)	>	(5)	>	(6)
≥10	ΔTO[y]		SUV[z]		Spread		σ_{frcsts_2}		$\sigma_{returns}$		σ_{frcsts}
≥20	ΔTO[y]		SUV[z]		Spread		σ_{frcsts_2}		$\sigma_{returns}$		σ_{frcsts}
≥30	ΔTO[y]		SUV[z]		Spread		σ_{frcsts_2}		$\sigma_{returns}$		σ_{frcsts}
≥50	ΔTO[y]		SUV[z]		Spread		σ_{frcsts_2}		$\sigma_{returns}$		σ_{frcsts}
≥100	ΔTO[x]		SUV[z]		Spread		σ_{frcsts_2}		$\sigma_{returns}$		σ_{frcsts}

Panel A of the table presents regressions of opinion divergence on proxies for it and control variables for firm size and stock price (coefficients on controls are not reported). There are six regressions in each row. Each row's sample varies with the condition on number of orders required to calculate opinion divergence. Control variables: Market value of equity quintile is assigned based on which quintile of NYSE firms the firm's market value of equity falls in (at the end of 2001). P_inverse equals one over stock price, and is specific to the firm/day. Proxies: ΔTO is calculated as market-adjusted turnover on the day, minus the average of market-adjusted turnover during a control period (December 2001). SUV is the scaled (by estimation window (November/December 2001) standard deviation of prediction errors) one-day prediction error from a market model-style regression of volume on absolute valued returns. Spread is the average bid-ask spread across all quotes during normal trading hours on the day. σ_{returns} is the standard deviation of transaction-to-transaction returns, across all transactions during normal trading hours on the day. σ_{frcsts} is the standard deviation of analysts' forecasts, divided by the average forecast. Forecasts used are submitted during the month corresponding to the day of analysis. If the dependent variable is from January 30, 2002, then forecasts issued during the month of January 2002 are used. $\sigma_{\text{frcsts\_2}}$ is the standard deviation of analysts' forecasts, divided by the firm's average stock price during the month. [a,b,c] indicates significance of coefficient at the 10%, 5%, and 1% levels, respectively. *Adjusted R^2s for the full regression (including the control variables) are reported below each coefficient in parentheses*. Panel B presents the rank ordering of explanatory power of the proxies for DIVOP (based on adjusted R^2). Tests of differences in explanatory power are between the variable and its adjacent variable to the right (ΔTO is compared to SUV, SUV is compared to spread, etc.), and are based on Vuong [1989]. *Proxies carrying negative coefficients are italicized and treated as having lower explanatory power for DIVOP (they appear to the right of proxies with positive coefficients in Vuong tests).*

[x,y,z] indicates significant Vuong test values at the 10%, 5%, and 1% levels, respectively.

(in table 1). I report only the coefficients on the proxy (for readability reasons) and the full model's adjusted R^2 in parentheses below the coefficient.[21]

Across all five rows of results (which vary by the criteria for minimum number of orders to calculate a daily DIVOP measure), we see a consistent set of results. DIVOP is significantly positively correlated with both unexplained

[21] Indications of significance are provided (as a superscript) to the right of the proxy's coefficient.

volume proxies and with bid-ask spread. It is not significantly positively correlated with the other three proxies common to the extant literature on investor opinion divergence.

However, these results say nothing about a dominant proxy—whether one proxy is superior to any others (i.e. they do not assess relative explanatory power). To address this question, I employ the methodology of Vuong [1989]. Specifically, I conduct a likelihood ratio test that evaluates competing non-nested models in terms of their explanatory power. The Vuong [1989] test specifies a null hypothesis that does not presume either proxy is reliable. Therefore, the test statistic (a Z-statistic) allows for a directional conclusion regarding which of the competing proxies is more correlated with DIVOP. The Vuong [1989] test improves upon competing tests of non-nested model selection because it can reject one proxy in favor of another, even if both are significantly positively correlated with DIVOP (even *incrementally*).

The results from Vuong [1989] tests are presented in panel B. Given my goal of selecting a dominant proxy, I simply compare one proxy's explanatory power with the *next best* proxy's explanatory power. I continue this through all six models, comparing the second best proxy with the third best, the third best proxy with the fourth best, etc. I order the proxies' explanatory power based on adjusted R^2 and whether the coefficient on the proxy is positive (because high explanatory power in a regression with a proxy that is *negatively* related to DIVOP is inconsistent with the extant literature's treatment of the proxy).

My relative explanatory power results are quite similar across the five different criteria for minimum number of orders in the DIVOP calculation. The unexplained volume measure "change in turnover" (ΔTO) is always the best proxy. Vuong [1989] tests always indicate that it dominates the next best proxy, SUV. The fact that SUV is the next best proxy, and that it always dominates the third best proxy (Spread) suggests that for the average situation, unexplained volume is the best possible proxy for investor opinion divergence.

Notably, the other three proxies common to the extant literature are all negatively correlated with DIVOP. This suggests that they are not good proxies. In particular, the negative correlations between DIVOP and the analysts' forecast dispersion measures calls into question a broad swath of conclusions that presume forecast dispersion proxies for investor opinion divergence.[22] Overall, the evidence in table 2 strongly suggests that unexplained volume is the best proxy for investor opinion divergence in an average setting. Future research might justifiably rely on it as such.

4.2 CONDITIONAL ANALYSIS

The earlier results suggest that unexplained volume is the best proxy for investor opinion divergence, but these are average results across the

[22] Again, see footnote 1 for details.

main sample of firm/days. They do not address whether conditioning on either firm or order characteristics affects my inferences about which proxy for opinion divergence is best. This section conditions on several potential characteristics that may associate with different proxies dominating as explainers of DIVOP.

4.2.1. Conditional on IBES Following. Hong, Lim, and Stein [2000] provide evidence consistent with IBES analyst following changing the information environment for a stock. Specifically, information is incorporated into prices faster when there is analyst following. This may influence order submission strategies, as well as market maker behavior, with the upshot that correlations between proxies and opinion divergence vary. This could change inferences regarding which proxy is dominant.

Table 3, panel A explores the relation between DIVOP and the usual proxies for two mutually exclusive subsamples: firms with IBES measures of analysts' forecast dispersion and firms without. Consistent with the results in table 2, when firms have data on forecast dispersion ΔTO is the dominant proxy and SUV is second best, while Spread is third best. By contrast, for the sample without forecast dispersion data, ΔTO is no longer positively related to DIVOP. SUV still is, so unexplained volume may yet be a good proxy in this subsample. However, Spread is too, as is stock return volatility. Panel A1 (containing Vuong [1989] tests) indicates that the spread proxy is the best possible proxy for opinion divergence in this subsample, followed by stock return volatility and only then, SUV. The contrasting results for the analyst-followed and nonfollowed firm samples are consistent with Hong, Lim and Stein [2000]. I further investigate this result below (in section 4.3).

4.2.2. Conditional on Firm Size. A host of papers in accounting and finance find differential stock performance and firm behavior across different firm sizes. Panel B of table 3 conditions on firm size, by market value of equity quintile.[23] I find that unexplained volume is significantly positively correlated with DIVOP, regardless of which firm size quintile I examine.

However, the altered sample construction appears to influence my ability to distinguish significant explanatory power differences between the two unexplained volume proxies. Vuong [1989] tests do not indicate a dominant proxy among the two. To further explore this issue, I re-run the panel B tests conditioning on firm size *terciles*. For small and large firms (terciles one and three), ΔTO appears to dominate other proxies for opinion divergence. It is only for medium firms (tercile two) that there is no significant difference.

It is also worth noting that since these tests are conducted on the main (i.e. analyst followed) sample, cut into firm size quintiles, there are far fewer small firms (size quintiles one and two) than larger firms. To be sure that sample size differences are not influencing my conclusions, I re-run my

[23] To avoid perfect collinearity, I adjust my firm size control variable so that it's a size decile ranking rather than size quintile ranking.

TABLE 3

The Relation between Opinion Divergence and Proxies Categorized by IBES Availability, Firm Size, Inverse of Stock Price, and Volume

Panel A: Categorization by whether firm has IBES forecast dispersion data

σ_{frcsts} Data Available?	Proxy Variable ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$	N
NO	−0.0047	0.0002[c]	0.26715[c]	0.0073[c]	N/A	N/A	17,718
	(0.4312)	(0.4316)	(0.4595)	(0.4319)			
YES	0.1043[c]	0.0005[c]	0.0536[c]	*−0.0039*	*−0.000002*	*−0.0001*[c]	13,017
	(0.2815)	(0.2757)	(0.2704)	*(0.2679)*	*(0.2677)*	*(0.2683)*	

Panel A1: Differences in relative explanatory power (Vuong 1989)

Rankings of explanatory power of proxies based on adjusted R^2:

IBES	(1)	>	(2)	>	(3)	>	(4)	>	(5)	>	(6)
NO	Spread[z]		σ_{returns}		SUV		ΔTO				
YES	ΔTO[y]		SUV[z]		Spread		*σ_{frcsts_2}*		*$\sigma_{returns}$*		*σ_{frcsts}*

Panel B: Categorization by firm size (market value of equity) quintile

Firm Size	Proxy Variable ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$	N
Quintile 1	0.4552[c]	0.0016[c]	*−0.0234*	*−0.0120*	*−0.0005*[a]	*−0.0007*	658
(smallest)	(0.2911)	(0.2841)	*(0.2632)*	*(0.2654)*	*(0.2665)*	*(0.2643)*	
Quintile 2	0.2065[c]	0.0006[b]	0.0048	0.0066	*−0.0008*[b]	*−0.0001*[b]	961
	(0.1537)	(0.1210)	(0.1150)	(0.1156)	*(0.1197)*	*(0.1186)*	
Quintile 3	0.0607[c]	0.0004[c]	*−0.0153*	*−0.0025*	0.00003	0.0003	2,398
	(0.4214)	(0.4213)	*(0.4158)*	*(0.4157)*	(0.4158)	(0.4157)	
Quintile 4	0.1041[c]	0.0006[c]	0.0882[c]	*−0.0053*	*−0.0001*[c]	*−0.0022*[b]	3,353
	(0.2020)	(0.1890)	(0.1759)	*(0.1650)*	*(0.1668)*	*(0.1664)*	
Quintile 5	0.0911[c]	0.0004[c]	0.0878[c]	*−0.0007*	0.00001	0.0007	5,647
(largest)	(0.2730)	(0.2710)	(0.2749)	*(0.2659)*	(0.2660)	(0.2660)	

Panel B1: Differences in relative explanatory power (Vuong 1989)

Rankings of explanatory power of proxies based on adjusted R^2:

Size	(1)	>	(2)	>	(3)	>	(4)	>	(5)	>	(6)
1	ΔTO		SUV		*σ_{frcsts}*		*$\sigma_{returns}$*		*σ_{frcsts_2}*		*Spread*
2	ΔTO		SUV		σ_{returns}		Spread		*σ_{frcsts}*		*σ_{frcsts_2}*
3	ΔTO		SUV[y]		σ_{frcsts}		$\sigma_{\text{frcsts\_2}}$		*Spread*		*$\sigma_{returns}$*
4	ΔTO		SUV[y]		Spread		*σ_{frcsts}*		*σ_{frcsts_2}*		*$\sigma_{returns}$*
5	Spread		ΔTO		SUV[z]		σ_{frcsts}		$\sigma_{\text{frcsts\_2}}$		*$\sigma_{returns}$*

Panel C: Categorization by firm's control period (December 2001) volume quintile

Market-Adjusted Turnover	Proxy Variable ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$	N
Qintile 1	0.2894[c]	0.0007[c]	0.0464[b]	*−0.0067*[a]	0.0002	*−0.0024*[c]	2,584
(lowest turnover)	(0.2148)	(0.2105)	(0.2009)	*(0.1998)*	(0.1994)	*(0.2028)*	
Quintile 2	0.1227[c]	0.0005[c]	0.0613[c]	*−0.0029*	*−0.0002*	0.0011[b]	2,605
	(0.2534)	(0.2522)	(0.2460)	*(0.2426)*	*(0.2430)*	(0.2442)	
Quintile 3	0.1609[c]	0.0007[c]	0.0620[c]	*−0.0053*	0.00002	*−0.0025*[c]	2,628
	(0.3498)	(0.3512)	(0.3411)	*(0.3379)*	(0.3378)	*(0.3406)*	
Quintile 4	0.1833[c]	0.0005[c]	0.0156	0.0084	*−0.00003*	*−0.0001*[c]	2,594
	(0.2477)	(0.2106)	(0.2028)	(0.2030)	*(0.2030)*	*(0.2049)*	
Quintile 5	0.0624[c]	0.0006[c]	0.0550[c]	0.0072	0.0004[c]	0.0014	2,606
(highest turnover)	(0.4296)	(0.4234)	(0.4174)	(0.4153)	(0.4171)	(0.4154)	

(*Continued*)

TABLE 3 — *Continued*

Panel C1: Differences in relative explanatory power (Vuong 1989)

Rankings of explanatory power of proxies based on adjusted R^2:

Volume	(1)	>	(2)	>	(3)	>	(4)	>	(5)	>	(6)
1	ΔTO		SUV^{y}		Spread		σ_{frcsts}		*σ_{frcsts_2}*		*$\sigma_{returns}$*
2	ΔTO		SUV		Spread		σ_{frcsts_2}		*σ_{frcsts}*		*$\sigma_{returns}$*
3	SUV		ΔTO^{x}		Spread		σ_{frcsts}		*σ_{frcsts_2}*		*$\sigma_{returns}$*
4	ΔTO^{y}		SUV^{y}		$\sigma_{returns}$		Spread		*σ_{frcsts_2}*		*σ_{frcsts}*
5	ΔTO		SUV		Spread		σ_{frcsts}		σ_{frcsts_2}		$\sigma_{returns}$

Panel D: Categorization by firm's inverse stock price (1/P) quintile

	Proxy Variable						
(1/Stock Price)	ΔTO	SUV	Spread	$\sigma_{returns}$	σ_{frcsts}	σ_{frcsts_2}	*N*
Quintile 1	0.1093^{c}	0.0004^{c}	0.0244^{b}	0.0008	0.0004^{c}	0.0024^{c}	2,602
(highest price)	(0.0434)	(0.0225)	(0.0034)	(0.0010)	(0.0061)	(0.0048)	
Quintile 2	0.1043^{c}	0.0005^{c}	0.0741^{c}	*−0.0021*	0.0004^{c}	0.0015	2,605
	(0.0258)	(0.0220)	(0.0162)	*(0.0104)*	(0.0133)	(0.0104)	
Quintile 3	0.0709^{c}	0.0005^{c}	0.0439^{c}	0.0023	0.00004^{b}	*−0.0002*	2,602
	(0.0447)	(0.0465)	(0.0363)	(0.0338)	(0.0352)	*(0.0338)*	
Quintile 4	0.1136^{c}	0.0008^{c}	0.0826^{c}	*−0.0062*	0.00005^{a}	*−0.0004*	2,605
	(0.0468)	(0.0348)	(0.0153)	*(0.0038)*	(0.0045)	*(0.0035)*	
Quintile 5	0.1238^{c}	0.0004^{c}	0.0596^{c}	*−0.0042*	*−0.0002*^{c}	*−0.0001*^{a}	2,603
(lowest price)	(0.2510)	(0.2441)	(0.2434)	*(0.2416)*	*(0.2442)*	*(0.2424)*	

Panel D1: Differences in relative explanatory power (Vuong 1989)

Rankings of explanatory power of proxies based on adjusted R^2:

[1/P]	(1)	>	(2)	>	(3)	>	(4)	>	(5)	>	(6)
1	ΔTO^{z}		SUV^{y}		σ_{frcsts}		σ_{frcsts_2}		Spread		$\sigma_{returns}$
2	ΔTO		SUV		Spread^{x}		σ_{frcsts}		σ_{frcsts_2}		*$\sigma_{returns}$*
3	SUV		ΔTO^{x}		Spread		$\sigma_{frcsts}{}^{x}$		$\sigma_{returns}$		*σ_{frcsts_2}*
4	ΔTO		SUV^{y}		Spread^{y}		σ_{frcsts}		*$\sigma_{returns}$*		*$\sigma_{returns}$*
5	ΔTO		SUV		Spread		*σ_{frcsts}*		*σ_{frcsts_2}*		*$\sigma_{returns}$*

Table presents regressions of opinion divergence on proxies for it and control variables for firm size and stock price (coefficients on controls are not reported). There are six regressions in each row. Each row's sample varies with the condition listed (categorization criteria). Categorization is (in panel A) IBES availability (sufficient data to calculated forecast dispersion); (in panel B) firm size (market value of equity) quintiles; (in panel C) 1/(Stock Price), in quintiles (where stock price is the closing price on the day); and (in panel D) ex ante volume (market-adjusted volume in December 2001) in quintiles. Control variables: Market value of equity quintile is assigned based on which quintile of NYSE firms the firm's market value of equity falls in (at the end of 2001). To avoid perfect collinearity problems, the control variable for firm size is market value of equity decile in the firm size categorized regressions. P_inverse equals one over stock price, and is specific to the firm/day. Proxies: ΔTO is calculated as market-adjusted turnover on the day, minus the average of market-adjusted turnover during a control period (December 2001). SUV is the scaled (by estimation window (November/December 2001) standard deviation of prediction errors) one-day prediction error from a market model-style regression of volume on absolute valued returns. Spread is the average bid-ask spread across all quotes during normal trading hours on the day. $\sigma_{returns}$ is the standard deviation of transaction-to-transaction returns, across all transactions during normal trading hours on the day. σ_{frcsts} is the standard deviation of analysts' forecasts, divided by the average forecast. Forecasts used are submitted during the month corresponding to the day of analysis. If the dependent variable is from January 30, 2002, then forecasts issued during the month of January 2002 are used. σ_{frcsts_2} is the standard deviation of analysts' forecasts, divided by the firm's average stock price during the month. ^{a,b,c} indicates significance of coefficient at the 10%, 5%, and 1% levels, respectively. *Adjusted R^2s for the full regression (including the control variables) are reported below each coefficient in parentheses.* Panels A1, B1, C1, and D1 present the rank orderings of explanatory power of the proxies for DIVOP (based on adjusted R^2). Tests of differences in explanatory power are between the variable and its adjacent variable to the right (for the IBES—YES sample, ΔTO is compared to SUV, SUV is compared to spread, etc.), and are based on Vuong [1989]. *Proxies carrying negative coefficients are italicized and treated as having lower explanatory power for DIVOP (they appear to the right of proxies with positive coefficients in Vuong tests).*

^{x,y,z} indicates significant Vuong test values at the 10%, 5%, and 1% levels, respectively.

regressions with bootstrapped coefficients and standard errors, re-sampling 10,000 times. This does not influence my conclusions.

Finally, I find that spread is equally as effective as unexplained volume in explaining DIVOP among the largest firms (size quintile 5).[24] This appears to be due to the reduced correlation between unexplained volume and DIVOP for this subsample. In particular, while the coefficient on spread is quite similar in regressions for firm size quintiles 4 and 5 (0.0882 vs. 0.0878, respectively), it's lower for both ΔTO (0.1041 vs. 0.0911, respectively) and SUV (0.0006 vs. 0.0004, respectively).[25]

Why is the correlation between unexplained volume and DIVOP lower among the largest firms? One possible explanation is in the patterns of the two metrics across size quintiles 4 and 5. Average DIVOP is nearly unchanged (0.0082 vs. 0.0081, respectively). By contrast, unexplained volume is much lower in quintile 5. For example, ΔTO is roughly four times larger in quintile 4 versus quintile 5 (0.00135 vs. 0.00033).[26] Moreover, ΔTO's cross-sample standard deviation is roughly 43% smaller in size quintile 5. Thus, there is similar DIVOP but much less unexplained volume and less variability in it to correlate with DIVOP.

Overall, the firm size conditioning results suggest that the superior ability of unexplained volume to capture DIVOP is nonlinear in size.[27] Among nonfollowed firms (which tend to be much smaller), the explanatory power of unexplained volume is weaker than spread's. Among followed firms of most sizes (quintiles 1 to 4) unexplained volume carries the strongest explanatory power for DIVOP. However, among the largest firms both spread and unexplained volume carry similar explanatory power for DIVOP.

4.2.3. Conditional on Firm Volume. Stocks with low trading activity may have less measurement error in DIVOP. One way I have already attempted to control for this is with the number of orders criterion cutoffs in the table 2 analysis. Another way I (now) control for it is by conditioning on the firm's typical volume, measured as market-adjusted turnover during December 2001 (in quintiles).

Panel C presents the results. Again, unexplained volume and spread are significantly positively related to DIVOP. Vuong [1989] tests in panel C1 suggest that unexplained volume dominates spread, but it is difficult to establish a dominant proxy from among the two unexplained volume choices.

4.2.4. Conditional on Firm Stock Price. My main tests control for stock price (inverse) because it has a direct influence on the calculated value of DIVOP.

---

[24] Size quintile 2 indicates no significant difference in explanatory power for DIVOP between SUV and spread. However, ΔTO is significantly better than spread in the Vuong test (p-value = .08).

[25] Note the adjusted R^2 are higher in size quintile 5 regressions. This is due to the tigher fit of the overall model, particularly the explanatory power of the inverse stock price control.

[26] This is likely because typical volume is increasing in firm size (Chordia, Huh, and Subrahmanyam [2007]).

[27] I thank the referee for highlighting this.

However, it is also possible that this control may be a useful conditioning variable for the regression analysis.

Panel D of table 3 presents conditional analysis with categorization by P_inverse (1/Stock Price). Again, unexplained volume appears to dominate other proxies for opinion divergence.[28] However, Vuong [1989] tests generally fail to indicate a significant difference between the explanatory power of ΔTO and SUV. This may be due to sample size. For example, the coefficients on ΔTO are similar across all five P_inverse quintiles, and similar to the full sample coefficient in table 2. But, the explanatory power is lower (adjusted R^2) perhaps because of the smaller sample.

On the other hand, the relation between DIVOP and analyst forecast dispersion (scaled by stock price) does appear to hinge on inverse stock price quintile. There is a marked U-shape to the coefficients across the quintiles. I further examine this relationship as follows. I construct an interactive variable ($\sigma_{\text{frcst\_2}}$ * P_inverse) and include it in the regression of DIVOP on $\sigma_{\text{frcst\_2}}$ (and controls). The coefficient on $\sigma_{\text{frcst\_2}}$ by itself is now significantly positive and the coefficient on the interactive variable is negative. Higher price stocks (with lower P_inverse) associate with a more positive relation between DIVOP and $\sigma_{\text{frcst\_2}}$. In other words, there is a positive relation between DIVOP and $\sigma_{\text{frcst\_2}}$ among higher priced stocks.

Why might the above result occur? First, $\sigma_{\text{frcst\_2}}$ is lower among high priced stocks by construction. If high price stocks attract fewer uninformed (small wealth constrained) investors, this may also lower DIVOP because informed investors seem more likely to cluster their bids and offers around an informed value. To test this intuition, I correlate DIVOP with the fraction of orders that come from individual (rather than institutional) investors and I also correlate stock price with the same. I expect negative relations in both cases and this is what I find. Thus, higher stock prices associate with both reduced DIVOP (through the lower proportion of individual investor trades) and lower $\sigma_{\text{frcst\_2}}$ by construction. This is consistent with the positive relation between DIVOP and $\sigma_{\text{frcst\_2}}$ among high priced stocks.

4.2.5. Conditional on Order Type. Shalen [1993] hypothesizes that risk-weighted hedging demand may influence the tripartite relationship between volume, stock return volatility, and opinion divergence (which she calls dispersion). I investigate this possibility by conditioning on program trading orders in my data. While this is not a perfect analogy, program trades seem more likely to be used in conjunction with portfolio optimization strategies, suggesting portfolio risk-weighting motivated trade.

Table 4 panels A and B present the analysis. I condition on program trading in two ways. First, I construct DIVOP using orders drawn from two distinct samples: program trade orders and nonprogram-trade orders. I then run my tests separately using the two different DIVOP series. Second, I construct

[28] In all cases, the unexplained volume proxies have the highest adjusted R^2. In three of those five cases, they are significantly better than the third-best proxy.

TABLE 4

The Relation between Opinion Divergence and Proxies
Categorization by Preponderance of (Non-) Program Trade Orders in DIVOP and by PIN

Panel A: Categorization by order-type usage (program trade orders vs. nonprogram trade orders) in DIVOP calculation

Program trade orders?	Proxy Variable						
	ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$	N
None in DIVOP	0.1074[c]	0.0007[c]	0.0818[c]	*−0.0076*[b]	*−0.00004*[a]	*−0.0002*[c]	13,017
	(0.2053)	(0.2046)	(0.1996)	*(0.1957)*	*(0.1955)*	*(0.1966)*	
Only these in DIVOP	0.1034[c]	0.0007[c]	0.0828[c]	*−0.0056*	*−0.00004*[a]	*−0.0002*[c]	12,546
	(0.1977)	(0.1985)	(0.1924)	*(0.1879)*	*(0.1879)*	*(0.1890)*	
Relatively Fewer	0.1066[c]	0.0006[c]	0.0581[c]	*−0.0059*[a]	*−0.00004*	*−0.0002*[c]	6,307
	(0.3533)	(0.3439)	(0.3361)	*(0.3335)*	*(0.3333)*	*(0.3348)*	
Relatively More	0.0958[c]	0.0004[c]	0.0664[c]	*−0.0023*	0.000001	0.0001[a]	6,308
	(0.1071)	(0.1080)	(0.1061)	*(0.1000)*	(0.0999)	(0.1003)	

Panel B: Differences in relative explanatory power (Vuong 1989)
Rankings of explanatory power of proxies based on adjusted R^2:

Prgrm	(1)	>	(2)	>	(3)	>	(4)	>	(5)	>	(6)
NO	ΔTO		SUV[y]		Spread		$\mathit{\sigma_{returns}}$		$\mathit{\sigma_{frcsts\_2}}$		$\mathit{\sigma_{frcsts}}$
YES	SUV		ΔTO[z]		Spread		$\mathit{\sigma_{frcsts\_2}}$		$\mathit{\sigma_{frcsts}}$		$\mathit{\sigma_{returns}}$
Few	ΔTO[y]		SUV[z]		Spread		$\mathit{\sigma_{frcsts\_2}}$		$\mathit{\sigma_{returns}}$		$\mathit{\sigma_{frcsts}}$
More	SUV		ΔTO		Spread[y]		$\sigma_{\text{frcsts\_2}}$		σ_{frcsts}		$\mathit{\sigma_{returns}}$

Panel C: Categorization by PIN (probability of informed trading) level

PIN Level	Proxy Variable						
	ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$	N
High	0.1073[c]	0.0007[c]	0.0768[c]	*−0.0079*[b]	*−0.00002*	*−0.0001*[b]	5,398
	(0.2313)	(0.2257)	(0.2169)	*(0.2121)*	*(0.2114)*	*(0.2121)*	
Low	0.0843[c]	0.0003[c]	0.0567[c]	0.0002	0.00003	*−0.0013*[b]	5,405
	(0.1868)	(0.1790)	(0.1797)	(0.1749)	(0.1753)	*(0.1756)*	

Panel D: Differences in relative explanatory power (Vuong 1989)
Rankings of explanatory power of proxies based on adjusted R^2:

PIN	(1)	>	(2)	>	(3)	>	(4)	>	(5)	>	(6)
High	ΔTO		SUV[y]		Spread		$\mathit{\sigma_{frcsts\_2}}$		$\mathit{\sigma_{returns}}$		$\mathit{\sigma_{frcsts}}$
Low	ΔTO		Spread		SUV[y]		σ_{frcsts}		σ_{returns}		$\mathit{\sigma_{frcsts\_2}}$

Table presents regressions of opinion divergence on proxies for it and control variables for firm size and stock price (coefficients on controls are not reported). There are six regressions in each row. Each row's sample varies with the condition listed, based on availability or propensity of specific order types (program or nonprogram) to calculate DIVOP. Control variables: Market value of equity quintile is assigned based on which quintile of NYSE firms the firm's market value of equity falls in (at the end of 2001). P_inverse equals one over stock price, and is specific to the firm/day. Proxies: ΔTO is calculated as market-adjusted turnover on the day, minus the average of market-adjusted turnover during a control period (December 2001). SUV is the scaled (by estimation window (November/December 2001) standard deviation of prediction errors) one-day prediction error from a market model-style regression of volume on absolute valued returns. Spread is the average bid-ask spread across all quotes during normal trading hours on the day. σ_{returns} is the standard deviation of transaction-to-transaction returns, across all transactions during normal trading hours on the day. σ_{frcsts} is the standard deviation of analysts' forecasts, divided by the average forecast. Forecasts used are submitted during the month corresponding to the day of analysis. If the dependent variable is from January 30, 2002, then forecasts issued during the month of January 2002 are used. $\sigma_{\text{frcsts\_2}}$ is the standard deviation of analysts' forecasts, divided by the firm's average stock price during the month. [a,b,c] indicates significance of coefficient at the 10%, 5%, and 1% levels, respectively. *Adjusted R^2s for the full regression (including the control variables) are reported below each coefficient in parentheses.* Panel B presents the rank ordering of explanatory power of the proxies for DIVOP (based on adjusted R^2). Tests of differences in explanatory power are between the variable and its adjacent variable to the right (for the prgrm–NO sample, ΔTO is compared to SUV, SUV is compared to spread, etc.), and are based on Vuong [1989]. *Proxies carrying negative coefficients are italicized and treated as having lower explanatory power for DIVOP (they appear to the right of proxies with positive coefficients in Vuong tests).* Panel C presents regression results from DIVOP on proxies, categorized by the relative level of PIN (probability of informed trading—Easley, Hvidkjaer and O'Hara [2002]). PIN is calculated monthly. Panel D is analogous to panel B in presenting Vuong [1989] tests of relative explanatory power of proxies.

[x,y,z] indicates significant Vuong test values at the 10%, 5%, and 1% levels, respectively.

DIVOP from the main sample, and then condition on whether the *fraction* of orders that were program trade orders was above or below the median (within the main sample of daily DIVOP values).

In general, my results continue to indicate that unexplained volume is a better proxy than others for opinion divergence.[29] The lone (albeit weak) exception is when the fraction of orders that are program trade oriented is above the median, in which case spreads are not significantly worse than unexplained volume as a proxy for opinion divergence. One possible explanation for this result is as follows.

A higher proportion of program trading may present an additional concern to market makers when opinion divergence is high. In particular, high DIVOP by construction implies limit orders are on average farther away from the consensus price. If program trades execute at the best price available, larger DIVOP makes it more likely that the market maker's bid or ask will represent the best price. To the extent that market makers are concerned with trading against the program (which may be informed trading in some cases), they may raise spreads when DIVOP is high and there is substantial program trading to mitigate the likelihood of offering the best price.

4.2.6. Conditional on Probability of Informed Trading. The above analysis conditions on program trading characteristics to study DIVOP's link with typical proxies. Another perspective on trading characteristics' influence is perhaps more penetrating. Easley, Hvidkjaer, and O'Hara [2002] measure the PIN by studying the characteristics of order flow. They interpret high PIN levels as an indicator of greater adverse selection (or AI) concerns faced by market makers. I therefore condition on PIN levels and re-examine the relationships between investor opinion divergence and proxies.[30]

I construct firm-level PINs over the three separate months of my data window. In preliminary analysis, regressing DIVOP on PIN, there is a significantly negative relation, suggesting that PIN is not a good proxy for investor opinion divergence.[31] Nevertheless, conditioning on PIN might still influence the relation between DIVOP and typical proxies. Following the (program trading) conditional analysis in table 4 panels A and B, I segment my sample into observations with relatively higher (above the sample median) versus relatively lower levels (below the sample median) of PIN. For each subsample, I regress DIVOP on the typical proxies. I present the results in panels C and D of table 4.

The results for high PIN firms are similar to the main results presented in table 2. For this subsample, unexplained volume is the dominant proxy.

---

[29] While these results are inconsistent with the predictions of Shalen [1993], because volume is always a better proxy than stock return volatility for opinion divergence, this may simply be due to my lack of an adequate measure for risk-weighted hedging demand.

[30] I use the methods of Easley, Hvidkjaer, and O'Hara [2002] to construct PIN. Marginally different PIN construction approaches are discussed in Brown, Hillegeist, and Lo [2004], Venter and de Jongh [2004], and Brown and Hillegeist [2007].

[31] Results are not tabled for brevity, but are available from the author upon request.

Vuong [1989] tests do not allow discrimination between ΔTO and *SUV* as better proxies, but both carry significantly greater explanatory power than the next best proxy (Spread).

By contrast, among low PIN firms, unexplained volume and spreads carry similar explanatory power for opinion divergence. This is inconsistent with the interpretation of evidence from sampling on firm/days with relatively more program trades. One possible reason for this inconsistency is that PIN may not effectively measure AI concerns of market makers. Duarte and Young [2009] suggest this and further find that low PIN implies a high level of liquidity, which also suggests low spreads. If DIVOP is also low when PIN is low, this would imply the strong positive correlation between DIVOP and spreads when sampling on low PIN stocks. This is intuitively appealing. Low levels of PIN occur when factors leading to low DIVOP are in place. Specifically, relatively fewer limit orders that are far from the respective previous transaction price suggests more depth around the most recent trade price. This suggests better liquidity.

4.2.7. Summary. Overall, my inference that unexplained volume is generally a better proxy for opinion divergence than spreads, stock return volatility, or analyst forecast dispersion, survives most conditional analyses. The most striking exception appears to be when there is insufficient analyst following to estimate forecast dispersion. In this case, spreads appear be the dominant proxy for opinion divergence. Perhaps the lack of analyst activity in a stock increases AI problems and market makers protect against it. I investigate this possibility next.

4.3 DIVOP AND COMPONENTS OF THE BID-ASK SPREAD

Hong, Lim, and Stein [2000] suggest analysts help to incorporate firm specific information into stock prices quickly (reducing the profitability of momentum strategies). If a firm is not followed by analysts, stock prices may be less informative and traders with private information could be relatively more informed (i.e. AI problems are more pronounced). This could increase market makers' concerns with adverse selection, causing them to widen spreads. If greater AI also associates with more divergent expressed opinions (via orders), increasing DIVOP, this could explain the dominance of spread as a proxy for DIVOP when analyst coverage doesn't exist. The key link is the larger AI component of the spread.

I test whether the AI component of spreads correlates more strongly with DIVOP when I include nonanalyst-followed firms in the sample. This necessitates decomposing spreads into components for OP and AI. I use the methodology of George, Kaul, and Nimalendran [1991].[32]

[32] I do not estimate inventory holding costs because George, Kaul, and Nimalendran [1991] find that inventory holding costs are effectively zero. In fact, their methodology (which I follow) assumes they are zero. They later document the reasonableness of this assumption.

TABLE 5

The Relation between Opinion Divergence and Two Components of Spread

	OP	AI
Panel A: Regressions		
Coefficient	0.0275[a]	0.0231
Adj. R^2	(0.2086)	(0.2085)
Panel B: Full sample (with and without $\sigma_{\mathbf{frcsts}}$ data) Regressions		
Coefficient	0.2938[c]	0.2354[c]
Adj. R^2	(0.4772)	(0.4721)

Table presents regressions of opinion divergence on two components of bid-ask spread (coefficients on control variables are not reported for readability): Order Processing component and AI component. The sample is the 10,760 observations from the main sample with sufficient data to calculate the two components of the spread. Control variables: Market value of equity quintile is assigned based on which quintile of NYSE firms the firm's market value of equity falls in (at the end of 2001). P_inverse equals one over stock price, and is specific to the firm/day. Proxies: OP is the order processing component of the bid-ask spread, calculated using the methodology of George, Kaul and Nimalendran [1991]. AI is the asymmetric information components of the bid-ask spread, calculated using the methodology of George, Kaul and Nimalendran [1991]. Spread is the average bid-ask spread across all quotes during normal trading hours on the day. *Adjusted R^2s for the full regression (including the control variables) are reported below each coefficient in parentheses.* Panel B presents regressions on the full sample of 21,763 observations, regardless of whether there is data on analysts' forecast dispersion for the observation.

[a,b,c] indicates significance of coefficient at the 10%, 5%, and 1% levels, respectively.

The results are presented in table 5. It appears that in the main sample—with data on analysts' forecast dispersion available—only the OP component is significantly positively related to DIVOP (panel A). The coefficient on the OP component variable is significant at the 10% level. The coefficient on the AI component variable is not.

By contrast, when I re-investigate the relation between spread components and DIVOP (in panel B) for the larger sample *including* firms without IBES coverage, I find that the coefficient on AI is significantly positively related to DIVOP. The coefficient is an order of magnitude larger than when I restrict my sample to analyst-followed firms. The evidence supports the above interpretation. Analyst coverage reduces AI problems and disconnects a potential link between spreads and investor opinion divergence.

4.4 FACTOR ANALYSIS TESTS

To this point, my statistical approach has been standard regression analysis. To add an element of robustness to my conclusions, I now view the relations between opinion divergence and proxies from a factor analysis perspective.

I begin by extracting the first common factor from the six typical proxies. Second, I assess the relation between each proxy and this factor. I then examine the relation between the first common factor and DIVOP. My results generally suggest that unexplained volume proxies are best.

Table 6, panel A1 presents raw correlations between the first common factor from the six typical proxies for opinion divergence, and the individual proxies. Only the two unexplained volume proxies correlate positively with the first factor drawn from all six. This evidence suggests that the

TABLE 6
The Relationship between DIVOP and Proxies—Factor Analysis

Panel A1: Correlations between proxies and first common factor of the proxies

Proxy	ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$
Weight	0.79845	0.55129	−0.08147	−0.02234	−0.00431	−0.04258

Panel A2: Weights on proxies in the first common factor of the proxies

Proxy	ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$
Weight	0.70194	0.18748	−0.07400	−0.01403	0.00094	−0.02168

Panel B: Regression

	Factor	ΔTO
Coefficient	0.0009[c]	0.1047[c]
Adj. R^2	(0.2825)	(0.2815)

Panel B1: Differences in relative explanatory power [Vuong [1989]]

Rankings of explanatory power of proxies based on adjusted R^2:

(1)	$\geq$	(2)
Factor[y]		ΔTO

Panel A1 presents the correlation between proxies and the first common factor. Panel A2 presents the weight of each proxy in a linear combination of proxies and weights that yields the first principal component from factoring the proxies. These are the standardized scoring coefficients. Proxies: ΔTO is calculated as market-adjusted turnover on the day, minus the average of market-adjusted turnover during a control period (December 2001). SUV is the scaled (by estimation window (November/December 2001) standard deviation of prediction errors) one-day prediction error from a market model-style regression of volume on absolute valued returns. Spread is the average bid-ask spread across all quotes during normal trading hours on the day. σ_{returns} is the standard deviation of transaction-to-transaction returns, across all transactions during normal trading hours on the day. σ_{frcsts} is the standard deviation of analysts' forecasts, divided by the average forecast. Forecasts used are submitted during the month corresponding to the day of analysis. If the dependent variable is from January 30, 2002, then forecasts issued during the month of January 2002 are used. $\sigma_{\text{frcsts\_2}}$ is the standard deviation of analysts' forecasts, divided by the firm's average stock price during the month. Panel B presents the coefficient from regressing DIVOP on the first common factor from the proxies, and from (separately) regressing on ΔTO. *Adjusted R^2s for the full regression (including the control variables) are reported below each coefficient in parentheses.* Panel B1 presents the rank ordering of explanatory power of the proxies for DIVOP (based on adjusted R^2). Tests of differences in explanatory power are between the variable and its adjacent variable to the right (Factor1 is compared to ΔTO), and are based on Vuong [1989]. Panel C presents raw correlations between each proxy and the first principal component from five different values of DIVOP (each value of DIVOP is an average within one of the five different size quintiles). An additional proxy is included in the correlations analysis: $\sigma_{\text{frcsts\_numerator}}$ is the numerator from both analyst forecast dispersion proxies.

[x,y,z] indicates significant Vuong test values at the 10%, 5%, and 1% levels, respectively.

common underlying determinant associated with all opinion divergence proxies (which one might reasonably presume is an indicator of opinion divergence) is most closely related to unexplained volume. Confirmatory evidence is seen in panel A2 of table 6. This presents the weights necessary on each proxy (in a linear combination of weights and proxies) to yield the factor's value.[33] Again, only the two unexplained volume proxies carry positive weights.

Taken together, the first common factor from all six opinion divergence proxies appears to be largely attributable to unexplained volume. However, this need not indicate that unexplained volume is actually a dominant proxy for opinion divergence. It's possible that the factor is not a strong explainer of DIVOP.

[33] These are standardized scoring coefficients.

TABLE 7

The Relation between Opinion Divergence and Proxies Around Earnings Announcements

Panel A: Regressions

Proxy	ΔTO	SUV	Spread	σ_{returns}	σ_{frcsts}	$\sigma_{\text{frcsts\_2}}$
Proxy	0.0461[b]	0.0018[c]	0.0183	0.0072	*−0.0002*	0.0987
	(0.3403)	(0.3518)	(0.3196)	(0.3192)	*(0.3200)*	(0.3240)

Panel B: Differences in relative explanatory power (Vuong 1989)

Rankings of explanatory power of proxies based on adjusted R^2:

(1)	>	(2)	>	(3)	>	(4)	>	(5)	>	(6)
SUV		ΔTO^{y}		$\sigma_{\text{frcsts\_2}}$		Spread		σ_{returns}		$\sigma_{\textit{frcsts}}$

Table presents regressions of opinion divergence on proxies (coefficients on control variables are not reported for readability). There are six regressions. The sample is 248 earnings announcements by firms with complete data to calculate proxies. Control variables: Market value of equity quintile is assigned based on which quintile of NYSE firms the firm's market value of equity falls in (at the end of 2001). P_inverse equals one over stock price, and is specific to the firm/day. Proxies: ΔTO is calculated as market-adjusted turnover on the day, minus the average of market-adjusted turnover during a control period (December 2001). SUV is the scaled (by estimation window (November/December 2001) standard deviation of prediction errors) one-day prediction error from a market model-style regression of volume on absolute valued returns. Spread is the average bid-ask spread across all quotes during normal trading hours on the day. σ_{returns} is the standard deviation of transaction-to-transaction returns, across all transactions during normal trading hours on the day. σ_{frcsts} is the standard deviation of analysts' forecasts, divided by the average forecast. Forecasts used are submitted during the 30 calendar days following the earnings announcement. $\sigma_{\text{frcsts\_2}}$ is the standard deviation of analysts' forecasts, divided by the firm's stock price two days before the earnings event. [a,b,c] indicates significance at the 10%, 5%, and 1% levels, respectively. *Adjusted R^2s for the full regression (including the control variables) are reported below each coefficient in parentheses.* Panel B presents the rank ordering of explanatory power of the proxies for DIVOP (based on adjusted R^2). Tests of differences in explanatory power are between the variable and its adjacent variable to the right (SUV is compared to ΔTO, ΔTO is compared to $\sigma_{\text{frcsts\_2}}$, etc.), and are based on Vuong [1989]. *Proxies carrying negative coefficients are italicized and treated as having lower explanatory power for DIVOP (they appear to the right of proxies with positive coefficients in Vuong tests).*

[x,y,z] indicates significant Vuong test values at the 10%, 5%, and 1% levels, respectively.

In fact though, panel B of table 6 shows that the factor is significantly positively correlated with DIVOP. Moreover, a Vuong [1989] test in panel B1 illustrates that it is a significantly better proxy than ΔTO, which was previously the dominant proxy. Overall, the combined evidence suggests that either unexplained volume or a common factor derived from all six proxies (but largely driven by unexplained volume) is the best proxy for opinion divergence.

4.5 AN ALTERNATIVE PERSPECTIVE: OPINION DIVERGENCE AROUND EARNINGS ANNOUNCEMENTS

Up to this point, my analysis focuses on "average" days. However, research into earnings events suggests that they may actually engender diverging investor opinions (e.g. Stice [1991], Lee, Mucklow, and Ready [1993], Kim and Verrecchia [1994], Krinsky and Lee [1996], and Bamber, Barron, and Stober [1997]). Therefore, I re-examine the correlation between my new measure of opinion divergence and the typical proxies, in the three-day window surrounding earnings announcements for my sample.[34]

[34] Defining the Compustat earnings date as day 0, the window I examine is [−1, +1].

My sample of firms makes 248 earnings announcements during the January–March 2002 calendar window. For each of these events, I calculate three-day versions of DIVOP and the following proxies: both unexplained volume measures, bid-ask spread, and stock return volatility. I also calculate analyst forecast dispersion using forecasts from the 30 calendar days following the earnings announcement date. I believe it is critically important to not use forecasts prior to the earnings event in this analysis, as the event is likely to contain information that could change the analyst's forecast. Again, McNichols and O'Brien [1997] note that analysts may drop coverage rather than adjusting their forecast in the face of new information. This behavior could bias up the measure of analyst forecast dispersion if I used forecasts issued prior to the event. Such bias seems more likely when focusing on the time period around earnings events.

Table 7 presents results from regressions that mimic those in table 2. They indicate that only unexplained volume correlates reliably positively with DIVOP. No other proxies correlate significantly with DIVOP around earnings events. A Vuong [1989] test fails to indicate a dominant proxy among the two unexplained volume measures. Conditional on an earnings announcement, it appears that only unexplained volume is a reliable proxy for investor opinion divergence.

5. *Conclusion*

This paper examines the empirical validity of extant proxies for opinion divergence. I derive a new measure for opinion divergence using investors' expressions of interest in stocks through their limit and market orders. My sample is all NYSE-listed firms with tickers beginning with the letters A–D, and with order (and other) data from January to March 2002. In regressions examining the link between my new measure of opinion divergence and the proxies, only spreads and unexplained volume carry significant explanatory power. These tests suggest that variability in stock returns and dispersion in analysts' forecasts are weaker proxies.

Given three potentially useful proxies for opinion divergence, I conduct Vuong [1989] tests of relative explanatory power. By in large, unexplained volume—particularly "change in turnover" (ΔTO) appears to be the dominant proxy.

Conditional tests on subsamples of firms (and orders) generally confirm the above conclusion. However, there are some exceptions. The most significant is that for firms with (little to) no analyst following, spread is the dominant proxy for opinion divergence. This is consistent with work by Hong, Lim, and Stein [2000], who show that analyst coverage changes the information environment. When analysts cover a firm, AI problems appear to be low and unrelated to investor opinion divergence. By contrast, when there is no analyst coverage, AI concerns are much higher and strongly positively related to investor opinion divergence. A fuller exploration of the importance of analyst coverage for the link between AI concerns and investor opinion divergence may be fruitful.

I also examine the correlation between investor opinion divergence and the proxies around earnings announcements. Again, unexplained volume proxies appear to be the best proxy for investor opinion divergence. Finally, factor analysis also suggests that unexplained volume is a significantly better proxy for investor opinion divergence.

Taken together, the bulk of my evidence supports the use of unexplained volume as the dominant proxy for investor opinion divergence. In regression tests, my new construct for opinion divergence is most highly correlated with two measures of unexplained volume. Second, forecast dispersion and stock return volatility are *negatively* related to my construct. This calls into serious question prior work treating these proxies as indicators of dispersed *investor* opinions. Finally, my factor analysis results suggest that combining several proxies into a single construct may yield a better overall proxy. But it is noteworthy that the largest determinants of this factor are unexplained volume proxies.

Further work might also be conducted focusing on opinion divergence associated with securities traded on NASDAQ, as my results are based on a sample of NYSE stocks.[35] In particular, the unique specialist system on NYSE may lead to faster opinion convergence than on NASDAQ (see Hasbrouck [1995] and Garfinkel and Nimalendran [2003] for studies of the effects of market structure on price discovery). Also, as noted earlier, future research into the tripartite relation between analyst coverage, AI, and investor opinion divergence appears warranted.

Finally, while this paper is the first to use investors' orders to derive the new measure for investors' opinion divergence, work in this area remains. For example, my data necessarily focus attention on investors who believe strongly enough to submit orders for trade. Perhaps experimental work can help us understand the strength of my new measure relative to one that employs different information sets given to different potential investors, not all of whom choose to submit orders.

REFERENCES

ABARBANELL, J.; W. LANEN; AND R. VERRECCHIA. "Analysts' Forecasts as Proxies for Investor Beliefs in Empirical Research." *Journal of Accounting and Economics* 20 (1995): 31–60.

AGRAWAL, A.; S. CHADHA; AND M. CHEN. "Who Is Afraid of Reg FD? The Behavior and Performance of Sell-Side Analysts Following the SEC's Fair Disclosure Rules." *Journal of Business* 79 (2006): 2811–34.

AHN, H.; K. BAE; AND K. CHAN. "Limit Orders, Depth and Volatility: Evidence from the Stock Exchange of Hong Kong." *The Journal of Finance* 56 (2001): 767–88.

AJINKYA, B.; R. ATIASE; AND M. GIFT. "Volume of Trading and the Dispersion in Financial Analysts' Earnings Forecasts." *The Accounting Review* 66 (1991): 389–401.

AJINKYA, B.; R. ATIASE; AND M. GIFT. "Heterogeneous Prior Beliefs, Differential Interpretation and the Consensus Effect of Quarterly Earnings Signals." Working paper, University of Texas at Austin, Austin, TX, 2004.

[35] Of course, the data are likely to be much harder to obtain, given the lack of a single trading venue.

ALEXANDRIDIS, G.; A. ANTONIOU; AND D. PETMEZAS. "Divergece of Opinion and Post-Acquisition Performance." *Journal of Business, Finance* 34 (2007): 439–60.

BAGEHOT, W. [pseudo.]. "The Only Game in Town." *Financial Analyst Journal* 27 (1971): 12–14.

BAMBER, L. "Unexpected Earnings, Firm Size and Trading Volume around Quarterly Earnings Announcements." *The Accounting Review* 62 (1987): 510–32.

BAMBER, L.; O. BARRON; AND T. STOBER. "Trading Volume and Different Aspects of Disagreement Coincident with Earnings Announcements." *The Accounting Review* 72 (1997): 575–97.

BAMBER, L.; O. BARRON; AND T. STOBER. "Differential Interpretations and Trading Volume." *Journal of Financial and Quantitative Analysis* 34 (1999): 369–86.

BARRON, O.; M. STANFORD; AND Y. YU. "Further Evidence on the Relation between Analysts' Forecast Dispersion and Stock Returns." *Contemporary Accounting Research* 26 (2009): 329–57.

BEAVER, W. "The Information Content of Annual Earnings Announcements." *Journal of Accounting Research* 6 (1968): 67–92.

BENSTON, G., AND R. HAGERMAN. "Determinants of Bid-Ask Spreads in the Over-the-Counter Market." *Journal of Financial Economics* 1 (1974): 353–64.

BLOOMFIELD, R.; M. O'HARA; AND G. SAAR. "The 'Make or Take' Decision in an Electronic Market: Evidence on the Evolution of Liquidity." *Journal of Financial Economics* 75 (2005): 165–200.

BOEHME, R. D.; B. R. DANIELSEN; AND S. M. SORESCU. "Short-Sale Constraints, Dispersion of Opinion and Overvaluation." *Journal of Financial and Quantitative Analysis* 41 (2006): 455–87.

BOEHMER, E.; G. SAAR; AND L. YU. "Lifting the Veil: An Analysis of Pre-Trade Transparency at the NYSE." *The Journal of Finance* 60 (2005): 783–815.

BRANCH, B., AND W. FREED. "Bid-Asked Spreads on the AMEX and the Big Board." *The Journal of Finance* 32 (1977): 159–63.

BROWN, L.; G. FOSTER; AND E. NOREEN. *Security Analyst Multi-Year Earnings Forecasts and the Capital Market*. Studies in Accounting Research, No. 21. Sarasota, FL: American Accounting Association, 1985.

BROWN, S., AND S. HILLEGEIST. "How Disclosure Quality Affects the Level of Information Asymmetry." *Review of Accounting Studies* 12 (2007): 443–77.

BROWN, S.; S. HILLEGEIST; AND K. LO. "Conference Calls and Information Asymmetry." *Journal of Accounting and Economics* 37 (2004): 343–66.

BYARD, D. "Firm Size and the Benefit of Analysts' Heterogeneous Interpretations," Working paper. Web site, http://ssrn.com/abstract=353200, 2002.

CHATTERJEE, S.; K. JOHN; AND A. YAN. "Takeovers and Divergence of Investor Opinion," Working paper. Web site, http://ssrn.com/abstract=1327289, 2009.

CHEN, X., AND Q. CHENG. "Issuing Stock Recommendations: Theory and Evidence." Working paper, University of British Columbia, 2003.

CHORDIA, T.; S. W. HUH; AND A. SUBRAHMANYAM. "The Cross-Section of Expected Trading Activity." *Review of Financial Studies* 20 (2007): 709–40.

CORE, J., AND W. GUAY. "Estimating the Value of Employee Stock Option Portfolios and Their Sensitivities to Price and Volatility." *Journal of Accounting Research* 40 (2002): 613–30.

CRABBE, L., AND M. POST. "The Effect of a Rating Downgrade on Outstanding Commercial Paper." *The Journal of Finance* 46 (1994): 39–56.

DAIGLER, R., AND M. WILEY. "The Impact of Trader Type on the Futures Volatility-Volume Relation." *The Journal of Finance* 54 (1999): 2297–316.

DANIELSON, B., AND S. SORESCU. "Why Do Option Introductions Depress Stock Prices? A Study of Diminishing Short Sale Constraints." *Journal of Financial and Quantitative Analysis* 36 (2001): 451–84.

DECHOW, P.; A. HUTTON; AND R. SLOAN. "The Relation between Analysts' Forecasts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings." *Contemporary Accounting Research* 17 (2000): 1–32.

DIETHER, K. B.; C. J. MALLOY; AND A. SCHERBINA. "Differences of Opinion and the Cross-Section of Stock Returns." *The Journal of Finance* 57 (2002): 2113–42.

DOUKAS, J.; C. KIM; and C. PANTZALIS. "Divergence of Opinion and Equity Returns." *Journal of Financial and Quantitative Analysis* 41 (2006): 573–606.

DUARTE, J., AND L. YOUNG. "Why Is PIN Priced?." *Journal of Financial Economics* 91 (2009): 119–38.

EASLEY, D.; S. HVIDKJAER; AND M. O'HARA. "Is Information Risk a Determinant of Asset Returns?," *The Journal of Finance* 57 (2002): 2185–221.

FLEMING, M., AND E. REMOLONA. "Price Formation and Liquidity in the U.S. Treasury Market: the Response to Public Information." *The Journal of Finance* 54 (1999): 1901–15.

FOUCAULT, T. "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market." *Journal of Financial Markets* 2 (1999): 99–134.

FRANKEL, J., AND K. FROOT. "Chartists, Fundamentalists and Trading in the Foreign Exchange Market." *American Economic Review* 80 (1990): 181–5.

GARFINKEL, J. A., AND M. NIMALENDRAN. "Market Structure and Trader Anonymity: An Analysis of Insider Trading." *Journal of Financial and Quantitative Analysis* 38 (2003): 591–610.

GARFINKEL, J. A., AND J. S. SOKOBIN. "Volume, Opinion Divergence and Returns: A Study of Post-Earnings Announcement Drift." *Journal of Accounting Research* 44 (2006): 85–112.

GEBHARDT, W.; C. LEE; AND B. SWAMINATHAN. "Toward an Implied Cost of Capital." *Journal of Accounting Research* 39 (2001): 135–76.

GEORGE, T.; G. KAUL; AND M. NIMALENDRAN. "Estimation of the Bid-Ask Spread and Its Components: A New Approach." *Review of Financial Studies* 4 (1991): 623–56.

GOYAL, A., AND P. SANTA-CLARA. "Idiosyncratic Risk Matters." *The Journal of Finance* 58 (2003): 975–1008.

GRAHAM, J.; J. KOSKI; AND U. LOEWENSTEIN. "Information Flow and Liquidity around Anticipated and Unanticipated Dividend Announcements." *Journal of Business* 79 (2006): 2301–36.

GUTTMAN, I. "The Timing of Analysts' Earnings Forecasts." Stanford University Working Paper, 2005.

HANDA, P.; R. SCHWARTZ; AND A. TIWARI. "Quote Setting and Price Formation in an Order Driven Market." *Journal of Financial Markets* 6 (2003): 461–89.

HARRIS, L., AND J. HASBROUCK. "Market vs. Limit Orders: The SuperDOT Evidence on Order Submission Strategy." *Journal of Financial and Quantitative Analysis* 31 (1996): 213–31.

HARRIS, M., AND A. RAVIV. "Differences of Opinion Make a Horse Race." *Review of Financial Studies* 6 (1993): 473–506.

HASBROUCK, J. "One Security, Many Markets: Determining the Contributions to Price Discovery." *The Journal of Finance* 50 (1995): 1175–99.

HEALY, P., AND K. PALEPU. "Information Asymmetry, Corporate Disclosure, and the Capital Markets: A Review of the Empirical Disclosure Literature." *Journal of Accounting and Economics* 31 (2001): 405–40.

HOLLIFIELD, B.; R. MILLER; P. SANDAS; AND J. SLIVE. "Estimating the Gains from Trade in Limit Order Markets." *The Journal of Finance* 61 (2006): 2753–804.

HOLTHAUSEN, R., AND R. VERRECCHIA. "The Effect of Informedness and Consensus on Price and Volume Behavior." *The Accounting Review* 65 (1990): 191–208.

HONG, H.; T. LIM; AND J. STEIN. "Bad News Travels Slowly: Size, Analyst Coverage and the Profitability of Momentum Strategies." *The Journal of Finance* 55 (2000): 265–95.

HONG, H., AND J. STEIN. "A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets." *The Journal of Finance* 54 (1999): 2143–85.

HOUGE, T.; T. LOUGHRAN; G. SUCHANEK; AND X. YAN. "Divergence of Opinion, Uncertainty and the Quality of Initial Public Offerings." *Financial Management* 30 (2001): 5–23.

KANDEL, E., AND N. PEARSON. "Differential Interpretation of Public Signals and Trade in Speculative Markets." *Journal of Political Economy* 103 (1995): 831–72.

KANIEL, R., AND H. LIU. "So What Orders Do Informed Traders Use?." *Journal of Business* 79 (2006): 1867–913.

KARPOFF, J. "The Relation between Price Changes and Trading Volume: A Survey." *Journal of Financial and Quantitative Analysis* 22 (1987): 109–26.

KIM, O., AND R. VERRECCHIA. "Market Liquidity and Volume around Earnings Announcements." *Journal of Accounting and Economics* 17 (1994): 41–67.

KRINSKY, I., AND J. LEE. "Earnings Announcements and the Components of the Bid-Ask Spread." *The Journal of Finance* 51 (1996): 1523–35.

LEE, C.; B. MUCKLOW; AND M. READY. "Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis." *Review of Financial Studies* 6 (1993): 345–74.

LERMAN, A.; J. LIVNAT; AND R. MENDENHALL. "The High-Volume Return Premium and Post-Earnings Announcement Drift." Working paper, NYU, 2008.

LEV, B., AND J. OHLSON. "Market-Based Empirical Research in Accounting: A Review, Interpretation, and Extension." *Journal of Accounting Research* 20 (1982): 249–322.

LIN H., AND M. MCNICHOLS. "Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations." *Journal of Accounting and Economics* 25 (1998): 101–27.

LJUNGQVIST, A.; C. MALLOY; AND F. MARSTON. "Rewriting History." *The Journal of Finance* (2009): Forthcoming.

MCNICHOLS, M., AND P. O'BRIEN. "Self-Selection and Analyst Coverage." *Journal of Accounting Research* 35 (supplement) (1997): 167–99.

PETERSEN, M. A., AND D. FIALKOWSKI. "Posted Versus Effective Spreads: Good Prices or Bad Quotes?." *Journal of Financial Economics* 35 (1994): 269–92.

SCHERBINA, A., AND R. SADKA. "Analyst Disagreement, Forecast Bias and Stock Returns." *The Journal of Finance* 62 (2007): 2367–403.

SHALEN, C. "Volume, Volatility, and the Dispersion of Beliefs." *Review of Financial Studies* 6 (1993): 405–34.

STICE, J. "Using Financial and Market Information to Identify Pre-Engagement Factors Associated with Lawsuits against Auditors." *The Accounting Review* 66 (1991): 516–33.

TKAC, P. "A Trading Volume Benchmark: Theory and Evidence." *Journal of Financial and Quantitative Analysis* 34 (1999): 89–114.

VENTER, J., AND D. DE JONGH. "Extending the EKOP Model to Estimate the Probability of Informed Trading," Proceedings of the First African Finance Conference, 2004.

VUONG, Q. "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses." *Econometrica* 57 (1989): 307–33.

WANG, F. "Strategic Trading, Asymmetric Information and Heterogeneous Prior Beliefs." *Journal of Financial Markets* 1 (1998): 321–52.

X. ZHANG. "Information Uncertainty and Analyst Forecast Behavior." *Contemporary Accounting Research* 23 (2006a): 565–90.

X. ZHANG. "Information Uncertainty and Stock Returns." *The Journal of Finance* 61 (2006b): 105–36.