THE ACCOUNTING REVIEW
Vol. 90, No. 5
2015
pp. 2079–2106

American Accounting Association
DOI: 10.2308/accr-51025

Volatility Forecasting Using Financial Statement Information

Suhas A. Sridharan
University of California, Los Angeles

ABSTRACT: This paper examines whether financial statement information can predict future realized equity volatility incremental to market-based equity volatility forecasts. I use an analytical framework to identify accounting-based drivers of realized volatility. My main hypothesis is that accounting-based drivers can be used to forecast future realized volatility incremental to either past realized volatility or option-implied volatility. I confirm this empirically and document abnormal returns to an option-based trading strategy that takes a long (short) position in firms with financial statement information indicative of high (low) future realized volatility. These results suggest that accounting-based volatility drivers may serve as useful indicators of variance risk. Finally, I demonstrate that the incorporation of accounting-based fundamental information into forecasting models yields lower forecast errors relative to models based solely on past realized volatility.

Keywords: *volatility; option pricing; fundamental analysis; variance risk.*

I. INTRODUCTION

This article examines how accounting information can be used to supplement existing equity volatility forecasting techniques to generate more precise equity volatility forecasts. It documents a new channel through which accounting information is useful for equity valuation. Studies relating corporate disclosure to equity valuation typically focus on how accounting information can enhance predictions of the mean of future equity returns. However, this focus overlooks a basic result from asset pricing theory: equity investors seek to maximize mean returns while minimizing the risk associated with their investments (Markowitz 1952). The duality of the equity investor's decision implies that simply linking disclosures with mean returns is insufficient to fully capture the informativeness of these disclosures for equity valuation. An equally relevant consideration is how these disclosures relate to the risk in equity investment.

I appreciate the helpful comments from Mary Barth, Jack Hughes, Travis Johnson, Ron Kasznik, Arthur Korteweg, Charles Lee, Maria Ogneva, Eric So, two anonymous reviewers, Leslie D. Hodder (editor), and seminar participants at the Stanford University Graduate School of Business, London Business School, The University of Utah, University of California, Los Angeles, New York University, Columbia University, The George Washington University, Boston University, and The Ohio State University.

Editor's note: Accepted by Leslie D. Hodder.

Submitted: May 2012
Accepted: January 2015
Published Online: January 2015

Equity volatility realizations provide an estimate of the total risk in equity investment, including idiosyncratic risk. In theory, investors can eliminate exposure to idiosyncratic risk through diversification. In practice, however, limits to diversification can leave investors exposed to some amount of idiosyncratic risks, making realized equity volatility a relevant risk metric for equity valuation (Ang, Hodrick, Xing, and Zhang 2006; Fu 2009). As measures of market uncertainty, equity volatility estimates are also frequently used for macroeconomic analyses, making them important to academics, practitioners, and market regulators.

Equity volatility forecasts are also useful to debt market investors and other potential creditors for calibrating assessments of default risk. The Merton (1973) model is commonly used to assess a firm's likelihood of default, and explicitly calls for an estimate of future asset volatility, of which future equity volatility is a predictor. Similarly, volatility forecasts are central to derivatives trading because they are a key input in many derivative pricing models. Both the debt and derivatives markets are substantial components of the global capital markets. Trading in equity options alone accounted for more than six trillion U.S. dollars during 2012, and trading in corporate bond markets has doubled since 2007. The magnitude and recent growth in both the public debt and derivatives markets highlights the importance of improvements in our volatility forecasting techniques.

Prior literature demonstrates that equity returns can be characterized as a function of accounting-based variables (Penman, Reggiani, Richardson, and Tuna 2012). I use this literature as a foundation to develop an expression of equity return variance as a function of information from financial statements. I decompose equity return variance into three components: variance of the earnings yield, variance of the equity price premium, and the covariance of the earnings yield with the equity price premium. I find evidence that accounting-based volatility drivers explain a significant portion of the difference between option-implied volatility and future realized volatility. This difference is often attributed to the premium that investors demand for exposure to variance risk, or to uncertainty about future variance realizations. However, little is known about how variance risk relates to firm fundamentals. This article fills this difference in the literature by demonstrating a systematic relation between accounting-based fundamental information and the difference between implied and realized volatilities.

My evidence is based on a sample of 47,398 quarterly observations from 3,078 firms between 1996 and 2012. The sample is constrained by the data requirements necessary to measure firm-specific implied volatility using the Britten-Jones and Neuberger (2000) model-free approach, which my study is the first to do. My analyses reveal that accounting-based fundamental information is relevant for volatility prediction and is incrementally relevant to the model-free implied volatility in forecasting future realized equity volatility. These results are insensitive to the inclusion of past realized volatility and liquidity measures as controls.

I also document the ability to generate positive returns by taking a long (short) straddle position in firms with financial statement information indicative of low (high) future realized volatility. A straddle position is formed by purchasing at-the-money call and put options on the same underlying asset. This generates a payoff that is increasing in the absolute price change of the underlying asset, but is insensitive to the direction of the change. Consistent with the hypothesis that financial statement information helps to predict future realized equity volatility, I find that sorting 30-day straddle positions using accounting-based volatility drivers improves profitability by 1.5 percent relative to a sorting based solely on implied and past realized volatilities. These returns persist after taking into account the transactions costs in the options market.

Using the methodology developed by Carr and Wu (2009), I examine the ability of accounting-based volatility drivers to predict future option straddle returns after controlling for each firm's *ex ante* variance risk. The results of the analysis suggest that the correlation of accounting-based volatility drivers with variance risk is not the sole driver of their usefulness in estimating option straddle returns. These results also suggest that variance risk does not fully explain the difference between implied and realized volatilities. In addition to variance risk, the difference between implied and realized volatilities might also reflect incomplete processing of relevant accounting information by option market participants analogous to the well-documented under-reaction of equity prices to financial statement information in forecasting future cash flows.

My accounting-based model of equity variance demonstrates how investors can assess equity volatility for firms in the absence of implied volatility estimates. In out-of-sample testing, I find that volatility forecasts that incorporate accounting-based volatility drivers generate significantly lower volatility forecast errors than forecasts based solely on past realized volatility. Past realized volatility is a highly relevant benchmark as it is the best available forecast for the many firms without actively traded equity options. By improving upon this benchmark, my study has practical relevance to debt and equity investors in these firms.

This study also contributes to the literature exploring the usefulness of financial statement information to capital markets. The extent to which accounting-based fundamentals are relevant in forecasting the first moment of equity returns is the subject of a large literature. In contrast, how these disclosures relate to the second moment of equity returns remains relatively unexplored. As there is a fundamental distinction between predicting the mean of equity returns and predicting the volatility of equity returns, this difference in the literature is non-trivial. By linking financial statement information to the realized equity variances, I shed light on how investors might use accounting disclosures to assess risk in equity investment. In doing so, I help develop a more complete understanding of how disclosure can be useful for equity valuation.

II. RELATED LITERATURE AND HYPOTHESIS DEVELOPMENT

This article is related to two primary streams of literature. The first is the literature on financial statement analysis. The second is the literature on implied volatility forecasting. I describe my study in the context of each of these bodies of work in the sections that follow.

The Relation between Mean and Variance

A substantial literature in accounting and finance documents the usefulness of financial statement information for predicting mean returns. If informational inefficiencies exist in the equity market, then it is plausible that financial statement information could predict mean equity returns without providing any information about the uncertainty of equity returns. Without uncertainty of equity returns, there is no reason to expect such information to be useful in the prediction of equity returns. Consider the general definition of the variance of a random variable r with mean μ:

$$Var[r] = E[(r - \mu)^2] = E[r^2] - \mu^2$$

This expression of variance highlights the fact that the exact relationship between the mean and variance of random variable r depends upon the value of the second uncentered moment $E[r^2]$, which, in turn, depends upon the probability distribution of r.

A common assumption in asset pricing models is that equity returns are normally distributed. In the normal distribution, there is no structural relationship between mean and variance. However, normality of a random variable implies an infinite range, which contradicts the reality of bounded investment losses. A popular alternative that reflects the lower bound on equity returns is the lognormal distribution. If equity return r is lognormally distributed with parameters μ and σ^2, then its mean and variance are calculated as follows:

$$E[r] = e^{\mu + \frac{1}{2}\sigma^2} \qquad Var[r] = \left((E[r])^2\right)\left(e^{\sigma^2} - 1\right)$$

The above equations illustrate that even with full knowledge of $E[r]$, one cannot know $Var[r]$

precisely in the lognormal distribution. Variation in σ^2 could still lead two different lognormally distributed variables of equal mean to have significantly different variances.

Identifying Fundamentals

To assess how accounting information might relate to future realized equity returns volatility, I begin with the clean surplus relation:

$$B_{t+1} = B_t + Earn_{t+1} - d_{t+1}$$

Using the clean surplus relation, Penman et al. (2012) develop an expression for expected equity returns from t to $t+1$ in terms of accounting quantities:

$$r_{t+1} = \frac{P_{t+1} + d_{t+1} - P_t}{P_t} = \frac{Earn_{t+1}}{P_t} - \frac{P_{t+1} + B_{t+1} - (P_t - B_t)}{P_t}$$

Extending this model by taking the variance of the alternative expression of r yields:

$$Var[r_{t+1}] = Var\left[\frac{Earn_{t+1}}{P_t}\right] - Var\left[\frac{P_{t+1} + B_{t+1} - (P_t - B_t)}{P_t}\right] \\ -2Cov\left[\frac{Earn_{t+1}}{P_t}, \frac{P_{t+1} + B_{t+1} - (P_t - B_t)}{P_t}\right] \quad (1)$$

Equation (1) reveals that expected future variance in equity returns is driven by future variance in the earnings yield and future variance in the change in premium of market value over book value. The extent to which the earnings yield and the change in premium of market value over book value of equity covary reduces the expected future variance of equity returns. In Equation (1), the accounting variables are defined over the same time period as the volatility realization. To facilitate volatility prediction, I need an expression of future returns volatility as a function of *ex ante* accounting information. To generate such an expression, I assume that current-period realizations of each accounting variable serve as proxies for their future realizations. This leads to my first hypothesis:

H1: Expected future equity volatility is increasing (decreasing) in current earnings yield volatility and current volatility of the change in market to book premium (covariance of earnings yield and change in market to book premium).

Implied Volatility

Several studies explore the use of option-implied volatility as a benchmark forecast for future realized equity volatility (Latane and Rendleman 1976; Chiras and Manaster 1978; Lamoureux and Lastrapes 1993; Christensen and Prabhala 1998). Implied volatility estimates are inferred from observable contemporaneous option prices. The first estimates of option-implied volatility were generated using the Black-Scholes option-pricing formula. Since implied volatility theoretically reflects the market's rational expectation of future volatility, Merton (1973) argues that equity volatility implied by option prices using the Black-Scholes formula should equal the average realized volatility of equity returns over the remaining life of the option. This implies that, in the absence of risk-aversion, a regression of subsequent realized volatility on *ex ante* implied volatility should yield a coefficient of 1 on implied volatility and a coefficient of 0 on any other explanatory variables.[1]

[1] An assumption in this prediction is that the regression includes observations over the remaining life of the option. As the time sample shrinks relative to the remaining option life, the likelihood of observing a non-zero coefficient on other explanatory variables increases. This is further exacerbated for firms with non-stationary realized equity returns variances.

However, empirical evidence on this assertion is mixed. Using Standard & Poor's (S&P) index option data, Day and Lewis (1992) and Lamoureux and Lastrapes (1993) find that past volatility is predictive of future volatility incremental to implied volatility. From this, they conclude that implied volatility is an inefficient predictor of future returns volatility. Using the same data, Canina and Figlewski (1993) show that the correlation between implied and future realized volatilities disappears after one controls for past realized volatility. However, Christensen and Prabhala (1998) find that implied volatility is significantly correlated with future realized volatility and that past realized volatility is fully incorporated in the current market expectation. They find that the coefficient on implied volatility is significantly different from 1, indicating bias in implied volatility.

One limitation of prior research is the use of Black-Scholes implied volatility as a proxy for the market's expectation of future volatility. To the extent that its underlying assumptions are violated, the Black-Scholes formula will generate an implied volatility estimate that measures the option market's expectation of future volatility with error. The Black-Scholes model implicitly assumes that securities are infinitely divisible and that their prices follow a geometric Brownian motion. It also assumes that equity markets are weak-form informationally efficient. However, the large literature on equity mispricing raises questions about the degree of informational inefficiency in equity markets[2] (Lee 2001). Additionally, several studies document equity price jumps that would violate the assumed Brownian price process (Pan 2002).

Model-free implied volatility measures provide an alternative to implied volatility estimates generated from the Black-Scholes formula and other formulas that assume a functional form for the underlying asset price process. Britten-Jones and Neuberger (2000) generate an expression for the option market's expectation of future equity volatility as the area underneath the curve mapping option prices to the range of strike prices. Although Britten-Jones and Neuberger (2000) assume that asset prices follow a diffusion process in their development of model-free implied volatility, Jiang and Tian (2005) demonstrate that any asset price series that satisfies the generic properties of a martingale (including those with jumps) will lead to the Britten-Jones and Neuberger (2000) result. Therefore, unlike Black-Scholes implied volatility, model-free implied volatility does not suffer from measurement error induced by violations of distributional assumptions. Jiang and Tian (2005) show that model-free implied volatility measures subsume the information contained in the corresponding Black-Scholes implied volatility estimate, but also forecast future realized volatility with error.

Variance Risk Premia and Information Processing in the Options Market

That the difference between implied and realized volatilities persists despite innovations in the measurement of implied volatility supports two possible hypotheses. The first relates to the risk-aversion of equity market investors. The implication of equality between *ex ante* implied volatility and *ex post* realized volatility is made under the assumption of risk-neutrality. If equity investors are actually risk-averse, then it is unclear that implied volatility should exactly equal subsequent realized volatility. Implied volatility estimates are generated by assuming that risk-neutral investors demand no compensation for exposure to uncertainty about future volatility realizations. However, options prices might be set by risk-averse investors who do demand such compensation. Risk-aversion will lead implied volatility estimates to deviate from realized volatility when there is variance risk. Almost all prior studies on option pricing attribute the deviation of implied volatility from realized volatility to the existence of variance risk. Many studies have defined the difference

[2] In sensitivity analyses, I examine the extent to which my results depend upon assuming informational efficiency in equity markets. See Section V for more details.

American Accounting Association

between realized and implied volatility as the variance risk premium. However, prior research has largely overlooked the mechanism by which variance risk arises, offering little insight into the fundamental characteristics of the firm that might lead to, or at least be correlated with, increased uncertainty about future variance. I attempt to address this question by examining if the accounting-based volatility drivers from Section II can explain the difference between implied and realized volatility. This leads to my second hypothesis:

H2: The difference between implied and future realized volatility is predictable using accounting-based volatility drivers.

H2 could also arise as a test of an alternative hypothesis: that the difference between implied and realized volatilities is the result of a bias in the market's expectations of future volatility. The fundamental analysis literature documents the under-reaction of equity prices to financial statement information in forecasting future cash flows. That capital market participants could under-react in forming the expectation of future cash flow levels suggests the possibility of similar under-reaction in forming expectations of the volatility of equity returns. Moreover, in relation to that of equity markets, the informational efficiency of the options market remains relatively unexplored. Prior research reveals that the options market features a higher concentration of well-informed and sophisticated institutional traders for whom one would anticipate a low incidence of informational inefficiency (Jin, Livnat, and Zhang 2012). However, characteristics of the option market microstructure that limit trading, such as low trading volumes and high transactions costs, may counteract this effect (Pool, Stoll, and Whaley 2008; Roll, Schwartz, and Subrahmanyam 2010).

The literature on option pricing also suggests the possibility of informational inefficiency in the options market. Prior research documents the predictability of option returns using both historical option prices and accounting-based fundamentals. Goyal and Saretto (2009) demonstrate that the difference between historical and implied volatility positively predicts long straddle portfolio returns. Straddles are constructed of a single call and put on the same underlying asset. Straddle returns are increasing in extreme stock price movement. Goodman, Neamtiu, and Zhang (2014) demonstrate that, after controlling for implied volatility, the residual expected absolute equity return implied by a set of fundamentals is also positively associated with long straddle returns. One interpretation of these studies is that they hint at the possibility of informational inefficiency. An equally plausible alternative is that they provide evidence of variance risk premia in options prices. To date, there has been no study of informational efficiency in straddle prices that directly addresses the pricing implications of variance risk. To consider this missing piece of the literature, I test a third hypothesis regarding information processing in the options market:

H3: Options markets do not fully capture the information in financial statements in forecasting future volatility.

In my second and third hypotheses, I focus on option-implied volatility as the leading forecast of future equity volatility. This focus stems from the outperformance of option-implied volatility relative to other categories of volatility forecasts. However, option-implied volatility forecasts are only available for firms with actively traded equity options, which is a small subset of the universe of firms with publicly traded equity. For firms without option-implied volatility forecasts, investors must rely solely on past realized volatility as a predictor of future volatility. By identifying accounting-based volatility drivers, this article offers a new class of information for use in volatility prediction. Unlike option-implied volatility, accounting-based volatility drivers can be calculated for every publicly traded firm, as they are based only on information from the four main financial statements. I hypothesize that incorporating these accounting-based volatility drivers into a volatility forecasting model will generate substantial improvements in forecast errors relative to a model based solely on historical realized volatility. This hypothesis is stated below in alternative form:

H4: Volatility forecasts formed using accounting-based volatility drivers outperform forecasts based on historical realized volatility.

III. RESEARCH DESIGN

My empirical analysis consists of four stages. I test my first hypothesis by examining the relations between each of my accounting-based volatility drivers and future realized volatility. To test my second hypothesis, I examine the significance of each driver incremental to model-free option-implied volatility in the prediction of future realized volatility using regression framework.[3] I also demonstrate that a trading strategy based on the accounting-based volatility drivers can generate positive returns on investments in option straddle portfolios. To test my third hypothesis, I examine the predictability of straddle returns using accounting-based volatility drivers after controlling for variance risk exposure. Finally, I test my last hypothesis by exploring how accounting-based volatility forecasts perform out of sample, relative to forecasts based on historical volatility.

Testing H1

H1 predicts that there are three accounting-based drivers (*ABD*) of volatility: standard deviation of the earnings yield, standard deviation of the change in premium of market value over book value, and the covariance of the two.[4] I test this hypothesis by estimating the following equation for each driver:

$$\sigma^{RV}_{i,t,\tau} = \alpha + \beta_j ABD^j_{i,t-1} + \gamma Year + \delta Industry + \varepsilon_{i,t} \tag{2}$$

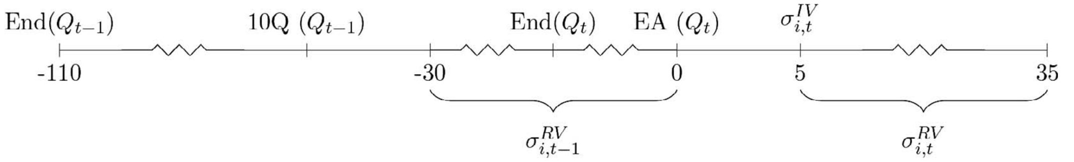
In Equation (2), $\sigma^{RV}_{i,t,\tau}$ is the logarithm of the sum of squared five-minute returns for firm i over the $\tau = 30$ days prior to the earnings announcement date for quarter t. In estimating each accounting-based volatility driver $(ABD^j_{i,t-1})$, I ensure that all information used is available to market participants five days after the quarter t earnings announcement date. At that point in time, the market will have access to quarter t earnings, but will only have equity book values as of the end of quarter t–1. For this reason, I use lagged accounting data in my estimation of each accounting-based driver. Figure 1 provides a timeline of events for firm i in quarter t. Day 0 on the timeline refers to the announcement date of quarter t earnings. Implied volatility is measured at day 5. Future realized volatility is measured as the logarithm of the sum of squared five-minute returns for firm i over the 30 days starting five days after the earnings announcement date for quarter t. To ensure that there is no peek-ahead bias in my estimation, all fundamental variables are calculated using accounting information released in the 10-Q for quarter t–1 or earlier. As Figure 1 shows, this information is typically released prior to day −30, well in advance of the quarter t earnings announcement date. This ensures that there has been sufficient time for options prices to incorporate the information. I identify announcement dates by using the earlier of the I/B/E/S and Compustat announcement dates. If one database does not report an announcement date, but the other does, then I use the date available. If both I/B/E/S and Compustat are missing announcement dates, then I

[3] My study is the first to estimate model-free implied volatility at the firm level. To provide continuity with prior research, I also explore the use of Black-Scholes implied volatility in untabulated analyses. The use of Black-Scholes implied volatility provides further support for my conclusions.

[4] I use standard deviations rather than variances of each accounting-based driver in my empirical estimation since my dependent variable is realized volatility rather than realized variance.

American Accounting Association

FIGURE 1
Timeline of Events



This figure presents the timeline of events for each firm-quarter of my sample. *EA*(Q_t) denotes the earnings announcement date for quarter *t*. *10Q*(Q_{t-1}) denotes the 10-Q release date for quarter *t*−1. $\sigma^{RV}_{i,t-1}$ is the sum of squared five-minute returns over the 30-day window starting 30 days prior to the quarter *t* earnings announcement. $\sigma^{RV}_{i,t}$ is the sum of squared five-minute returns over the 30-day window starting five days after the quarter *t* earnings announcement date. $\sigma^{IV}_{i,t}$ is the model-free implied volatility inferred from closing option prices on the fifth day after quarter *t* earnings announcement date.

eliminate the observation from my sample. Following Barth and So (2014), I adjust the announcement date one trading day forward when the announcement occurs after the market close.

To estimate the three accounting-based drivers identified in H1, I first construct for each firm in my sample a monthly time-series of two variables: earnings yield and change in market to book premium. I define earnings yield as the ratio of quarterly earnings to the closing equity price on the last trading date of each month. I ensure that the earnings numbers used in the earnings yield are publicly available at least one week prior to the last trading day of the month (i.e., it should be incorporated into the closing equity price). I similarly calculate the change in market to book premium to ensure that there is no peek-ahead bias. Market value is measured on the last trading day of each month, and I match each market value to the most recently reported quarterly book value of equity figure that is publicly available at least one week prior to the last trading day of the month. I scale this change by the closing stock price on the last trading day of the month.

For each firm-quarter, I estimate earnings yield volatility as the standard deviation of the monthly earnings yield over the ten months prior to the quarter *t* announcement date. I estimate the volatility of the change in market to book premium as the standard deviation of the monthly change in market to book premium over the ten months prior to the quarter *t* announcement date. I also estimate the covariance of the monthly earnings yield and change in market to book premium for same ten months. These three variables constitute the three accounting-based drivers from H1. For each accounting-based driver of volatility, an estimate of β_j that is significantly different from zero supports my first hypothesis that the particular variable is significantly correlated with future returns volatility. In estimating Equation (2), I also include industry and year fixed effects. Industries are defined using the Fama and French (1997) 48 industry classifications. Because my sample exhibits time-series and cross-sectional correlation, I use two-way firm and quarter clustered standard errors when testing coefficient significance (Gow, Ormazabal, and Taylor 2010).

Testing H2

To test my second hypothesis, I estimate Equation (3) below:

$$\sigma^{RV}_{i,t,\tau} = \alpha + \beta_1\sigma^{RV}_{i,t,-\tau} + \beta_2\sigma^{MFIV}_{i,t,\tau} + \beta_j ABD^j_{i,t-1} + \beta_4 Spread + \gamma Year + \delta Industry + \varepsilon_{i,t} \quad (3)$$

In Equation (3), $\sigma^{MFIV}_{i,t,\tau}$ is the logarithm of model-free implied volatility of options on firm *i*'s equity measured five days after quarter *t*'s earnings announcement date with $\tau = 30$ days remaining until expiration. Measuring implied volatility five days after the quarterly earnings announcement date helps avoid capturing announcement-induced volatility in my measurement of these amounts. I limit my analysis to options with 30 days remaining until expiration to avoid including multiple earnings announcements in the volatility measurement period.[5]

I estimate implied volatility using the model-free method derived by Britten-Jones and Neuberger (2000):

$$\sigma_T^{MFIV} = E_0^F\left[\int_0^T \left(\frac{dF_t}{F_t}\right)^2\right] = \int_{-\infty}^{0} \frac{P^F(T,K) - \max\{0, F_0 - K\}}{K^2} dK + \int_0^{\infty} \frac{C^F(T,K) - \max\{0, F_0 - K\}}{K^2} dK \quad (4)$$

In Equation (4), F_t is the forward price of the underlying asset at time *t*, and $P^F(T, K)$ ($C^F[T, K]$) is the price of a put (call) option with strike price *K* and remaining time to maturity *t*. Appendix A provides a replication of the Breeden and Litzenberger (1978) result underlying this model-free implied volatility estimate and further details on the derivation of the above approximation. Equation (4) requires option price observations for the entire continuum of strike prices, but regulations on most options exchanges prevent the trading of options with very high or low strikes. This truncation is the largest source of error in model-free implied volatility estimates, but Jiang and Tian (2007) show that the truncation error becomes negligible if the range of available strikes used is at least two standard deviations around the current underlying asset value. For each firm-date on which I estimate model-free implied volatility, I require that firms have at least five traded options (two out of the money, one at the money, and two in the money). These five traded options must also satisfy basic no-arbitrage conditions (the bid price must be strictly positive and less than the ask price, the bid-ask spread must exceed the minimum tick size, and there must be non-zero open interest in the option).

In Equation (3), I include controls for past volatility, $\sigma_{i,t,-\tau}^{RV}$, and liquidity, *Spread*. I measure past volatility for firm *i* in quarter *t* as the sum of squared five-minute returns for firm *i* over the 30 days prior to the earnings announcement date for quarter *t*. *Spread* is the logarithm of the median volume-weighted bid-ask spread for all options on firm *i*'s equity over the year ending on the relevant earnings announcement date. I include these variables as controls because the extensive literature on volatility forecasting finds that past volatility is informative of future volatility and that liquidity and volatility are significantly inversely related (Christoffersen, Goyenko, Jacobs, and Karoui 2011). In estimating Equation (3), I also include industry and year fixed effects, defined identically as in Equation (2). Equation (3) allows me to test the incremental informativeness of each accounting-based volatility driver after controlling for implied volatility. My second hypothesis predicts that the coefficients β_j will be non-zero.

H2 also implies the possibility of using accounting-based volatility drivers to predict option returns. I test this implication by measuring the returns to holding a long straddle portfolio for each firm-quarter. Long straddle portfolios consist of a single at-the-money call and put option on the same underlying asset and have the following payoff function:

$$V^{Straddle} = |S - K| - P \quad (5)$$

where *S* is the value of the underlying stock, *K* is the strike price of the call and put options in the straddle, and *P* is the purchase price of the straddle. I measure straddle returns using as a purchase price the closing prices of one put and one call option five days after the quarterly earnings announcement date. I measure the straddle payoff using Equation (5) and the closing price of the

[5] Untabulated results show that using options with 60 days until expiration does not qualitatively change my conclusions.

underlying security on the day of expiration. I only consider options with moneyness between 0.975 and 1.025 to mitigate concerns about pricing anomalies associated with the volatility smile (Hull 2009). I also eliminate observations with bid prices equal to zero or bid prices greater than ask prices to minimize potential recording errors (Goyal and Saretto 2009).

For each firm quarter, I pick the option pair closest to being at-the-money. I first restrict my analysis to options with 30 days until expiration to ensure consistency with my main analyses, because variation in the time horizon of volatility estimation could affect the informativeness of the financial statement variables. I estimate straddle returns under the assumption that investors always buy options at the ask price and write options at the bid price. In other words, I assume that the profitability of each transaction is lowered by the amount of the full bid-ask spread as reported by OptionMetrics. Prior research reveals that closing bid-ask spreads on the OptionMetrics database are often larger than the effective spreads investors actually face (Mayhew 2002; De Fontnouvelle, Fishe, and Harris 2003; Battalio, Hatch, and Jennings 2004). Consequently, using OptionMetrics' bid-ask spread estimates provides me with a conservative estimate of the transactions costs that options investors face. I use the closing bid-ask spread to proxy for the average bid-ask spread, since intraday data on traded options are not available through OptionMetrics.

My second hypothesis suggests that the profitability of using accounting-based volatility drivers to predict straddle returns will not be completely subsumed by the profitability of using model-free implied volatility and historical volatility. To test this, I generate volatility forecasts using the following equations:

$$\sigma^{RV}_{i,t,\tau} = \alpha_1 + \beta_1 \sigma^{MFIV}_{i,t,\tau} + \beta_2 \sigma^{RV}_{i,t,-\tau} + \varepsilon_{i,t} \quad (6)$$

$$\sigma^{RV}_{i,t,\tau} = \alpha_2 + \beta_3 \sigma^{MFIV}_{i,t,\tau} + \beta_4 \sigma^{RV}_{i,t,-\tau} + \beta_j ABD^{j}_{i,t-1} + \varepsilon_{i,t} \quad (7)$$

I estimate the parameters α_1, α_2, β_1, β_2, β_3, β_4, and β_j in Equations (6) and (7) for each quarter in my sample using only data from that quarter and the previous quarter. Specifically, I use quarter t–1 and earlier accounting data for firm i to calculate the accounting based drivers. $\sigma^{MFIV}_{i,t,\tau}$ is model-free implied volatility measured five days after the earnings announcement for quarter t using the prices of options with 30 days to maturity on firm i's equity, and $\sigma^{RV}_{i,t,-\tau}$ is the sum of squared five-minute returns for firm i over the 30 days prior to the earnings announcement date for quarter t. This structure avoids the possibility of any peek-ahead bias in the construction of the expected volatility estimates. I use the parameter estimates generated from quarter t data to calculate expected values for future realized volatility in quarter t+1. This process generates two different expectations of future realized volatility for each quarter, which I label *ERV_NoABD* and *ERV_ABD:*

$$ERV\_NoABD = \hat{\alpha}_1 + \widehat{\beta}_1 \sigma^{MFIV}_{i,t,\tau} + \widehat{\beta}_2 \sigma^{RV}_{i,t,-\tau} + \varepsilon_{i,t}$$
$$ERV\_ABD = \hat{\alpha}_2 + \widehat{\beta}_3 \sigma^{MFIV}_{i,t,\tau} + \widehat{\beta}_4 \sigma^{RV}_{i,t,-\tau} + \widehat{\beta}_j ABD^{j}_{i,t-1} + \varepsilon_{i,t}$$

The difference between *ERV_NoABD* and *ERV_ABD* arises from the incremental informativeness of the accounting-based drivers (*ABD*). I use both *ERV_NoABD* and *ERV_ABD* as partitioning variables in the construction of straddle hedge portfolios and examine the relative performance of the two portfolio classes. Based upon my second hypothesis, I predict that the portfolio constructed with *ERV_ABD* as a partitioning variable will exhibit higher returns than the portfolio constructed with *ERV_NoABD* as a partitioning variable.

Testing H3

My third hypothesis examines whether the findings established in the course of testing H2 can be attributed to insufficient information processing in the options market. To test this, I must control for the effect of correlation between accounting-based volatility and variance risk. I accomplish this by constructing a direct measure of variance risk and including it as an additional control variable in straddle returns regressions. Carr and Wu (2009) outline a process for estimating a firm's variance risk analog to the capital asset pricing model (CAPM) beta. Specifically, they define "variance beta" as:

$$VarBeta = \frac{Cov(\log RV_j, \log RV_{SPX})}{Var(\log RV_{SPX})}$$

where RV_j is the realized variance of firm j's equity returns and RV_{SPX} is the realized variance of the S&P 500 index's returns. This variance beta measure captures the extent to which the variance of a firm's equity returns is correlated with the variance of market returns. A high *VarBeta* indicates that the two are highly correlated and investors should demand higher premiums for being exposed to such variance risk. Consequently, the option returns for firms with high variance beta should be higher than for those with low variance beta.

I use quarterly returns regressions to explore the incremental informativeness of each accounting-based driver for assessing straddle returns after controlling for each firm's variance risk beta. Specifically, I estimate:

$$r^s_{i,t} = \alpha + \beta_1 Rank\sigma^{MFIV}_{i,t,\tau} + \beta_2 Rank\sigma^{RV}_{i,t,-\tau} + \beta_j ABD_{i,t-1} + \beta_4 VarBeta_{i,t} + \varepsilon_{i,t} \quad (8)$$

where $r^s_{i,t}$ is the 30-day straddle return for firm i in quarter t. All other variables are as previously defined. I use the Fama and MacBeth (1973) approach to addressing cross-sectional correlation in returns by first estimating Equation (8) quarterly and then averaging coefficients and estimating standard errors from the coefficient distribution before evaluating statistical significance. To account for time-series correlation in straddle returns, I employ the Newey-West correction technique with four lags. H3 predicts that the coefficients β_j from Equation (8) are significantly non-zero.

Testing H4

My fourth hypothesis examines whether accounting-based drivers help predict future equity volatility on an out-of-sample basis. I test this by estimating the following regressions over ten-quarter windows (up to and including quarter t–1) to generate coefficient values:

Base model: $\sigma^{RV}_{i,t,\tau} = \alpha_1 + \beta_t \sigma^{RV}_{i,t,-\tau} + \varepsilon_{i,t}$
ABD model: $\sigma^{RV}_{i,t,\tau} = \alpha_2 + \delta_t \sigma^{RV}_{i,t,-\tau} + \zeta^j_t ABD^j_{i,t-1} + \varepsilon_{i,t}$

All coefficient definitions are identical to those in Equation (3), above. To best approximate the extant forecasting techniques available for firms without active options markets, the base model uses only past realized volatility as a predictor and the accounting-based model uses past realized volatility, as well as the three accounting-based drivers of equity volatility.

I use the coefficients generated from the above rolling regressions to generate predicted volatilities by applying the resulting coefficient estimates to quarter t values:

Base model: $E[\sigma^{RV}_{i,t,\tau}]_{BASE} = \widehat{\alpha_1} + \widehat{\beta}_t \sigma^{RV}_{i,t,-\tau}$
ABD model: $E[\sigma^{RV}_{i,t,\tau}]_{ABD} = \widehat{\alpha_2} + \widehat{\delta}_t \sigma^{RV}_{i,t,-\tau} + \widehat{\zeta^j_t} ABD^j_{i,t-1}$

In the above equations, $E[\sigma^{RV}_{i,t,\tau}]$ is a prediction of future equity volatility for the 30 days beginning five days after the quarter t+1 announcement date. I use a mean squared error function of the form:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\sigma^{RV}_{i,t,\tau} - E\left[\sigma^{RV}_{i,t,\tau}\right]\right)^2$$

to compare the accuracy of the base and ABD predictions. Patton (2011) finds mean squared error functions to be most appropriate when evaluating predictions of volatilities calculated using high-frequency data. My hypothesis predicts that the mean squared error associated with the accounting-based driver model will be lower than the mean squared error associated with the base model.

Sample

My sample comprises all firms with standardized implied volatility data on OptionMetrics and sufficient Compustat and CRSP data available to construct my variables. OptionMetrics provides price data from 1996 to the present for all Chicago Board Options Exchange-listed options on U.S. equities. In addition to reported prices, OptionMetrics provides interpolated option prices over a range of strike prices (commonly referred to as a volatility surface). I use OptionMetrics' volatility surface data to construct estimates of model-free implied volatility. I conduct my analysis using the implied volatility of 30-day options, although untabulated results indicate that the conclusions are unaffected by the use of 60-day options. I obtain accounting data from Compustat and daily equity returns from CRSP. I require firms to have earnings announcement dates on I/B/E/S or Compustat and require ten quarters of data prior to each quarterly observation to construct variables. The resulting sample consists of 47,398 observations from 3,078 firms from 1996 to 2012.

Panel A (B) of Figure 2 provides density plots for the level (logarithm) of the implied and realized volatility sample distributions. The plots in Panel A indicate that both implied and realized volatilities are highly skewed and leptokurtic. From Panel B, it appears that both volatility series are roughly lognormal. Therefore, I conduct my analysis using the log-series of both implied and realized volatility. Table 1 provides details on the composition of my sample by industry and year. The dominant industries in my full sample are business equipment, healthcare, financial services, and wholesale and retail trade. Panel B of Table 1 reveals a general increase in the number of observations per year over time (from 665 firm-quarter observations in 1996 to 4,256 firm-quarter observations in 2011), which is consistent with the increase in options trade over the past decade. The slight decline in observations in 2012 reflects data availability constraints at the time of data collection.

Table 2 presents univariate descriptive statistics for the key variables in my full sample. The univariate statistics for the logarithms of future and past realized volatilities are very similar, which is consistent with past volatility being predictive of future volatility. The statistics in Table 2 reveal that, on average, model-free implied volatility is higher than realized volatility. The mean and median of the logarithm of model-free implied volatility are −3.98 and −3.99 and are higher than the mean and median of future realized volatility (−3.79 and −3.81). Table 3 provides Pearson and Spearman correlation coefficients for the key variables in my analysis. Consistent with the large literature on time-series volatility estimation, past and future realized volatility exhibit a Pearson (Spearman) correlation of 0.74 (0.79). Future realized volatility is also significantly positively correlated with current model-free implied volatility with a Pearson (Spearman) correlation of 0.78 (0.82).

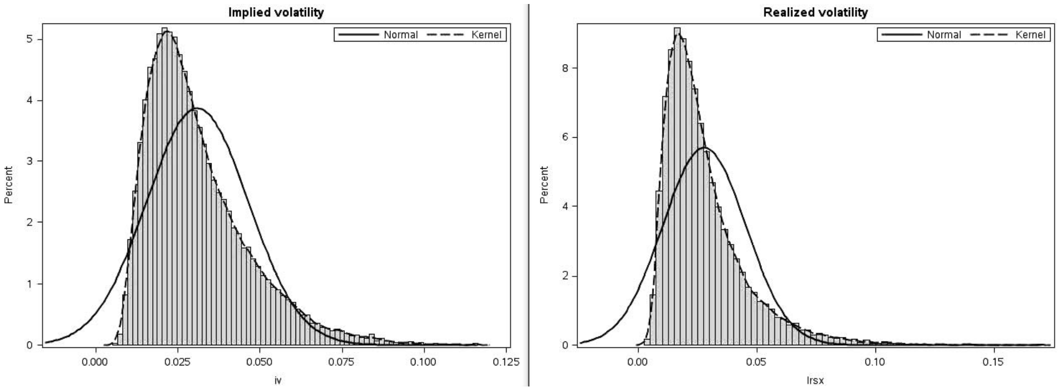
IV. EMPIRICAL RESULTS

Fundamentals and Future Equity Volatility

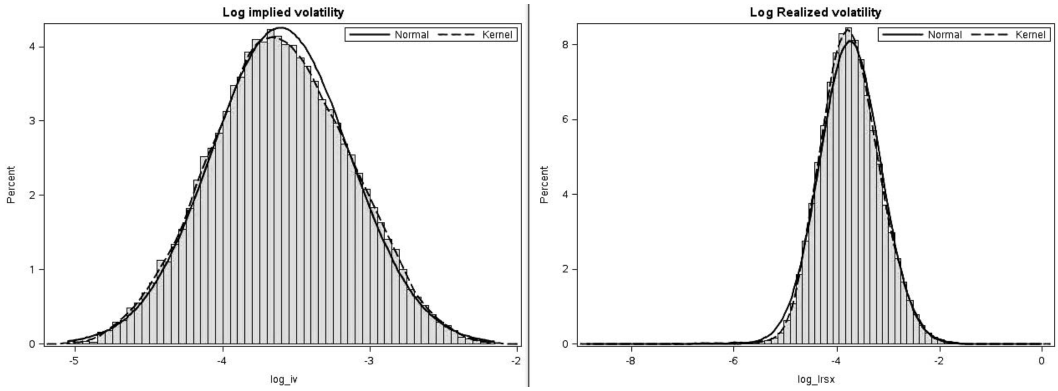
Table 4 presents summary statistics from the estimation of Equation (2), which is designed to test H1 using each of the accounting-based volatility drivers discussed in Section II. Table 4 reveals that each driver exhibits the predicted relation with future equity volatility. Equity returns volatility is significantly positively related to the earnings yield volatility (coefficient = 0.035, t-statistic = 21.553) and the volatility of the change in market to book premium (coefficient = 0.018, t-statistic = 6.769). Volatility is significantly negatively related to the covariance of the earnings yield and

FIGURE 2
Density Plots of Implied and Realized Volatility

Panel A: Density Plots for the Level of Implied and Realized Volatility



Panel B: Density Plots for the Logarithm of Implied and Realized Volatility



change in market to book premium (coefficient = −0.016, t-statistic = −7.468). Column IV of Table 4 confirms that these relations persist when all three accounting-based drivers are included in a single regression, suggesting that they have incremental explanatory power to one another. The explanatory power of the regressions in Table 4 varies from 28 percent to 32 percent. Overall, the results from Table 4 support my first hypothesis that the three accounting-based drivers (earnings yield volatility, volatility of the change in market to book premium, and the covariance of the earnings yield and change in market to book premium) are associated with future equity volatility.

Fundamentals and the Difference between Option-Implied Volatility and Future Realized Volatility

Table 5 presents coefficient estimates from Equation (3), which is designed to test my second hypothesis. H2 predicts that β_3 will be significantly different from 0. The results in Table 5 support this prediction. Columns I–III confirm the findings of prior research in demonstrating that model-

TABLE 1
Sample Composition

Panel A: Sample Composition by Industry

Industry	Observations	Percent
Consumer Nondurables	2,039	4.30
Consumer Durables	1,066	2.25
Manufacturing	5,358	11.31
Energy	2,945	6.21
Chemicals	1,418	2.99
Business Equipment	10,926	23.05
Telecommunications	932	1.97
Utilities	885	1.87
Wholesale and Retail Trade	5,429	11.45
Healthcare	5,469	11.54
Financial Services	5,651	11.92
Other	5,280	11.14
Total	47,398	100.00

Panel B: Sample Composition by Year

Year	Observations	Percent
1996	665	1.40
1997	1,256	2.65
1998	1,617	3.41
1999	1,966	4.15
2000	2,120	4.47
2001	2,186	4.61
2002	2,194	4.63
2003	2,389	5.04
2004	2,887	6.09
2005	3,019	6.37
2006	3,248	6.85
2007	3,544	7.48
2008	4,130	8.71
2009	4,093	8.64
2010	4,166	8.79
2011	4,256	8.98
2012	3,662	7.73
Total	47,398	100.00

This table provides information on the composition of my sample. Panel A (Panel B) provides a breakdown of my sample by industry (year). The column "Observations" indicates the number of firm-quarter observations for each industry; "Percent" indicates the percentage of the total sample attributable to each industry. Industries are defined using the 12 Fama-French classifications.

TABLE 2

Descriptive Statistics

	n	Mean	Std. Dev.	Min.	Q1	Med.	Q3	Max.
$\sigma_{i,t,\tau}^{RV}$	47,398	−3.96	0.51	−5.06	−4.32	−3.99	−3.61	3.58
$\sigma_{i,t,-\tau}^{RV}$	47,398	−3.79	0.5	−4.93	−4.15	−3.81	−3.45	3.85
$\sigma_{i,t,\tau}^{MFIV}$	47,398	−3.98	0.41	−4.95	−4.27	−3.99	−3.69	3.14
$\sigma_{i,t}^{EY}$	47,398	1.1	2.76	0.03	0.28	0.53	1.1	60.73
$\sigma_{i,t}^{CMMB}$	47,398	0.86	1.21	0.05	0.19	0.4	0.96	9.63
$\sigma(EY, CMMB)_{it}$	47,398	−1.68	5.99	−73.88	−0.58	−0.08	−0.01	0
$I\left\{\sigma_{i,t,\tau}^{RV} < \sigma_{i,t,\tau}^{MFIV}\right\}$	47,398	0.51	0.5	0	0	1	1	1

This table provides univariate statistics for the main variables in my analysis.

Variable Definitions:
$\sigma_{i,t,-\tau}^{RV}$ = sum of squared five-minute returns on firm *i*'s equity over the 30-day window starting 30 days prior to the quarter *t* earnings announcement;
$\sigma_{i,t,\tau}^{RV}$ = sum of squared five-minute returns on firm *i*'s equity over the 30-day window starting five days after the quarter *t* earnings announcement date;
$\sigma_{i,t,\tau}^{MFIV}$ = model-free implied volatility inferred from closing option prices on firm *i*'s equity on the fifth day after the quarter *t* earnings announcement date;
$\sigma_{i,t}^{EY}$ = standard deviation of earnings yield for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;
$\sigma_{i,t}^{CMMB}$ = standard deviation of the change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;
$\sigma(EY, CMMB)_{it}$ = covariance of earnings yield and change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date; and
$I\left\{\sigma_{i,t,\tau}^{RV} < \sigma_{i,t,\tau}^{MFIV}\right\}$ = an indicator equaling 1 if $\sigma_{i,t,\tau}^{RV}$ is less than $\sigma_{i,t,\tau}^{MFIV}$ for firm *i* in quarter *t*.

free implied volatility outperforms historical volatility as a predictor of future volatility. Estimates of β_2, the coefficient on model-free implied volatility, range from 0.720 to 0.736 across all of the models and are consistently higher than the coefficients on past realized volatility.

Column III of Table 5 reveals that the addition of volume-weighted equity market spread as a liquidity proxy does not significantly affect the equation's explanatory power. Overall, explained variation is unchanged by the inclusion of the variable *Spread*. *Spread* has a significantly negative coefficient of −0.006 (t-statistic = −0.814) that reflects the inverse relation between option market liquidity and prices. Columns IV to VIII provide the regression coefficients from the estimation of Equation (3) in full. Each of the variables I examine has a coefficient (β_3) that is significantly non-zero. Earnings yield volatility has a coefficient of 0.051 and the volatility of change in market to book premium has a coefficient of −0.032 (t-statistic = −3.450). Overall, the significant non-zero β_3 estimates for these variables suggest that these variables are useful in predicting the difference between implied and realized volatilities.[6] One potential interpretation of these findings is that earnings yield volatility and change in market to book premium volatility are useful indicators of variance risk.

The persistent significance of accounting-based volatility drivers, incremental to option-implied volatility, in the prediction of future realized volatility suggests that these accounting-based

[6] Underlying this inference is the assumption that the options market uses market price as the price of equity in determining the price of the option. I relax this assumption in Section V.

TABLE 3

Correlation Matrices

	$\sigma^{MFIV}_{i,t,\tau}$	$\sigma^{RV}_{i,t,-\tau}$	$\sigma^{RV}_{i,t,\tau}$	$\sigma^{EY}_{i,t}$	$\sigma^{CMMB}_{i,t}$	$\sigma(EY, CMMB)_{it}$	$Spread_{it}$	$VolSpread_{it}$
$\sigma^{MFIV}_{i,t,\tau}$	1	0.84	0.78	0.31	0.24	−0.31	0.03	−0.02
$\sigma^{RV}_{i,t,-\tau}$	0.86	1	0.74	0.29	0.23	−0.31	−0.01	−0.01
$\sigma^{RV}_{i,t,\tau}$	0.82	0.79	1	0.28	0.23	−0.31	0	−0.01
$\sigma^{EY}_{i,t}$	0.34	0.28	0.28	1	0.03	0	0	−0.02
$\sigma^{CMMB}_{i,t}$	0.18	0.18	0.18	0.18	1	0.39	0.25	0.02
$\sigma(EY, CMMB)_{it}$	−0.31	−0.31	−0.29	−0.22	−0.66	1	−0.08	0
$Spread_{it}$	0.05	0.02	0.04	0.04	−0.42	0.33	1	−0.01
$VolSpread_{it}$	0.03	0.02	0.03	−0.02	0.02	−0.02	0	1

This table presents correlation matrices for the main variables in my analysis. Pearson (Spearman) correlations are above (below) the diagonal.

Variable Definitions:
$\sigma^{RV}_{i,t,-\tau}$ = sum of squared five-minute returns on firm *i*'s equity over the 30-day window starting 30 days prior to the quarter *t* earnings announcement;
$\sigma^{RV}_{i,t,\tau}$ = sum of squared five-minute returns on firm *i*'s equity over the 30-day window starting five days after the quarter *t* earnings announcement date;
$\sigma^{MFIV}_{i,t,\tau}$ = model-free implied volatility inferred from closing option prices on firm *i*'s equity on the fifth day after the quarter *t* earnings announcement date;
$\sigma^{EY}_{i,t}$ = standard deviation of earnings yield for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;
$\sigma^{CMMB}_{i,t}$ = standard deviation of the change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;
$\sigma(EY, CMMB)_{it}$ = covariance of earnings yield and change in market to book premium firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;
$Spread_{it}$ = logarithm of the median volume-weighted bid-ask spread for all options on firm *i*'s equity over the year ending on the quarter *t* earnings announcement date; and
$VolSpread_{it}$ = open interest-weighted average difference in implied volatilities between call options and put options (with the same strike price and maturity) on firm *i*'s equity measured on the quarter *t* earnings announcement date.

drivers could be useful in the prediction of option returns. I explore this implication by examining the returns to straddle portfolios by quintile of two different expected future volatility estimates: *ERV_NoABD* and *ERV_ABD*. Differences between *ERV_NoABD* and *ERV_ABD* arise from the incremental informativeness of the variable. My second hypothesis suggests that hedge returns to the straddle portfolio constructed with *ERV_ABD* as a partitioning variable will exhibit higher returns than the portfolio constructed with *ERV_NoABD* as a partitioning variable. I use both *ERV_NoABD* and *ERV_ABD* as partitioning variables in the construction of straddle hedge portfolios and examine the relative performance of the two portfolio classes.

Figure 3 presents the mean returns by quintile of *ERV_NoABD* and *ERV_ABD*. The solid columns in Figure 3 represent the returns to straddle portfolios partitioned by quintile of *ERV_ABD*. Straddle portfolios in the lowest quintile of *ERV_ABD* earn −1.52 percent return over the sample period, while straddle returns in the highest quintile earn 1.74 percent over the same period. In relation to this, the mean returns earned by portfolios partitioned on *ERV_NoABD* (shown in Figure 3 in the striped columns) is lower in the highest quintile (mean = 0.52 percent) and higher in the lowest quintile (−1.20 percent). This suggests that *ERV_ABD* is helpful to investors, particularly in identifying instances of more extreme future volatility or lack thereof. Particularly, the disparity in mean returns in the highest quintile of *ERV_ABD* and *ERV_NoABD* reveals that *ERV_ABD*'s usefulness to investors may be in identifying profitable straddle purchasing opportunities. The

TABLE 4

Regressions of Realized Volatility on Accounting-Based Volatility Drivers

$$\sigma^{RV}_{i,t,\tau} = \alpha + \beta_j ABD^j_{i,t-1} + \gamma Year + \delta Industry + \varepsilon_{i,t}$$

	Pred.	I	II	III	IV
$\sigma^{EY}_{i,t}$	+	0.035***			0.036***
		(21.553)			(21.907)
$\sigma^{CMMB}_{i,t}$	+		0.018***		0.018***
			(6.769)		(8.499)
$\sigma(EY, CMMB)_{it}$	−			−0.016***	−0.004***
				(−7.468)	(−2.796)
Obs.		47,398	47,398	47,398	47,398
Adj. R^2		0.31	0.29	0.28	0.32
Year FE		Yes	Yes	Yes	Yes
Industry FE		Yes	Yes	Yes	Yes

*, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively, of the hypothesis that the given coefficient is different from zero.

This table provides the coefficient estimates from Equation (3), above, using each of the fundamental indicator variables discussed in Section II. In parentheses below each coefficient are t-statistics from two-way firm and quarter clustered standard errors.

Variable Definitions:

$\sigma^{RV}_{i,t,\tau}$ = sum of squared five-minute returns on firm *i*'s equity over the 30-day window starting five days after the quarter *t* earnings announcement date;

$\sigma^{EY}_{i,t}$ = standard deviation of earnings yield for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;

$\sigma^{CMMB}_{i,t}$ = standard deviation of the change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date; and

$\sigma(EY, CMMB)_{it}$ = covariance of earnings yield and change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date.

hedge return to a strategy of taking a long position in straddles for firms with high *ERV_ABD* and a short position in straddles for firms with low *ERV_ABD* is 3.26 percent over the sample period. In comparison, the hedge return to an analogous strategy using *ERV_NoABD* is 1.72 percent. Both returns are significantly different from zero. Overall, the results in Figure 3 confirm the prediction that using accounting-based information can generate profitable trading opportunities in the options market beyond what is available through market-based data alone.

Variance Risk Beta

Table 6 presents coefficient estimates from Equation (8), which is designed to test my third hypothesis. H3 predicts that the coefficients β_3 should be significantly different from 0. The results in Table 6 support this prediction. Column IV of Table 6 provides construct validity by showing that variance beta exhibits a significantly positive relation to option straddle returns (coefficient = 0.096, t-statistic = 5.02). Having established this, I then proceed to estimate Equation (8) using all three accounting-based volatility drivers. The results in Table 6 provide mixed support for the hypothesis by showing that the coefficients on earnings yield volatility and the volatility of change in market to book premium are both significantly different from 0 after controlling for variance risk. Earnings yield volatility has a coefficient of 0.087 (t-statistic = 2.45) and the change in market to book premium volatility has a coefficient of 0.023 (t-statistic = 2.01). These results suggest that the

TABLE 5

Regressions of Realized Volatility on Implied Volatility and Accounting-Based Volatility Drivers

$$\sigma^{RV}_{i,t,\tau} = \alpha + \beta_1\sigma^{RV}_{i,t,-\tau} + \beta_2\sigma^{MFIV}_{i,t,\tau} + \beta_j ABD^j_{i,t-1} + \beta_4 Spread + \gamma Year + \delta Industry + \varepsilon_{i,t}$$

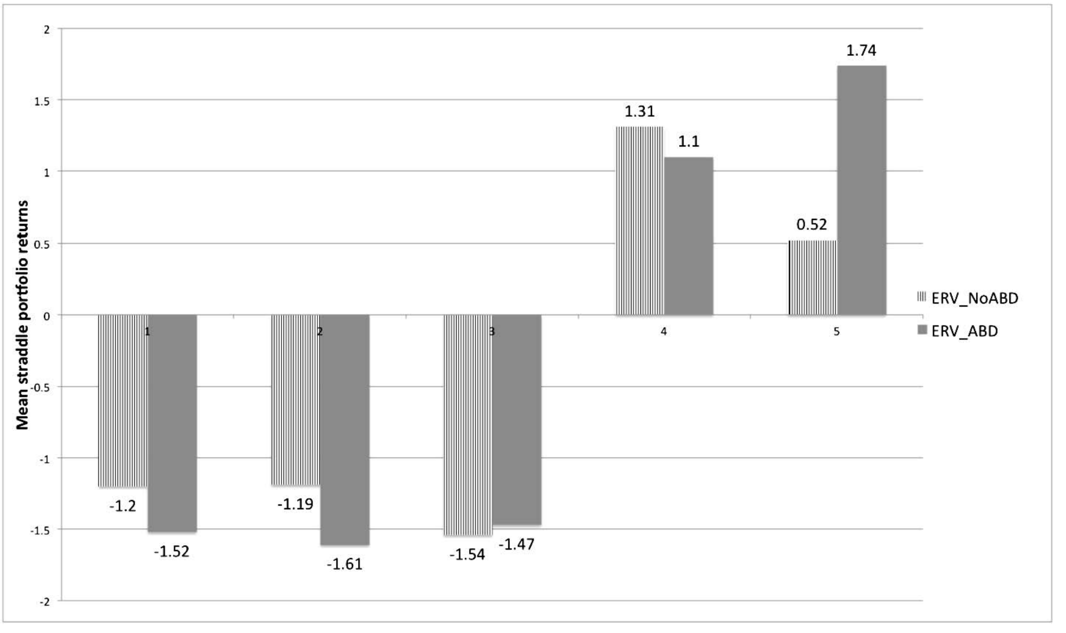
	I	II	III	IV	V	VI	VII
$\sigma^{RV}_{i,t,-\tau}$	0.616$^+$	0.179$^+$	0.177$^+$	0.176$^+$	0.178$^+$	0.177$^+$	0.176$^+$
	(−26.108)	(−26.108)	(−52.852)	(−53.143)	(−52.898)	(−52.802)	(−53.172)
$\sigma^{MFIV}_{i,t,\tau}$		0.732$^+$	0.735$^+$	0.720$^+$	0.736$^+$	0.734$^+$	0.721$^+$
		(−12.352)	(−12.544)	(−12.883)	(−12.489)	(−12.581)	(−12.765)
Spread			−0.006	−0.006	−0.013	−0.010	−0.014*
			(−0.814)	(−0.793)	(−1.634)	(−1.340)	(−1.785)
$\sigma^{EY}_{i,t}$				0.051***			0.042***
				(4.550)			(4.837)
$\sigma^{CMMB}_{i,t}$					−0.032***		−0.031***
					(−3.450)		(−3.361)
$\sigma(EY, CMMB)_{it}$						0.035***	0.011
						(3.308)	(1.235)
Obs.	47,398	47,398	47,398	47,398	47,398	47,398	47,398
Adj. R^2	0.58	0.60	0.60	0.64	0.62	0.61	0.65
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively, of the hypothesis that the given coefficient is different from zero.
$^+$ Indicates significance at the 1 percent level of the hypothesis that the given coefficient is different from 1.
This table provides the coefficient estimates from Equation (3), above, using each of the fundamental indicator variables discussed in Section II and model-free implied volatility. In parentheses below each coefficient are t-statistics from two-way firm and quarter clustered standard errors.

Variable Definitions:
$\sigma^{RV}_{i,t,\tau}$ = sum of squared five-minute returns on firm *i*'s equity over the 30-day window starting five days after the quarter *t* earnings announcement date;
$\sigma^{MFIV}_{i,t,\tau}$ = model-free implied volatility inferred from closing option prices on firm *i*'s equity on the fifth day after the quarter *t* earnings announcement date;
$\sigma^{EY}_{i,t}$ = standard deviation of earnings yield for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;
$\sigma^{CMMB}_{i,t}$ = standard deviation of the change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;
$\sigma(EY, CMMB)_{it}$ = covariance of earnings yield and change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date; and
$Spread_{it}$ = logarithm of the median volume-weighted bid-ask spread for all options on firm *i*'s equity over the year ending on the quarter *t* earnings announcement date.

FIGURE 3
Hedge Straddle Returns to an Estimated Recovery Value (ERV) Strategy

Panel A: Average Straddle Returns by ERV Quintile



Panel B: Average Hedge Returns to ERV Strategy

Partitioning Variable	Hedge (Q5 – Q1) Returns
ERV_ABD	3.26%***
ERV_NoABD	1.72***

*** Indicates significance at the 1 percent level of the hypothesis that the given statistic is different from zero.
Panel A provides average 30-day returns from a long straddle portfolio averaged within quintiles of *ERV_ABD* or *ERV_NoABD* (see Appendix B for variable definitions). The long straddle portfolio is formed by taking a long position in both an at-the-money (ATM) call and ATM put option of equivalent maturity. All portfolios are held to maturity in order to reduce biases due to limited trading. Returns are calculated as the change in portfolio value over the period as a percentage of initial investment costs, assuming a 100 percent effective spread as an approximation of transactions costs. Panel B provides average hedge returns earned by each strategy.

predictive ability that earnings yield volatility and change in market to book premium volatility have over the difference in implied and realized volatilities is not entirely driven by their correlation with variance risk. However, the same is not true for the covariance of earnings yield with changes in market to book premium. That variable has a coefficient of 0.002 (t-statistic = 1.15), which is not distinguishable from 0 after controlling for variance risk. This suggests that the earlier observed relations between this variable and the difference in implied and realized volatility is driven by the relation this variable has with variance risk.

TABLE 6

Regressions of Straddle Returns on Fundamental Variables and Variance Risk Factor

$$r^s_{i,t} = \alpha + \beta_1 Rank\sigma^{MFIV}_{i,t,\tau} + \beta_2 Rank\sigma^{RV}_{i,t,-\tau} + \beta_j ABD_{i,t-1} + \beta_4 VarBeta_{i,t} + \varepsilon_{i,t}$$

	I	II	III	IV
$\sigma^{MFIV}_{i,t,\tau}$	0.040*	0.077***	0.067***	
	(1.69)	(3.11)	(3.00)	
$\sigma^{RV}_{i,t,-\tau}$	0.071***	0.064***	0.061***	
	(3.39)	(3.13)	(3.04)	
$\sigma^{EY}_{i,t}$		0.087***	0.083***	
		(2.45)	(2.43)	
$\sigma^{CMMB}_{i,t}$		0.031**	0.23**	
		(2.09)	(2.01)	
$\sigma(EY, CMMB)_{it}$		0.002	0.002	
		(1.06)	(1.15)	
$VarBeta_{it}$			0.021	0.096***
			(0.15)	(5.02)
Adj. R^2	0.020	0.057	0.062	0.017
No. of quarters	66	66	66	66
Obs. per quarter	710	710	710	710

*, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively, of the hypothesis that the given coefficient is different from zero.
This table provides coefficient estimates from Equation (8). The dependent variable in each model is firm-level 30-day option straddle returns. In parentheses below each coefficient are Fama-MacBeth t-statistics from Newey-West corrected standard errors.

Variable Definitions:
$\sigma^{RV}_{i,t,\tau}$ = sum of squared five-minute returns on firm i's equity over the 30-day window starting five days after the quarter t earnings announcement date;
$\sigma^{MFIV}_{i,t,\tau}$ = model-free implied volatility inferred from closing option prices on firm i's equity on the fifth day after the quarter t earnings announcement date;
$\sigma^{EY}_{i,t}$ = standard deviation of earnings yield for firm i over the most recent ten months ending prior to quarter t's earnings announcement date;
$\sigma^{CMMB}_{i,t}$ = standard deviation of the change in market to book premium for firm i over the most recent ten months ending prior to quarter t's earnings announcement date;
$\sigma(EY, CMMB)_{it}$ = covariance of earnings yield and change in market to book premium for firm i over the most recent ten months ending prior to quarter t's earnings announcement date; and
$VarBeta_{it}$ = coefficient β in regression of realized variance of firm i's equity returns on the realized variance of the S&P 500 index returns, estimated quarterly using data from the prior ten quarters.

Out-of-Sample Tests

Table 7 presents results for out-of-sample tests of H4. Panel A presents pooled results for the mean squared prediction error in from the "Base" and "ABD" models. The mean squared error generated by the base model is 0.149. This is substantially higher than the mean squared error generated by the ABD model (0.141). The mean and median differences between the two mean squared errors (0.00723 and 0.00064, respectively) are also significantly positive, confirming the hypothesis that inclusion of accounting-based volatility drivers in volatility forecasting models improves forecasting accuracy for firms without option-implied volatility estimates. In additional to being statistically relevant, this reduction in forecast error also has economic significance. Incorporating accounting information leads to a 5 percent reduction of the *ex ante* forecast error. It

TABLE 7

Out-of-Sample Tests

Base model: $\sigma^{RV}_{i,t,\tau} = \alpha_1 + \beta_t \sigma^{RV}_{i,t,-\tau} + \varepsilon_{i,t}$

ABD model: $\sigma^{RV}_{i,t,\tau} = \alpha_2 + \delta_t \sigma^{RV}_{i,t,-\tau} + \zeta^j_t ABD^j_{i,t-1} + \varepsilon_{i,t}$

	Mean	**Median**
$\sigma^{RV}_{i,t,\tau} - E[\sigma^{RV}_{i,t,\tau}]_{BASE}$	0.149	0.038
$\sigma^{RV}_{i,t,\tau} - E[\sigma^{RV}_{i,t,\tau}]_{ABD}$	0.141	0.035
Difference (Base − ABD)	0.00723***	0.00064***

*, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively, of the hypothesis that the given estimate is different from zero.
This table provides the prediction errors from out-of-sample testing of the Base and ABD volatility prediction models discussed in Section III.

also has meaningful implications for the estimation of stock-based compensation expense. Volatility forecasts are necessary to value the options offered to employees as compensation. On average, a 100 basis point change in volatility forecasts is associated with a 3.23 percent change in stock-based compensation expense.[7] This implies that the 70 basis point improvement in volatility forecasting obtained by including accounting information in the forecasts would correspond to approximately a 2.26 percent change in stock-based compensation expense.

V. SENSITIVITY ANALYSES

Put-Call Parity

A key assumption underlying my analyses is that the options market uses market price as the price of equity in determining the price of the option. This assumption is consistent with empirical evidence revealing that deviations from put-call parity are rare and typically temporary (Cremers and Weinbaum 2010). Nonetheless, I explore the sensitivity of my results to this assumption by reestimating Equation (2) including an estimate of volatility spread as an additional regressor. Volatility spread captures option market perceptions of equity mispricing. If option market perceptions of equity mispricing are driving my primary findings, then including a volatility spread control variable should eliminate the incremental statistical significance of the accounting-based volatility drivers. I estimate the volatility spread for each firm-day as the open-interest weighted average difference between implied volatilities of call and put options on the firm's equity with the same maturity that were traded on that day. Table 8 presents the summary statistics from the modified Equation (2). Consistent with my predictions, there are significant correlations between future realized volatility and each accounting-based volatility driver, even after controlling for volatility spread in addition to implied volatility, equity market liquidity, and past realized volatility. More importantly, all variables maintain the same directional relation incremental to implied volatility.

[7] In calculating this estimate, I assume the following: options are issued at-the-money, the risk-free rate is zero, and the average vega for options issued as compensation is 3.23 percent. My estimate of option vega is based on the findings of Aboody, Barth, and Kasznik (2006).

TABLE 8

Regressions of Realized Volatility on Fundamental Variables and Model-Free Implied Volatility with Volatility Spread Controls

$$\sigma^{RV}_{i,t,\tau} = \alpha + \beta_1 \sigma^{RV}_{i,t,-\tau} + \beta_2 \sigma^{MFIV}_{i,t,\tau} + \beta_j ABD^{j}_{i,t-1} + \beta_4 Spread_{it} + \beta_5 VolSpread_{it} + \gamma Year + \delta Industry + \varepsilon_{i,t}$$

	I	II	III	IV
$\sigma^{MFIV}_{i,t,\tau}$	0.716^{+}	0.722^{+}	0.721^{+}	0.714^{+}
	(−12.156)	(−12.178)	(−12.157)	(−12.677)
$\sigma^{RV}_{i,t,-\tau}$	0.216^{+}	0.215^{+}	0.215^{+}	0.215^{+}
	(−33.344)	(−33.281)	(−33.302)	(−33.245)
$VolSpread_{it}$	0.023	0.024	0.023	0.027
	(0.588)	(0.630)	(0.605)	(0.697)
$\sigma^{EY}_{i,t}$	0.032***			0.041***
	(4.607)			(4.953)
$\sigma^{CMMB}_{i,t}$		−0.031***		−0.031***
		(−3.458)		(−3.425)
$\sigma(EY, CMMB)_{it}$			0.034***	0.001
			(3.182)	(1.127)
Obs.	44,348	44,348	44,348	44,348
Adj. R^2	0.64	0.62	0.61	0.65
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes

*, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively, of the hypothesis that the given coefficient is different from zero.

$^{+}$ Indicates significance at the 1 percent level of the hypothesis that the given coefficient is different from 1.

This table provides the coefficient estimates from Equation (3) using each of the accounting-based volatility drivers discussed in Section II and a control for volatility spread defined in Section V. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors.

Variable Definitions:

$\sigma^{RV}_{i,t,\tau}$ = sum of squared five-minute returns on firm *i*'s equity over the 30-day window starting five days after the quarter *t* earnings announcement date;

$\sigma^{MFIV}_{i,t,\tau}$ = model-free implied volatility inferred from closing option prices on firm *i*'s equity on the fifth day after the quarter *t* earnings announcement date;

$\sigma^{EY}_{i,t}$ = standard deviation of earnings yield for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;

$\sigma^{CMMB}_{i,t}$ = standard deviation of the change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;

$\sigma(EY, CMMB)_{it}$ = covariance of earnings yield and change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date; and

$VolSpread_{it}$ = open interest-weighted average difference in implied volatilities between call options and put options (with the same strike price and maturity) on firm *i*'s equity measured on the quarter *t* earnings announcement date.

Filing Dates

In my main analysis, I measure implied volatility immediately after the quarter *t* earnings announcement and match it to accounting information from quarter *t*–1. This creates a lag between the release of accounting information and my estimation of implied volatility. To reduce this lag, I also examine implied volatility estimates calculated five days after the 10-Q or 10-K filing date for each firm-quarter, which I obtain from the Securities and Exchange Commission's (SEC) EDGAR

TABLE 9

Regressions of Realized Volatility on Fundamental Variables and Model-Free Implied Volatility: Filing Dates Sample

$$\sigma^{RV}_{i,t,\tau} = \alpha + \beta_1 \sigma^{RV}_{i,t,-\tau} + \beta_2 \sigma^{MFIV}_{i,t,\tau} + \beta_j ABD^{j}_{i,t-1} + \gamma Year + \delta Industry + \varepsilon_{i,t}$$

	I	II	III	IV
$\sigma^{MFIV}_{i,t,\tau}$	0.720^{+}	0.736^{+}	0.734^{+}	0.721^{+}
	(−11.328)	(−11.418)	(−11.322)	(−11.345)
$\sigma^{RV}_{i,t,-\tau}$	0.176^{+}	0.178^{+}	0.177^{+}	0.176^{+}
	(−33.119)	(−34.735)	(−34.792)	(−32.979)
$\sigma^{EY}_{i,t}$	0.008***			0.007***
	(8.230)			(7.635)
$\sigma^{CMMB}_{i,t}$		−0.002**		−0.002*
		(−2.200)		(−1.740)
$\sigma(EY, CMMB)_{it}$			0.002	0.002*
			(1.617)	(1.801)
Obs.	41,002	41,002	41,002	41,002
Adj. R^2	0.63	0.62	0.61	0.67
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes

*, **, *** Indicate significance at the 10 percent, 5 percent, and 1 percent levels, respectively, of the hypothesis that the given coefficient is different from zero.

$^{+}$ Indicates significance at the 1 percent level of the hypothesis that the given coefficient is different from 1.

This table provides the coefficient estimates from Equation (3) using each of the accounting-based volatility drivers discussed in Section II on the sample of firm-quarters for which implied volatility is measured five days after the filing date. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors.

Variable Definitions:

$\sigma^{RV}_{i,t,\tau}$ = sum of squared five-minute returns on firm *i*'s equity over the 30-day window starting five days after the quarter *t* earnings announcement date;

$\sigma^{MFIV}_{i,t,\tau}$ = model-free implied volatility inferred from closing option prices on firm *i*'s equity on the fifth day after the quarter *t* earnings announcement date;

$\sigma^{EY}_{i,t}$ = standard deviation of earnings yield for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date;

$\sigma^{CMMB}_{i,t}$ = standard deviation of the change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date; and

$\sigma(EY, CMMB)_{it}$ = covariance of earnings yield and change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date.

database. I match these implied volatility estimates with accounting-based volatility driver calculations based on the most immediately released 10-Q or 10-K. Using this modified dataset, I reestimate Equation (3) to examine whether fundamental variables remain incrementally predictive of future volatility relative to implied volatility.

Table 9 presents the summary statistics from the estimation of Equation (3) with this modified dataset. As in the original estimation, the results in Table 9 reveal that the volatility of changes in the market to book premium (earnings yield volatility) is negatively (positively) associated with realized volatility incremental to model-free implied volatility. Although the coefficient on the covariance of earnings yield with changes in market to book premium is not statistically significant independently, it does become significant when all three volatility-based drivers are included in column IV. The consistency of these results with my main analyses suggests that the incremental explanatory power of accounting-based fundamentals relative to implied volatility is not solely due to the choice of implied volatility measurement date.

VI. CONCLUSION

In this article, I use an analytical framework to identify accounting-based drivers of equity volatility. I show that these variables are useful in the prediction of equity volatility and in the prediction of the difference between option-implied volatility and future realized volatility. Prior research establishes that implied volatility is an imperfect estimator of future realized volatility, but the precise cause of the difference has remained unclear. Studies often refer to this difference as the "variance risk premium," but do not offer any insight into what drives variation in the premium. My study is the first to hypothesize and find evidence that information about equity volatility from financial statements is useful for this purpose.

I first develop a model of future equity return volatility that consists of three components: earnings yield volatility, change in market to book premium volatility, and the covariance of the earnings yield with the change in market to book premium. My model describes equity return volatility as being increasing in the two volatility terms and decreasing in the covariance term. My empirical analyses reveal that approximately 30 percent of total variation in observed equity volatility is explained by variation in these drivers. As a benchmark forecast of future volatility, I use the expectation of volatility implied by options prices under the Britten-Jones and Neuberger (2000) model-free approach. I find that the accounting-based drivers of volatility that I identify can supplement implied volatility in predicting future volatility. In doing so, I identify firm fundamentals that are indicative of variance risk.

I also find that some of the observed differences between option-implied and realized equity volatilities are not fully explained by the variance risk premium. Specifically, I estimate directly a measure of variance risk exposure and show that two of the three accounting-based volatility drivers remain significantly predictive of future option returns even after controlling for this measure. This result suggests that the fundamentals I identify are not predictive of the difference between implied and realized volatility merely because they are indicative of a variance risk premium. By identifying a source of bias in implied volatility that is distinct from model misspecification, I contribute to the literatures on implied volatility estimation and volatility forecasting.

My results also have implications for options pricing, as I show that accounting-based fundamental information can be used to predict option returns. A trading strategy based on my accounting-based volatility drivers generates significantly positive straddle returns. The returns from this strategy are incremental to returns that investors can earn from a purely market-based strategy using historical or implied volatility. They also persist across individual years of the sample. These suggest that the option market's failure to fully process volatility-relevant fundamental information from financial statements explains some of the previously documented bias in implied volatility. Overall, this article provides a link between the literatures on financial statement analysis and on volatility forecasting. In doing so, I further our understanding of the source of the variance risk, a longstanding concept in the asset pricing literature with very little fundamental foundation heretofore.

REFERENCES

Aboody, D., M. Barth, and R. Kasznik. 2006. Do firms understate stock-based compensation expense disclosed under SFAS 123? *Review of Accounting Studies* 11 (4): 429–461.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61 (1): 259–299.

Barth, M. E., and E. C. So. 2014. Non-diversifiable volatility risk and risk premiums at earnings announcements. *The Accounting Review* 89 (5): 1579–1607.

Battalio, R., B. Hatch, and R. Jennings. 2004. Toward a national market system for U.S. exchange-listed equity options. *Journal of Finance* 59 (2): 933–962.

Breeden, D., and R. Litzenberger. 1978. Prices of state-contingent claims implicit in option prices. *Journal of Business* 51 (4): 621–651.

Britten-Jones, M., and A. Neuberger. 2000. Option prices, implied price processes, and stochastic volatility. *Journal of Finance* 55 (2): 839–866.

Canina, L., and S. Figlewski. 1993. The informational content of implied volatility. *Review of Financial Studies* 6 (3): 659–681.

Carr, P., and L. Wu. 2009. Variance risk premiums. *Review of Financial Studies* 22 (3): 1311–1341.

Chiras, D., and S. Manaster. 1978. The information content of option prices and a test of market efficiency. *Journal of Financial Economics* 6 (2): 213–234.

Christensen, B., and N. Prabhala. 1998. The relation between implied and realized volatility. *Journal of Financial Economics* 50 (2): 125–150.

Christoffersen, P., R. Goyenko, K. Jacobs, and M. Karoui. 2011. *Illiquidity Premia in the Equity Options Market*. Working paper, University of Toronto, McGill University, and University of Houston.

Cremers, M., and D. Weinbaum. 2010. Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis* 45 (2): 335–367.

Day, T., and C. Lewis. 1992. Stock market volatility and the informational content of stock index options. *Journal of Econometrics* 52 (1): 267–287.

De Fontnouvelle, P., R. P. Fishe, and J. H. Harris. 2003. The behavior of bid-ask spreads and volume in options markets during the competition for listings in 1999. *Journal of Finance* 58 (6): 2437–2464.

Fama, E., and J. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607–636.

Fama, E., and K. French. 1997. Industry costs of equity. *Journal of Financial Economics* 43 (2): 153–193.

Fu, F. 2009. Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics* 91 (1): 24–37.

Goodman, T., M. Neamtiu, and F. Zhang. 2014. *Fundamental Analysis and Option Returns*. Working paper, Purdue University, The University of Arizona, and Yale University.

Gow, I., G. Ormazabal, and D. Taylor. 2010. Correcting for cross-sectional and time-series dependence in accounting research. *The Accounting Review* 85 (2): 483–512.

Goyal, A., and A. Saretto. 2009. Cross-section of option returns and volatility. *Journal of Financial Economics* 94 (2): 310–326.

Hull, J. 2009. *Options, Futures and Other Derivatives*. Englewood Cliffs, NJ: Pearson Prentice Hall.

Jiang, G., and Y. Tian. 2005. The model-free implied volatility and its information content. *Review of Financial Studies* 18 (4): 1305–1335.

Jiang, G., and Y. Tian. 2007. Extracting model-free volatility from option prices: An examination of the VIX index. *Journal of Derivatives* 14 (3): 35–60.

Jin, W., J. Livnat, and Y. Zhang. 2012. Options prices leading equity prices: Do options traders have an information advantage? *Journal of Accounting Research* 50 (2): 401–432.

Lamoureux, C., and W. Lastrapes. 1993. Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. *Review of Financial Studies* 6 (2): 293–326.

Latane, H., and R. Rendleman. 1976. Standard deviations of stock price ratios implied in option prices. *Journal of Finance* 31 (2): 369–381.

Lee, C. M. C. 2001. Market efficiency and accounting research: A discussion of "Capital market research in accounting" by S. P. Kothari. *Journal of Accounting and Economics* 31 (1/3): 233–253.

Markowitz, H. 1952. Portfolio selection. *Journal of Finance* 7 (1): 77–91.

Mayhew, S. 2002. Competition, market structure, and bid-ask spreads in stock option markets. *Journal of Finance* 57 (2): 931–958.

Merton, R. 1973. The theory of rational option pricing. *Bell Journal of Economics and Management Science* 4 (1): 141–183.

Pan, J. 2002. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics* 63 (1): 3–50.
Patton, A. J. 2011. Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160 (1): 246–256.
Penman, S., F. Reggiani, S. Richardson, and I. Tuna. 2012. *A Characteristic Model for Asset Pricing*. Working paper, Columbia University, Bocconi University, and London Business School.
Pool, V., H. Stoll, and R. Whaley. 2008. Failure to exercise call options: An anomaly and a trading game. *Journal of Financial Markets* 11 (1): 1–35.
Roll, R., E. Schwartz, and A. Subrahmanyam. 2010. O/S: The relative trading activity in options and stock. *Journal of Financial Economics* 96 (1): 1–17.

APPENDIX A

Calculating Model-Free Implied Volatility

As discussed in Section II, I calculate a model-free implied volatility estimate as an alternative to Black-Scholes implied volatility. Britten-Jones and Neuberger (2000) define model-free implied volatility as follows:

$$\sigma_T^{MFIV} = \frac{2e^{rT}}{T}\int_0^{F_T}\frac{P(T,K)}{K^2}dK + \int_{F_T}^{\infty}\frac{C(T,K)}{K^2}dK$$

where T is the time to expiration in years; r is the annualized risk free rate; $\{K_i\}$ is the set of available strike prices; F_T is the forward price of the underlying security; $C(T, K)$ is the value of a call option; and $P(T, K)$ is the value of a put option. This result is driven by the observation of Breeden and Litzenberger (1978) that the second derivative of a call option price with respect to the strike price is equivalent to the risk-neutral density. The derivation of this result begins with the assumption that the price of a call option at any point in time is the expectation of its future payoff:

$$C(T,K) = E[\max\{S_T - K, 0\}]$$

Under the risk-neutral density ($\phi_t[S_t]$), the above expectation can be rewritten as the sum of two integrals:

$$C(T,K) = \int_{-\infty}^{K}\max\{S_T - K, 0\}\phi_T(S_T)dS_T + \int_K^{\infty}\max\{S_T - K, 0\}\phi_T(S_T)dS_T$$

When the price of the underlying security is less than the strike, the value of the call option will be zero. This reduces the above summation to a single term:

$$C(T,K) = \int_K^{\infty}(S_T - K)\phi_T(S_T)dS_T$$

Differentiating this expression for the call option price with respect to K yields:

$$\frac{\partial C(T,K)}{\partial K} = \int_K^{\infty}\phi_T(S_T)dS_T$$

and differentiating a second time with respect to K generates the Breeden and Litzenberger (1978) result:

$$\frac{\partial^2 C(T,K)}{\partial K^2} = \phi_T(K)$$

Using this result, Britten-Jones and Neuberger (2000) derive the above expression for model-free implied volatility under the assumption that asset prices follow a diffusion process:

$$\frac{dF_t}{F_t} = \sigma_t dW_t$$

By Ito's Lemma, this implies:

$$d\ln F_t = \sigma_t dW_t - \frac{1}{2}\sigma_t^2 dt$$

$$\sigma_t^2 dt = 2[d\ln F_t + \sigma_t dW_t]$$

Integrating over time yields:

$$\int_0^T \sigma_t^2 dt = 2[\ln F_0 - \ln F_t + \sigma W_t]$$

then taking expectations under the risk-neutral density and recalling the Breeden and Litzenberger (1978) result:

$$\begin{aligned} E_0^F\left[\int_0^T \sigma_t^2 dt\right] &= 2\left[\ln F_0 - E_0^F(\ln F_t)\right] \\ &= \int_0^\infty \frac{C^F(T,K) - \max\{0, F_0 - K\}}{K^2} dK \\ &= \int_{F_0}^\infty \frac{C^F(T,K)}{K^2} dK - \int_0^{F_0} \frac{F_0 - K}{K^2} dK \\ &= \int_{F_0}^\infty \frac{C^F(T,K)}{K^2} dK + \int_0^{F_0} \frac{P^F(T,K)}{K^2} dK \end{aligned}$$

Note that this expression requires integration over the entire range of possible strike prices. Since such integration is not empirically feasible, I employ the following approximation, derived by Jiang and Tian (2005), in my calculation of model-free implied volatility:

$$\sigma_{MFIV}^2 = \frac{2e^{rT}}{T}\left[\sum_{i=1}^{S} \frac{\Delta K_i}{K_i^2} P_T(K_i) + \sum_{i=1}^{M} \frac{\Delta K_i}{K_i^2} C_T(K_i)\right]$$

where q is the annual dividend rate (assumed 0); S_o is the price of the underlying asset at $t=0$; and ΔK_i is the change in strike price. The above approximation assumes:

$$F_T = S_0 e^{(r-q)T} = S_0 e^{rT}$$

APPENDIX B

Variable Definitions

Variable	Definition
$\sigma^{RV}_{i,t,\tau}$	The logarithm of the observed sum of firm *i*'s squared five-minute equity returns over τ days starting five days after quarter *t*'s earnings announcement date.
$\sigma^{MFIV}_{i,t,\tau}$	The logarithm of model-free implied volatility of an option on firm *i*'s equity measured five days after quarter *t*'s earnings announcement date with τ days remaining until expiration.
EY_{it}	The earnings yield for firm *i* in month *t,* which is the ratio of most recently reported quarterly earnings to closing stock price on the last trading day of month *t*.
$CMMB_{it}$	The change in market to book premium, defined as $(MV_t - BV_t) - (MV_{t-1} - BV_{t-1})$. MV_t is the market value of firm *i*'s equity on the last trading day of month *t,* and BV_t is the book value of equity from the most recently filed 10-Q that is publicly available on the last trading day of month *t*.
$\sigma^{EY}_{i,t}$	The standard deviation of earnings yield for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date.
$\sigma^{CMMB}_{i,t}$	The standard deviation of the change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date.
$\sigma(EY, CMMB)_{it}$	The covariance of earnings yield and change in market to book premium for firm *i* over the most recent ten months ending prior to quarter *t*'s earnings announcement date.
$Spread_{it}$	The logarithm of the median volume-weighted bid-ask spread for all options on firm *i*'s equity over the year ending on the quarter *t* earnings announcement date.
$VolSpread_{it}$	The open interest-weighted average difference in implied volatilities between call options and put options (with the same strike price and maturity) on firm *i*'s equity measured on the quarter *t* earnings announcement date.
ERV_NoABD_{it}	The predicted value from the following regression for firm *i* in quarter *t,* estimated using quarter $t-1$ data: $\alpha_1 + \beta_1\sigma^{MFIV}_{i,t,\tau} + \beta_2\sigma^{RV}_{i,t,-\tau} + \varepsilon_{i,t}$.
ERV_ABD_{it}	The predicted value from the following regression for firm *i* in quarter *t,* estimated using quarter *t*–1 data: $\alpha_2 + \beta_3\sigma^{MFIV}_{i,t,\tau} + \beta_4\sigma^{RV}_{i,t,-\tau} + \beta_5 ABD_{i,t} + \varepsilon_{i,t}$.
$VarBeta_{it}$	The coefficient β in the following regression: $\log RV_i = \alpha + \beta \log RV_{SPX} + \varepsilon$. In this regression, RV_j is the realized variance of firm *i*'s equity returns and RV_{SPX} is the realized variance of the S&P 500 index returns. It is estimated quarterly using data from the prior ten quarters.