FUNDAMENTAL ANALYSIS AND EQUITY VOLATILITY

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL OF BUSINESS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Suhas A. Sridharan

May 2013

This dissertation is online at: http://purl.stanford.edu/wd090fj9866

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Mary Barth, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Ron Kasznik

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Charles Lee

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Abstract

I examine whether financial statement information can predict future realized volatility incremental to the volatility implied by option market prices. Prior research establishes that option-implied volatility is a biased estimator of future realized volatility. I hypothesize that financial statement information, by providing information about economic events correlated with future volatility, is informative in the prediction of future volatility and not fully incorporated in either past volatility or the market's expectation of future volatility. I confirm this empirically and show that the finding is robust to the measurement of option-implied volatility using either the Black-Scholes formula or a model-free approach. I also 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 volatility. Additionally, I provide evidence that contradicts a risk-based explanation for the incremental predictive ability of accounting information. Taken together, my results indicate that the market's failure to fully process accounting-based fundamental information explains some of the previously documented bias in implied volatility.

Preface

This dissertation presents a fundamental analysis approach to volatility forecasting. I begin by discussing the motivation for volatility forecasting in the context of risk assessment and asset pricing. I also provide an overview of the prior literature on fundamental analysis as it relates to forecasting volatility. I then develop hypotheses and present evidence indicating the usefulness of accounting information for forecasting volatility. Finally, I explore practical applications of this research to asset pricing and discuss implications for future research.

My thesis addresses two research questions. The central question of my thesis asks if accounting information can be used by equity investors to form precise estimates of future equity volatility. To answer this question, I develop a fundamental analysis approach to assess future volatility. In doing so, I extend the literature on fundamental analysis beyond the prediction of mean equity returns. To give context to this new approach to volatility forecasting, I provide an overview of the existing literature on equity volatility forecasting. My survey of the literature highlights the persistent bias in option-implied volatility as a forecast of future equity volatility. This bias is particularly puzzling given the relative superiority of option-implied volatility to other extant forecasting techniques. This leads me to my second research question: is the bias in option-implied volatility driven by option-market participants' failure to incorporate accounting information in their assessments of future volatility? To answer this question, I demonstrate my fundamental analysis approach to volatility forecasting can effectively supplement option-implied volatility to generate more precise forecasts of future volatility.

The results of my thesis have several implications for practitioners and academics. I highlight one practical implication of my results by using my fundamental analysis approach to sort options into profitable straddle portfolios. The repeated implementation of this strategy could help eliminate possible mispricings in options markets. More generally, my thesis provides deeper insight into how capital markets participants process fundamental information. Prior research on fundamental analysis has been relatively limited in scope, focusing only on the assessment of mean equity returns. I extend the literature by choosing instead to focus on equity returns volatility. As a first-order approximation for risk, understanding equity volatility is essential to efficient portfolio allocation. Therefore a complete understanding of how volatility is assessed and priced by markets is of central importance to the asset pricing literature. The results of my dissertation also provide accounting practitioners and academics with a new perspective on the informativeness of corporate financial statement disclosures. Our understanding of accounting for risk assessment is limited. The results of my paper highlight how investors and academics might use existing financial statement disclosure requirements to assess risk.

Acknowledgements

It is not possible to mention by name everyone who deserves acknowledgement in this thesis, but I would like to thank a few people who played particularly significant roles.

I am deeply grateful to my dissertation committee - Mary Barth, Charles Lee, and Ron Kasznik - for their encouragement and guidance throughout my time at Stanford. I owe a special debt of gratitude to my advisor, Mary Barth. I thank her for generously investing so much of her time mentoring me, for patiently reading countless drafts of my work, for motivating me through her own example, and for supporting me along this journey. I feel incredibly fortunate to have had the opportunity to work with and learn from her. I am also lucky to have arrived at Stanford at the same time as Charles Lee. A doctoral student could not ask for a kinder or more supportive mentor than Charles. I have benefited tremendously from his experience, intellect, and friendship. I thank Ron Kasznik for all his feedback and guidance on my dissertation and for providing me invaluable counsel during the job-market process. I thank Dan Boneh for chairing my oral defense committee. I also owe a special thanks to Grace Pownall, without whom I would not have discovered the path to academia.

Several other students and faculty made my time at Stanford immensely rewarding. Most importantly, I owe a deep debt of gratitude to Eric So. He has been a constant source of support and encouragement from my very first day at Stanford. I feel quite fortunate to have met him, as I cannot imagine a more generous colleague or friend. I am also indebted to Zac Peskowitz, Travis Johnson, Maria Ogneva, Arthur Korteweg, Bella Rozenkrants, and Doron Israeli. In addition to providing me insightful feedback on my work, many of them have become good friends over the past years. For that I am quite thankful. I also owe thanks to loved ones far away: my friends Kyle Valenti, Emily Allen, Jonathan Beam, and Rachel Zelkowitz; and my brother Jay Krishna Sridharan. Their love and unwavering confidence in me has kept me grounded through the highs and lows of the past five years.

Lastly and most importantly, I dedicate this thesis to my parents, Drs. V and Uma Sridharan. They were my first examples of what it means to be an academic and continue to inspire me through their own dedication and hard work. They have always believed in me and have never hesitated to put my needs before their own. My debt of gratitude to them is too deep for words to adequately describe. This thesis, as do all of my accomplishments, belongs to them.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this paper I examine how accounting information can be used to supplement existing volatility forecasting techniques to generate more precise volatility forecasts. In doing so, I document a new channel through which accounting information is useful for equity valuation. A central theme in accounting research is the link between disclosure and equity valuation. This is a natural consequence of the Financial Accounting Standards Board's stated goal to provide relevant information for equity investors. Studies relating corporate disclosure to equity valuation typically focus on how accounting information can inform predictions of the mean of future equity returns. Though informative, 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 understand the informativeness of these disclosures for equity valuation. An equally relevant consideration is how these disclosures relate to the risk in equity investment. However, this relation has been relatively overlooked by both the accounting and finance literatures.

Equity volatility provides a estimate of the total risk in equity investment, including idiosyncratic risk. In theory, investors can eliminate exposure to idiosyncratic risk through diversified investment. In practice, however, limits to diversification can leave investors exposed to some amount idiosyncratic risks, making equity volatility a relevant risk metric for equity valuation [Ang et al., 2006, Fu, 2009]. As measures of market uncertainty, equity volatility estimates are used not only in investment decision-making but also for macroeconomic analyses, making them important to academics, practitioners, and market regulators. Volatility forecasts are also central to derivatives trading because they are a key input in many derivative pricing models.

The need for accurate predictions of future volatility has resulted in a large literature devoted to volatility forecasting, and a variety of different forecasting techniques exist. Comparisons of these techniques reveal that estimates inferred from traded options prices, also known as implied volatility estimates, consistently outperform other classes of volatility forecasts in terms of minimizing forecast errors [Poon and Granger, 2003]. Nonetheless, a significant gap between implied and realized volatility persists [Christensen and Prabhala, 1998]. This difference is puzzling since the option market's expectation of volatility should, on average, equal future realized volatility under the assumption of informationally efficient markets.[1] The idea that implied volatility, as an estimate of the option market's expectation of volatility, should equal future realized volatility is analogous to the idea that equity prices are an unbiased expectation of a firm's future cash flows. A substantial literature in accounting and finance documents that equity prices underreact to financial statement information in forecasting future cash flows. The literature attributes this underreaction to several

[1]This assertion is distinct from the assertion that Black-Scholes implied volatility should equal future realized volatility. My ability to accurately measure the option market's expectation of future volatility might be imperfect given data restrictions and modeling constraints. I discuss the limitations of existing measurement techniques in more detail in section 2.2 of the paper. However, that does not change the theoretical assertion that the market's expectation of volatility, if unbiased, should on average equal realized volatility.

nonconflicting explanations including market frictions, behavioral biases, and investor learning. I extend this literature by examining whether investors also underreact to financial statement information in forecasting future equity volatility realizations. I hypothesize that the difference between implied and realized volatility is partially attributable to the options market's oversight of relevant accounting-based fundamentals.

In demonstrating that financial statement information can predict future realized equity volatility incremental to implied volatility, I shed light on a longstanding empirical puzzle in the volatility forecasting literature: the bias in implied volatility as an estimator of future realized volatility. Prior research attributes differences between implied and realized volatility to a premium demanded by investors for exposure to variance risk. However, I document patterns in firm-level implied and realized volatilities that directly contradict a risk-based explanation for the difference between implied and realized volatility. In doing so, I mitigate concern that the fundamentals I identify are merely correlates with variance risk premia. Overall, my findings present an alternative to previously accepted explanations for the source of the difference between implied and realized volatility.

I measure the market's expectation of future volatility in two ways: the expected volatility implied by options prices using the Black-Scholes formula and using the Britten-Jones and Neuberger [2000] model-free approach. Black-Scholes implied volatility is a commonly used measure of expected volatility. However, the Black-Scholes model assumes equity prices follow a diffusion process and several studies document jumps in asset prices that violate this assumption. Prior literature identifies the violation of this assumption as a source of measurement error in Black-Scholes implied volatility as a proxy for the option market's expectation of future volatility. To mitigate these concerns I also measure the options market's expectation of future volatility using model-free implied volatility, which is less sensitive to measurement error concerns because it does not require assumptions about the distribution of asset returns as does the Black-Scholes formula.

The accounting literature demonstrates that financial statement information is useful in identifying firms with growth opportunities, quantifying default or crash risk, and predicting extreme returns. I expect equity volatility to be higher for firms with more expected growth, higher default or crash risk, and greater probability of extreme returns. I hypothesize that information useful in predicting these fundamentals is useful in predicting equity volatility as well. I focus my analysis on eight explanatory variables: firm size, equity book-to-market ratio, cash flow volatility, earnings opacity, research and development expenditure, sales growth, return on assets, and leverage. My tests examine whether these eight financial statement metrics of growth, uncertainty in operations, and default and crash risks are significantly associated with future equity volatility and whether these relations persist after controlling for the market's expectation of future volatility.

My evidence is based on a sample of 78,034 quarterly observations from 3,934 firms between 1996 and 2010. My analyses reveal that accounting-based fundamental information is relevant for volatility prediction but is not fully incorporated in the option market's assessment of future volatility. In particular, the options market systematically underestimates volatility for firms with higher research and development expenses, higher cash flow volatility, and greater earnings management relative to their industry medians. It overestimates volatility for large firms, highly levered firms, and firms with high return on assets or high equity book-to-market ratios. To eliminate possible confounding effects I include controls for both past volatility and liquidity and find that neither subsumes financial statement information in terms of informativeness for future volatility prediction.

The results of this paper have implications for asset pricing. I document the ability to generate positive returns by taking a long (short) straddle position in firms with financial statement information indicative of low (high) volatility. A straddle position is formed by purchasing an at-the-money call option and an at-the-money put option on the the same underlying asset. This generates a payoff that is increasing in absolute price change of the underlying asset but is insensitive to the direction of the change. Consequently, information about the volatility of the equity prices should be useful in determining straddle payoffs. Consistent with the hypothesis that financial statement information helps to predict future equity volatility, I find that an unconditional sort of straddle positions using my volatility score (comprised of the eight metrics previously discussed) generates a 12.9% annualized return. These results suggest that option markets fail to fully process the information available in financial statements when forming volatility expectations. In demonstrating this application, my study complements contemporaneous work by Goodman et al. [2013], which shows that straddle returns are predictable using the residual expected stock price change based on select accounting-based fundamentals.

Additionally, this study 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 gap 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 impacts equity valuation.

Chapter 2

Hypothesis development

My thesis 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.

2.1 Fundamental analysis and equity volatility

2.1.1 The relation between mean and variance

A substantial literature in accounting and finance documents the usefulness of financial statement information for predicting mean returns. This literature provides a framework for identifying pieces of information that might be relevant for volatility forecasting by providing insight into the information content of financial statement disclosures. Although the literature identifies several returns-predictive variables derived from financial statements, these results have limited applicability to my research questions for several reasons. The ability to draw inferences about volatility prediction from the extant research on mean returns prediction depends critically upon the mechanism through which returns predictability occurs.

Under the assumption of informationally efficient markets, any variable that predicts equity returns must be a priced risk factor. It is possible that certain priced risk factors suggest higher equity volatility along with higher mean equity returns. I discuss this possibility further in the section 2.1.2. However, not all priced risk factors necessitate higher equity volatility. For example, systematic risk is a *relative* volatility measure that does not have a clear monotonic relationship with the level of equity volatility. Firms can exhibit low equity volatility and still have high systematic risk as long as the market volatility is also low. The non-monotonic nature of the relation of systematic risk to equity volatility makes it impossible to forecast equity volatility using previously documented relationships between financial statement information and systematic risk. Separate empirical analysis is necessary to ascertain whether a direct relation exists between this information and equity volatility. Moreover, several studies argue that returns predictability is driven by informational inefficiency in the equity market rather than by predictive variables being priced risk factors. If this is true, prior research becomes even less informative for future volatility prediction. If informational inefficiencies exist in the equity market, 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 volatility.

The existing mean returns prediction literature cannot address questions of volatility prediction because the mean and variance of any random variable are fundamentally distinct concepts. To better understand the lack of relation between mean and variance of equity returns, consider the general definition of the variance of a random variable r with mean μ

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

This expression of variance highlights the fact that the exact relationship (or lack thereof) 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. While there are distributions for which the mean is a sufficient statistic of the variance (the Poisson distribution is one example), such distributions do not provide accurate descriptions of the returns generating process. In all of the commonly used distributions to describe equity returns, mean and variance are distinct quantities with limited relation.

A common assumption is asset pricing models is that equity returns are normally distributed. Under this assumption there would be no direct link between the literature on assessing mean equity returns and my research question, as there is no structural relationship between mean and variance in the normal distribution. 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, its mean and variance are calculated as follows

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

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. The lack of relation between mean and variance limits my ability to extend the results of past research on mean returns to my hypotheses about the variance of returns.

2.1.2 Identifying fundamentals

Prior literature documents significant relations between accounting information and measures of risk or uncertainty in operations that are likely to be indicative of future equity volatility. These constructs include default and priced equity risks, the incidence of extreme equity returns, and growth opportunities for the firm. Each of these constructs is the subject of prior research, and from these literatures I focus on the following eight fundamental metrics: size, equity book-to-market ratio, leverage, return on assets, R&D expenditure, cash flow volatility, earnings opacity, and growth in sales. In the subsections that follow, I provide a brief description of each variable and its hypothesized relation with volatility.

Indicators of uncertainty in future operations: size and cash flow volatility

Pastor and Veronesi [2003] develop and provide empirical support for a model of equity prices in which equity volatility decreases in a firm's size. This is consistent with a firm's size being negatively related to credit risk [Beaver et al., 2005], the likelihood of merger and acquisition targeting [Barros, 1998], and the likelihood of earnings manipulation [Dechow et al., 2010]. It is also consistent with the significant negative correlation between size and excess returns, which is indicative of increased risk for small firms or the equity market's mispricing of these firms. Consequently, I expect to observe a negative relation between size and volatility in my sample.

The Pastor and Veronesi [2003] model also predicts that equity volatility increases in the volatility of a firm's profits. Assuming that equity market prices are the present value of expected of future cash flows to the firm, changes in price reflect changes in this expectation. Equity volatility captures the rate of change in expectations of future cash flows and can be viewed as a measure of the market's uncertainty regarding the levels of the firm's future cash flows or the discount rate used by the market to value the stream of expected future cash flows. I expect uncertainty about cash flows to be greater when the firm's operations are more unpredictable. I use the volatility of historical cash flows to quantify the extent to which a firm's operations are unpredictable. I assume that cash flow volatility for a specific firm informs equity investors about the uncertainty of that firm's future cash flows but does not provide information about changes in the discount rate since the discount rate is a macroeconomic variable unlikely to be materially affected by a single firm's operational uncertainty. With this assumption, I predict a positive relation between cash flow volatility and equity volatility.

Indicators of growth opportunities: R&D expenditure and sales growth

I expect variables that are positively correlated with a firm's growth opportunities to also be positively correlated with equity volatility because growth opportunities are likely to be associated with increased uncertainty about future firm performance. Research and development (R&D) expenses are often used to measure a firm's growth opportunities because they are an expenditure made by the company in anticipation of future product development and revenue generation. Under U.S. GAAP, most research and development expenditures are expensed as incurred. One reason for this accounting treatment is the uncertain nature of associated benefits; it is unclear that R&D expenditures today will generate revenues in the future. To the extent that R&D expenditures are the result of activities for which future cash flows are uncertain, there should be a positive relation between R&D expense and returns volatility. Chan et al. [2001] document a positive relation between R&D expense and returns volatility but do not investigate whether the options market incorporates the positive relation between R&D expense and returns volatility in its implied volatility estimate.

Prior literature identifies rate of change in sales revenue as a way to identify firms with growth opportunities [Lakonishok et al., 1993, Beneish et al., 2001]. The expected relation between sales growth and growth opportunities is positive, following the assumption that innovations in revenue streams are persistent. Beneish et al. [2001] document a positive relation between the growth in sales revenue and extreme returns. To the extent that firms with growth opportunities are more subject to changes in equity price, I predict a positive relation between growth in sales and equity volatility.

Indicators of equity returns: equity book-to-market ratio and earnings opacity

A positive (negative) association with excess returns for a given variable is either the result of market mispricing or because the variable indicates greater (less) risk. To the extent that this risk is driven by uncertainty, the latter explanation suggests that the variable might be informative about future volatility. Firms with a high book-to-market ratio earn positive abnormal returns [Fama and French, 1992, Piotroski, 2000]. One explanation for this relation is that a high book to market ratio is the result of investor inattention, and positive returns arise as market participants correct the initial inefficiency. In this case, a relation between book to market ratio and future returns volatility is not readily apparent. Conversely, Fama and French [1992] argue that the book to market ratio is a proxy for financial distress and the risk associated with increased distress necessitates lower returns for firms with higher equity book-to-market ratios. Based on this hypothesis, I anticipate firms in financial distress to exhibit more volatility in equity returns. Therefore I predict a positive relation between equity book-to-market ratio and volatility.[1]

Hutton et al. [2009] introduce a measure of earnings opacity based on the sum of three years' absolute discretionary accruals and show that it is significantly positively associated with equity returns and crash risk. The earnings management literature reports that higher levels of discretionary accruals are indicative of more earnings manipulation. Though there is significant debate on how best to separate the discretionary and non-discretionary components of accruals, a consistent finding in the literature is that most discretionary earnings management techniques must be reversed in the future. The subsequent reversal of current-period discretionary accruals might lead to more extreme price movements and higher volatility. Consistent with this hypothesis, Rajgopal and Venkatachalam [2011] show that the increase in firm-specific volatility from 1962 to 1997 first documented by Campbell et al. [2001] is associated with a deterioration in earnings quality. Given this and the documented relation of the Hutton et al. [2009] opacity measure to crash risk, I predict a positive relation between earnings opacity and equity volatility.

Indicators of default risk: leverage and return on assets

The literature using accounting information to estimate probability of default also provides an indirect link between accounting information and future volatility since I anticipate volatility to increase in default risk. Kaplan and Urwitz [1979], Beaver [1966], Altman [1968], Ohlson [1980], Shumway [2001], Hillegeist et al. [2004], and

[1] Prior literature also uses market-to-book ratio as a measure of a firm's expected growth opportunities [Penman, 1996]. Under this interpretation, firms with high book-to-market ratios are expected to exhibit less future growth. Since I expect equity volatility to be increasing in a firm's growth opportunities, this interpretation of book-to-market ratio implies a negative relationship between book-to-market ratio and equity volatility. This contradicts the positive relation I hypothesize under the interpretation of book-to-market ratio as a measure of financial distress. Determining which effect dominates is a empirical question. In section 4.1, I document a positive relationship between book-to-market ratio and equity volatility which supports the interpretation of high book-to-market ratios as indicators of financial distress.

Chava and Jarrow [2004] all present predictive models of bankruptcy probabilities that incorporate current-period financial ratios. Beaver et al. [2005] note that two of the most commonly used ratios in these models are return on total assets and leverage. Firms with high ROA are more profitable and consequently have lower risk of bankruptcy. Because I expect volatility to increase in a firm's credit risk, I predict a negative relation between ROA and returns volatility.

A basic result from the corporate finance literature with consistent empirical support is that highly levered firms exhibit higher conditional probabilities of bankruptcy [Ross, 1977, Beaver et al., 2005]. However, it is also plausible that having high leverage indicates greater financial health because unstable firms would not have access to large amounts of debt. Given these contradictory predictions, the direction of the leverage-volatility relation is an empirical question. Univariate analyses in my sample suggest that the latter effect dominates and highly levered firms will experience lower returns volatility.

In summary, my first hypothesis is

H1: *Characteristics of the firm reflected in financial statements help predict future volatility. Volatility is increasing (decreasing) in book to market ratio, research and development expenditure, earnings opacity, and cash flow volatility measures (size, leverage, return on assets).*

2.2 Implied volatility forecasting

Several studies explore the use of option-implied volatility as a benchmark forecast for future equity volatility [Latane and Rendleman, 1976, Chiras and Manaster, 1978, Lamoureux and Lastrapes, 1993, Christensen and Prabhala, 1998]. In contrast to statistical volatility forecasts, implied volatility estimates are not calculated directly but rather inferred from an assumed options pricing model. In these models, equity volatility is defined as $\sigma_{t|t-1}$ from the equation below

$$r_t = \mu_{t|t-1} + \epsilon_t = \mu_{t|t-1} + \sigma_{t|t-1} z_t \tag{2.1}$$

In equation (2.1), r_t is the ex-dividend return for an asset over period t and the series z_t are independently and identically normally distributed [Poon and Granger, 2003]. $\mu_{t|t-1}$ and $\sigma^2_{t|t-1}$ are the conditional mean and variance of the returns process given the information set at time $t-1$. This equation is equivalent to assuming that asset prices follow a geometric Brownian motion with constant volatility over the period from $t-1$ to t. This assumption is central to the derivation of the Black and Scholes [1973] options pricing formula. In addition to the price diffusion assumption, the original Black-Scholes model assumes options are European, unlimited borrowing is possible at the risk-free rate, equity securities are infinitely divisible, equity investors receive no dividend payments, and there are no arbitrage opportunities. Under these constraints, the original Black-Scholes framework defines implied volatility as the value of σ that satisfies the following equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

where V is the option value, S is the price of the underlying asset, t is the remaining time to maturity, and r is the risk-free rate of return. For traded options, time to maturity, the price of the underlying asset, and current market value are observable and uncontroversial proxies for the risk-free rate of return are readily available. With these inputs, one can invert the Black and Scholes [1973] options-pricing formula to obtain implied volatility, an estimate of the market's expected future volatility of the underlying asset's returns. If markets are informationally efficient and the model is correctly specified, this expectation should be the best predictor of future equity volatility. Poon and Granger [2003] confirm this by comparing forecast errors generated from conditional heteroskedasticity models such as EGARCH, stochastic volatility models, and Black-Scholes implied volatility. Their results reveal that of the three classes of models, Black-Scholes implied volatility forecasts most closely map into realized volatility.

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 variance of equity returns over the remaining life of the option. This implies that a regression of subsequent realized volatility on ex-ante implied volatility should yield a coefficient of one on implied volatility and a coefficient of zero on any other explanatory variables. Empirical evidence on this assertion, however, is mixed. Using 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 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 return volatility. However, Christensen and Prabhala [1998] find that implied volatility is significantly correlated with future realized volatility and that past volatility is fully incorporated in the current market expectation. They find that the coefficient on implied volatility is significantly different from one, indicating bias in implied volatility. The difference between existing research and their findings are driven by their use of non-overlapping S&P 500 index option price data and their employment of a two stage least-square approximation to mitigate measurement error in implied volatility.

The measurement error with which Christensen and Prabhala [1998] are concerned is error induced by misspecification of the Black-Scholes option pricing formula that is inverted to estimate expected future volatility. Even though Black-Scholes implied volatility can be modified to allow for dividend payments and American options, it implicitly assumes that securities are infinitely divisible and that their prices follow a geometric Brownian motion. The Black-Scholes model 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]. 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. 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. The intuition for model free implied volatility estimates comes from Breeden and Litzenberger's [1978] result that the risk-neutral density of returns equals the second derivative of the call option price with respect to the strike price. From this result, Britten-Jones and Neuberger [2000] generate the following expression for option market's expectation of future equity volatility as the area underneath the curve mapping option prices to the range of strike prices.

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

In the above equation, F_t is the forward price of the underlying asset at time t and $C^F(T,K)$ is the price of a call option with strike price K and remaining time to maturity T. 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

[2] In sensitivity analyses I examine the extent to which my results depend upon assuming informational efficiency in equity markets. See section 5.2 for more details.

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] also show that model-free implied volatility measure subsumes the information contained in the corresponding Black-Scholes implied volatility estimate. Despite these improvements, model-free implied volatility appears biased in predicting future volatility [Jiang and Tian, 2005].

2.3 Information processing in the options market

That the gap between implied and realized volatilities persists despite innovations in the measurement of implied volatility suggests that part of it might be the result of a bias in the market's expectations of future volatility. 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 et al., 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 et al., 2008, Roll et al., 2010]. Prior research documents 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, which constructed of a single call and put on the same underlying asset, are increasing in extreme stock price movement. As they typically generate negative returns, they are primarily used as a hedge against other investment positions. In a study contemporaneous with this one, Goodman et al. [2013] demonstrate that, after controlling for implied volatility, the residual expected absolute equity return implied by sales growth and change in earnings per share is also positively associated with long straddle returns.

Prior literature on implied volatility does not acknowledge that informational inefficiencies in the option market could also explain the persistent bias in option-implied volatility relative to realized volatility. However, the fundamental analysis literature provides a foundation for this hypothesis by documenting the underreaction of equity prices to financial statement information in forecasting future cash flows. That capital market participants could underreact in forming the expectation of future cash flow levels suggests the possibility of similar underreaction in forming expectations of the volatility of equity returns. My first hypothesis predicts that financial statement analysis can provide information about future equity volatility. If this hypothesis is correct and the options market does not fully incorporate accounting-based fundamentals in a timely manner, there will be a significant relation between volatility-relevant financial statement information and the magnitude of the implied volatility bias. This leads to my second hypothesis:

H2: *Options markets do not fully capture the information in financial statements in forecasting future volatility.*

Chapter 3

Research design

My empirical analysis consist of three stages. First, I examine the relations between each of my fundamental variables and future realized equity volatility. I then examine the significance of each variable incremental to option-implied volatility in the prediction of future volatility. I first test this hypothesis in a regression framework using Black-Scholes implied volatility. To alleviate concerns about measurement error in Black-Scholes implied volatility, I also explore the use of model-free implied volatility in alternative specifications. Finally, I demonstrate that a trading strategy based on the accounting-based fundamentals used in the first and second stage analyses can generate positive returns on investments in option straddle portfolios. In the sections that follow, I discuss the tests and predictions of each stage of my analysis.

3.1 Testing hypothesis 1

Prior studies of the efficiency of options markets estimate a relation similar to that of equation (3.1) [Day and Lewis, 1992, Lamoureux and Lastrapes, 1993]:

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_1 \sigma^{IV}_{i,t,\tau} + \gamma Year + \epsilon_{it} \tag{3.1}$$

In equation (3.1) $\sigma^{IV}_{i,t,\tau}$ is the logarithm of 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. $\sigma^{RV}_{i,t,\tau}$ is the logarithm of the observed standard deviation of equity returns over the period starting five days after quarter t's earnings announcement date and ending τ days later. $Year$ is a vector of year fixed effects. Since many of the prior studies of option market efficiency use as data a single time series of index option prices, they focus only on correcting the time-series correlation in volatility [Day and Lewis, 1992, Lamoureux and Lastrapes, 1993, Christensen and Prabhala, 1998]. Unlike these studies I employ a large panel dataset with observations from multiple firms in each quarter and across multiple quarters for each firm. Consequently, my observations exhibit time-series and cross-sectional correlation. To address this, I use two-way industry and quarter clustered standard errors when testing coefficient significance [Gow et al., 2010].

For each firm-quarter, I measure the volatility implied by option closing prices five days after that quarter's earnings announcement date. 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 60 days remaining until expiration to avoid including multiple earnings announcements in the volatility measurement period.[1] The associated realized volatility for each implied volatility estimate is the standard deviation of the underlying equity returns over the remaining 60 days in the option horizon.

Figure 7.2 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 standard deviation

[1] Untabulated results show that using options with 30 days until expiration does not qualitatively change my conclusions.

of daily returns starting at day 5 and ending at day 65. To ensure 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 7.2 shows, this information is typically released prior to day -60, 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, I use the date available. If both I/B/E/S and Compustat are missing announcement dates, I eliminate the observation from my sample. Following Barth and So [2013] I adjust the announcement date one trading day forward when the announcement occurs after the market close.

I use two methods to estimate implied volatility: a modified version of the Black-Scholes formula that allows for dividend payments and early exercise and the model-free method derived by Britten-Jones and Neuberger [2000]. Each Black-Scholes implied volatility estimate is calculated using the price of an at-the-money call option. Prior research documents significant variation in Black-Scholes implied volatility estimates for options on the same underlying asset and time to maturity but with different moneyness levels. This By restricting my analysis to at-the-money options, I effectively hold constant the moneyness of the option contracts in my sample. This generates a sample of comparable option securities in which there is a single Black-Scholes implied volatility estimate for each firm quarter.

I calculate model-free implied volatility using the following equation:

$$\hat{\sigma}^2_{MFIV} = \frac{2e^{rT}}{T}\left[\sum_{i=1}^{S}\frac{\Delta K_i}{K_i^2}P_T(K_i) + \sum_{i=S}^{M}\frac{\Delta K_i}{K_i^2}C_T(K_i)\right] \tag{3.2}$$

The above equation is an approximation of the integral expression for model-free implied volatility from section 2.2. In this approximation, $C_T(K)$ ($P_T(K)$) is the price of a call (put) option with strike price K and remaining time to maturity T. r is the risk free rate of return and M denotes the number of options used in the summation. Appendix section 8.1 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. The original integral equation requires option price observations for a continuum of strike prices, but regulations on most options exchanges prevent the trading of options with very high or low strike prices. 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.

Hypothesis 1 predicts that accounting-based fundamentals are associated with future equity returns volatility. I test this hypothesis by estimating the following equation for each of the eight variables discussed in section 2.1.2:

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_2 V\_FSV^{j}_{it} + \gamma Year + \epsilon_{it} \tag{3.3}$$

I estimate equation (3.3) for eight financial statement variables: size, equity book-to-market ratio, cash flow volatility, leverage, return on assets, research and development expenditure, earnings opacity, and sales growth. At the implied volatility measurement date, the market will have access to quarter t earnings but will only have cash flow and asset or liability amounts as of the end of quarter $t-1$. For simplicity, I ignore the availability of more recent earnings information and measure all variables using only information from the end of quarter $t-1$.

In equation (3.3), $V\_FSV^{j}_{it}$ is an indicator variable that equals one if the level of fundamental variable j is above the industry median for fiscal quarter t and zero otherwise.[2] Industries are defined using the Fama and French [1997] 48 industry classifications. There is an indicator variable for each of the eight fundamental variables in my analysis. V_RND_{it} is an indicator equalling 1 when the level of R&D expenditure reported by firm i in quarter t exceeds firm i's industry median R&D expenditure for quarter t. Analogous indicators are defined for sales growth (V_SGI_{it}), cash flow volatility ($V\_\sigma_{it}^{CF}$), earnings opacity ($V\_Opacity_{it}$), equity book-to-market ratio (V_BTM_{it}), size (V_Size_{it}), return on assets (V_ROA_{it}), and leverage (V_Lvg_{it}). For each of these indicators, an estimate of β_2 that is significantly different from zero supports my first hypothesis that the particular variable is significantly correlated with future returns volatility.

I measure firm size ($Size$) as total assets at the end of quarter $t-1$ and leverage (Lvg) as the ratio of total liabilities to total assets at the end of quarter $t-1$. Cash flow volatility (σ_{it}^{CF}) for firm i in quarter t is the standard deviation of operating cash flows scaled by total assets over the 10 quarters prior to (and *not* including) quarter t. Return on assets (ROA) for firm i in quarter t is the average of the ratio of earnings before interest divided by total assets for quarters $t-1$ through $t-4$. Following Chan et al. [2001] and Beneish et al. [2001], I define R&D expenditure (RND) for quarter t as the ratio of research and development expense(assumed to be zero if not reported) to total assets, both measured at the end of quarter $t-1$. In calculating equity book-to-market ratio (BTM), I estimate book value of equity as the difference between total assets and total liabilities at the end of quarter $t-1$. Since market equity values are observable daily, I measure market value of equity for firm i in quarter t using the closing stock price on the day before the quarter t earnings announcement date.

[2] I explore the sensitivity of my results to this variable construction in section 5.1

3.2 Testing hypothesis 2

To test my second hypothesis, I combine equations (3.1) and (3.3) into:

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_1 \sigma^{IV}_{i,t,\tau} + \beta_2 V\_FSV^{j}_{it} + \beta_3 Spread_{it} + \beta_4 \sigma^{RV}_{t-1} + \gamma Year + \epsilon_{it} \tag{3.4}$$

In equation (3.4) I include controls for past volatility, $\sigma^{RV}_{i,t-1,\tau}$, and liquidity, *Spread*. I measure past volatility for firm i in quarter t as the standard deviation of returns for firm i over the 60 days prior to the earnings announcement date for quarter t. $Spread_{it}$ 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 quarter t 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 et al., 2011]. Equation (3.4) allows me to test the *incremental* informativeness of each financial statement variable after controlling for implied volatility. Fully efficient options markets would imply that the coefficient β_1 (β_2) is indistinguishable from one (zero). My second hypothesis predicts the opposite; if accounting-based fundamental information explains future volatility incremental to the other variables, then β_2 will be non-zero.

I also estimate equations (3.3) and (3.4) using a summary variable *VScore*, which is the sum of indicator variables generated from each fundamental variable:

$$VScore_{it} = V\_RND_{it} + V\_SGI_{it} + V\_\sigma^{CF}_{it} + V\_Opacity_{it} + V\_BTM_{it}$$

$$+(1 - V\_Size_{it}) + (1 - V\_ROA_{it}) + (1 - V\_Lvg_{it})$$

In the calculation of *VScore*, I subtract the indicators for size, ROA, and leverage to generate a new indicator that equals 1 when the firm-quarter observation of size, ROA, or leverage is *below* the industry median for the quarter and zero otherwise.

This modification ensures that there will be a positive relation between the variable $VScore$ and expected equity returns volatility. Since $VScore$ is constructed to be increasing in equity volatility, I predict the coefficient on $VScore$ will be positive and significant.

Hypothesis 2 also implies the possibility of using fundamentals 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

$$V^{straddle} = |S - K| - P \tag{3.5}$$

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 (3.5) and the closing price of the 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 less than ask prices to minimize potential recording errors [Goyal and Saretto, 2009]. I pick the option pair closest to being at-the-money for each firm-quarter. I first restrict my analysis to options with 60 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 explore the ability of $VScore$ to predict straddle returns in two ways. First, I sort firms by $VScore$ and construct a hedge portfolio by taking a long position in straddles for firms with high $VScore$ and a short position in straddles for firms with low $VScore$. I estimate these returns under two sets of assumptions. First, I assume that all trades occur at the mid-point of bid and ask prices. This is a common assumption in equity market research; however, it is a less realistic assumption in options markets because options markets are characterized by fewer participants and lower liquidity than equity markets. The inability of investors to trade at the midpoint of the bid and ask prices lowers the profitability of the trading strategy, since investors are forced to buy (write) options at higher (lower) prices.

Taking this into consideration, I also calculate straddle returns assuming 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 [Dennis and Mayhew, 2002, De Fontnouvelle et al., 2003, Battalio et al., 2004]. Using OptionMetrics' bid-ask spread estimates consequently 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 is not available through OptionMetrics. Under both return measurement methods, my hypothesis that $VScore$ can predict option returns implies that the returns to the $VScore$ hedge portfolio are positive.

Goyal and Saretto [2009] reveal that straddle returns are predictable using the difference between historical and implied volatility (*DiffVol*). I replicate their result and use a double sort to examine how $VScore$ interacts with *DiffVol* in predicting straddle returns. I use a nested double sort for this analysis. First, I separate firms by $VScore$. Then, within each $VScore$ class, I calculate decile cutoffs for *DiffVol* and sort forms into these deciles. I then estimate mean returns within each $VScore$-*DiffVol* pair category. In addition to sorting, I use quarterly returns regressions to explore the incremental informativeness of $VScore$ for assessing straddle returns. Specifically, I estimate

$$r^s_{i,t} = \alpha + \beta_1 \text{Rank}\sigma^{IV}_{i,t,\tau} + \beta_2 \text{Rank}\sigma^{RV}_{i,t-1,\tau} + \beta_3 VScore + \epsilon_{it} \quad (3.6)$$

where $r^s_{i,t}$ is the 60-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 address cross-sectional correlation in returns by first estimating equation (3.6) 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. Hypothesis 2 predicts that β_3 from equation (3.6) are significantly positive.

3.3 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 US equities.[3] In addition to reported prices, OptionMetrics provides several summary statistics for standardized 30- and 60-day call options. OptionMetrics's standardization procedure effectively generates Black-Scholes implied volatility estimates for at-the-money options of constant duration [Barth and So, 2013]. I conduct my analysis using the standardized implied volatility of 60-day options, although untabulated results indicate that the conclusions are unaffected by the use of 30-day options. I obtain accounting data from Compustat and

[3] Options in the United States trade on one of four exchanges: the Chicago Board Options Exchange (CBOE), the American Stock Exchange (AMEX), the Philadelphia Stock Exchange (PHLX), and the Pacific Exchange. These markets differ in structure; the CBOE and the Pacific Exchange feature an open-outcry structure. In contrast, the AMEX and PHLX organize options trade through specialists. Although standard in the volatility literature, the use of only CBOE data limits the generalizability of my findings to the extent that differences in the microstructure of the CBOE and other options exchanges affects pricing.

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 78,034 observations from 3,934 firms from 1996 to 2010. The subsample of observations for which model-free implied volatility is measurable consists of 7,275 firm quarter observations from 1,817 firms.

Panel A (B) of Figure 7.1 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 log-normal. Therefore, I conduct my analysis using the log-series of both implied and realized volatility. Tables 7.1 and 7.2 provide details on the composition of my sample by industry and year. Table 7.1 provides a description of my whole sample; table 2 focuses on the subsample of firms for which I can calculate model-free implied volatility as well. The dominant industries in my full sample are business equipment, healthcare, manufacturing, and a generic other category that includes mining, construction, and entertainment. A comparison of panels A across tables 7.1 and 7.2 reveals that the samples are cross-sectionally similar despite having different magnitudes; this alleviates potential concerns about selection bias in the estimation of model-free implied volatility. Panel B of table 7.1 reveals a general increase in the number of observations per year over time (from 1,917 firm quarter observations in 1996 to 7,772 firm quarter observations in 2009), which is consistent with the increase in options trade over the past decade. The decline in observations in 2010 reflects data availability constraints. Panel B of table 7.2 shows marked increases in the availability of model-free implied volatility estimates in 2000 and 2008. This is the result of more firms having sufficiently many traded options to facilitate the calculation of model-free implied volatility. As previously discussed, I only calculate model-free implied volatility when there are sufficiently many traded options with unique strike prices to mitigate concern about truncation error. The increase in traded options in 2000 and 2008 is likely a consequence of higher market volatility during these periods, since higher market volatility increases demand for options.

Table 7.3 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 7.3 reveal that, on average, implied volatility is higher than realized volatility. The mean and median of the logarithm of implied volatility are -3.58 and -3.59 and are higher than the mean and median of future realized volatility (-3.65 and -3.66). The distribution of model-free implied volatility reveals that model-free implied volatility is consistently lower than both implied and realized volatility. The logarithm of model-free implied volatility in my sample has a mean of -5.07 and a standard deviation of 0.45. The distribution of $VScore$ shows that $VScore$ has a mean of 3.72 and does not always equal zero or eight for each firm; rather, there is cross-sectional variation in its value. Table 7.4 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 Pearson and Spearman correlations of 0.79. Future realized volatility is also significantly positively correlated with current implied volatility.

Chapter 4

Empirical tests

4.1 Fundamentals and future equity volatility

Table 7.5 presents summary statistics from the estimation of equation (3.3), which is designed to test hypothesis 1 using each of the financial statement variables discussed in section 2.1.2. With the exception of leverage, each variable exhibits the predicted relation with future equity volatility. Table 7.5 reveals that volatility is significantly negatively related to size (coefficient = -0.256, t-statistic = -10.30), leverage (coefficient = -0.076, t-statistic = -4.10), and ROA (coefficient = -0.155, t-statistic = -6.34). Each of these coefficients is significantly different from zero. Table 7.5 also reveals that future volatility is significantly and positively related to R&D expenditure (coefficient = 0.231, t-statistic = 5.83), cash flow volatility (coefficient = 0.235, t-statistic = 14.04), earnings opacity (coefficient = 0.111, t-statistic = 7.12), and equity book-to-market ratio (coefficient = 0.053, t-statistic = 2.89). Because the dependent variable in these estimations is the logarithm of realized volatility, the coefficients have a multiplicative interpretation. For instance, the coefficient -0.256 on V_Size indicates that firms with total assets that are above their industry medians have 22% lower volatility than firms with assets below their industry medians. The significantly negative coefficient on leverage in table 7.5 reveals that firms with high leverage relative to the industry median exhibit lower volatility. This is consistent with the alternative relation posited in section 2.1.2 that financially stable firms have greater access to debt financing and consequently exhibit higher leverage. Overall, the results from table 7.5 support my first hypothesis that accounting-based fundamental information is associated with future equity volatility.

4.2 Fundamentals and the bias in option-implied volatility

4.2.1 Measuring expected volatility with Black-Scholes implied volatility

Table 7.6 presents coefficient estimates from equation (3.4), which is designed to test my second hypothesis. Hypothesis 2 has two testable implications for equation (3.4); the coefficients β_1 and β_2 should be significantly different from one and zero, respectively. The results in table 7.6 support both of my predictions. Estimates of β_1 range from 0.652 to 0.661 across all of the models and are consistently significantly less from one with t-statistics ranging from -18.05 to -17.81. These results confirm prior findings that the Black-Scholes implied volatility is a biased estimator of future volatility.

Of the eight variables I examine, seven have coefficients that are significantly nonzero. The coefficients on size, leverage, ROA, and BTM ratio are negative, indicating that the market overestimates volatility for firms that are larger, more profitable, more levered, or more undervalued by the market than the industry median. Size has a coefficient of -0.022 (t-statistic = -3.16), ROA has a coefficient of -0.015 (t-statistic

= -3.20), leverage has a coefficient of -0.009 (t-statistic = -2.57) and equity book-to-market ratio has a coefficient of -0.011 (t-statistic = -2.08). Again, these coefficients have a multiplicative interpretation because the dependent variable in equation (3.4) is the logarithm of future realized equity volatility. The coefficients on sales growth, cash flow volatility, and earnings opacity are positive, suggesting that the market underestimates volatility for firms with more growth opportunities, as measured by sales growth, or more unpredictable operations, as measured by earnings volatility and opacity. Sales growth has a coefficient of 0.010 (t-statistic = 2.91), cash flow volatility has a coefficient of 0.018 (t-statistic = 3.07), and earnings opacity has a coefficient of 0.010 (t-statistic = 2.90). Consistent with the market being able to better anticipate managerial discretion than fundamental variability in operations, the coefficient on cash flow volatility, 0.018, is larger than that of opacity, 0.010. Overall, the significant non-zero β_2 estimates for these seven variables suggest that the market does not efficiently process the information the variables reflect.[1]

Table 7.7 presents summary statistics from the estimation of equations (3.1) - (3.4), first including only implied volatility as a predictor of future volatility and then adding $VScore$ to capture firm characteristics. Column I reveals that implied volatility alone explains 70.5% of the variation in future realized volatility. However, consistent with prior research, the coefficient on implied volatility, 0.896, is significantly different from one (t-statistic = -5.80). Columns II and III add past realized volatility and spread, respectively, to the model in column I. The results in column II reveal that past volatility is incrementally significant in explaining future volatility with a coefficient of 0.251 (t-statistic = 10.36). This suggests that implied volatility is not an efficient estimator of future volatility. Unlike Canina and Figlewski [1993] but consistent with Christensen and Prabhala [1998], I find that, after controlling

[1] 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 5.2.

for past volatility, implied volatility is significantly correlated with future volatility. However, the coefficient on implied volatility, 0.655, is significantly different from one (t-statistic = -18.30), indicating a persistent bias.

Column III of table 7.7 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 increases only 0.1%, from 71.7% in column II to 71.8% in column III. *Spread* has a significantly negative coefficient of -0.019 (t-statistic = -2.46) that reflects the inverse relation between equity market liquidity and options prices. Column IV of table 7.7 presents summary statistics from the estimation of equation (3.4) with $VScore$ as a summary of financial statement information. Consistent with my hypotheses, the coefficient on $VScore$ is significantly positive (t-statistic = 3.95). The coefficient estimate of 0.011 indicates that the market on average underestimates volatility at the rate of 1% per $VScore$ unit. In other words, a firm with a $VScore$ of x (where x ranges from 0 to 8) has, on average, an implied volatility that is x% lower than the subsequent realized volatility. A comparison of R^2 values in columns III and IV reveals that adding $VScore$ to the model improves the overall explanatory power of the model by 3%, from 71.8% to 74.8%.

4.2.2 Measuring expected volatility with model-free implied volatility

Tables 7.6 and 7.7 provide evidence that the Black-Scholes implied volatility is a biased estimator of future volatility. However, they do not address whether this bias is a result of measurement error or reflects inaccurate market expectations. Table 7.8 presents coefficient estimates from the regression of equation (3.4) using model-free implied volatility in place of Black-Scholes implied volatility as the market's expectation of future volatility. The results confirm that using model-free implied volatility as a measure of the market's expectation of future volatility does not change the inferences obtained from tables 7.6 and 7.7. Column I of table 7.8 provides results of regressing future realized volatility on model free implied volatility and past realized volatility without additional controls for fundamentals. The coefficient on model-free implied volatility in this estimation is 0.651 and is significantly different from one at the 1% level (t-statistic -19.16). Of the eight financial statement variables, three have coefficients that are significantly negative. Size has a coefficient of -0.045 (t-statistic = -3.84), leverage has a coefficient of -0.014 (t-statistic = -1.77) , and BTM has a coefficient of -0.035 (t-statistic = -3.98). Another three variables have significantly positive coefficients; R&D expenditure has a coefficient of 0.020 (t-statistic = 1.76), cash flow volatility has a coefficient of 0.025 (t-statistic = 2.95), and sales growth has a coefficient of 0.014 (t-statistic = 1.69). All coefficients have the same sign in the model-free specification as they do in the Black Scholes specification. The coefficients on opacity, 0.007, and ROA, 0.001, are not significant (t-statistics of 0.80 and -0.09, respectively) when using model-free implied volatility as a measure of the market's volatility expectation despite being significantly different from zero when using Black-Scholes implied volatility as a measure of the market's volatility expectation. This loss of increment significance in the model-free implied volatility specification suggests that the statistical significance of opacity and ROA might be partially driven by measurement error in Black-Scholes implied volatility rather than expectations errors. In contrast, the persistent significance of the remaining six variables suggests measurement error is not the sole driver of their explanatory power over future realized volatility. Overall, the results in table 7.8 reinforces the conclusions drawn from tables 7.6 and 7.7 by alleviating concern that the incremental significance of the fundamentals in those analyses are an artifact of measurement error in the market's expectation of future volatility.

4.3 Predicting straddle returns using fundamentals

The persistent significance of accounting-based fundamentals, incremental to option-implied volatility, in the prediction of future realized volatility suggests that there is inadequate incorporation of this information in options prices. This in turn suggests that financial statement information could be useful in the prediction of option returns. I explore this implication in several ways. Table 7.9 presents straddle portfolio returns from a fundamental analysis trading strategy. Panels A and B provide straddle returns by decile based on my volatility score ($VScore$) under two different assumptions about transactions costs. Panel A assumes all trades occur at the mid-point of the closing bid and ask prices. Under this assumption, the return to a hedge portfolio based on $VScore$ is 12.9% (t-statistic = 2.52). This is higher than the 8.9% return generated by the fundamental strategy of Goodman et al. [2013]. However, as discussed in section 3.2, transactions costs in the options market are often substantial and must be considered when attempting to form trading strategies. The results in panel B of table provide straddles returns to a $VScore$ strategy under the assumption of a 100% effective spread; in other words, I assume in this alternative that straddle investors must buy at the ask price and sell at the bid price. This modification to the returns calculation reduces the overall hedge return from 12.9% to 3.5%. Though substantially reduced, the overall hedge return is still positive, suggesting that there is some amount of mispricing in traded options that outside the range of transactions costs. Untabulated results reveal that the positive return to the $VScore$ strategy is not driven by any single component of the variable.

Panel A of table 7.10 presents straddle portfolio returns by decile of the Goyal and Saretto [2009] *DiffVol* variable. In my sample, the hedge return to a *DiffVol* strategy is 17.8% return (t-statistic = 6.60). The *DiffVol* return in my sample is lower than that documented by Goyal and Saretto [2009] and Goodman et al. [2013], a difference that most likely reflects differences in sample construction attributable to data requirements. Panel B of table 7.10 presents straddle returns under a double sort by *VScore* and *DiffVol* deciles. In all *DiffVol* deciles, firms with the highest *VScore*s exhibit strictly non-negative returns. In four *DiffVol* deciles, there are significant positive returns to the *VScore* hedge strategy. A *VScore* hedge generates 14.8% return in *DiffVol* decile 1 (t-statistic = 2.04), a 28.1% return in *DiffVol* decile 4 (t-statistic = 2.29), a 16% return in *DiffVol* decile 6 (t-statistic = 2.10), and a 25.7% return in *DiffVol* decile 7 (t-statistic = 2.49). In the other six deciles, the *VScore* hedge generates returns indistinguishable from zero. Panel C also reveals that the returns to a *DiffVol* hedge portfolio are non-negative across all *VScore* levels. When *VScore* is equal to 0, 5, 6, or 8, *DiffVol* hedge returns are 25.4% (t-statistic = 3.37) 18.9% (t-statistic = 2.06), 18.6% (t-statistic = 2.43), and 10.7% (t-statistic = 1.76), respectively. Each of these returns is significantly different from zero. When *VScore* is outside this range, *DiffVol* hedge returns are still positive on average, but the return cannot be distinguished from zero statistically. These results suggest that the returns to a *VScore* strategy are not subsumed by a *DiffVol* strategy.

To address concern that the straddle returns are driven by a single anomalous year, Figure 7.3 provides a graph of *VScore* hedge returns for each year of my sample. The graph reveals that in the 10 of the 14 years of my sample, the hedge return from a *VScore* based strategy is positive. In three of the ten years the average return exceeds 20% and in eight of the ten years it exceeds 10%. Of the four years in which the hedge return is negative, only in the first year, 1996, is the return significantly different from zero. The results of figure 7.3 reveal that the profitability of a *VScore* trading strategy is relatively persistent over time, providing assurance that the pooled sample results from table 7.9 are not driven by a single anomalous year.

As an additional robustness check of the incremental informativeness of *VScore* for assessing straddle returns, I estimate equation (3.6) from section 3.2. Table 7.11 presents coefficient estimates from equation (3.6), which relates straddle returns to the rank of implied and past realized volatility and $VScore$. My second hypothesis predicts that the coefficient on $VScore$ will be significantly positive. The results in table 7.11 reveals that the coefficient on implied volatility ranges from -0.014 to -0.024 across columns I through III and is consistently significantly negative (t-statistics range from -2.96 to -4.96), suggesting that higher implied volatility is associated with negative option returns. This result is consistent with prior literature documenting a negative relation between idiosyncratic volatility and equity returns [Ang et al., 2006]. Columns II and III show that the coefficient on the rank of past realized volatility is not distinguishable from zero after controlling for implied volatility (t-statistic = 1.41). Consistent with my second hypothesis, Column III reveals a significant positive coefficient of 0.021 on $VScore$ (t-statistic = 2.85). This implies that a one-point increase in $VScore$ is associated with a 2% increase in straddle returns and is consistent with the hypothesis that $VScore$ is incrementally useful in the prediction of future volatility and, consequently, option returns.

Chapter 5

Robustness and additional analyses

5.1 Alternative variable measurement

In my main analyses, the fundamental variables of interest are measured as indicators, equalling one if the level of the underlying continuous fundamental variable is above the industry median for fiscal quarter t and zero otherwise. This approach has several benefits: it provides a simple framework for controlling for industry-quarter fixed effects and also facilitates the construction of the volatility score which is used for sorting option straddle portfolios. Since the indicator variables suppress variation in the underlying explanatory variables, it is likely that this research design biases against my hypotheses by reducing the likelihood of observing significant covariance between the fundamental variables I identify and future equity volatility. Nonetheless, to ensure that my results are not specific to this design choice I re-estimate the main equations of my empirical tests with continuous versions of the eight fundamental variables in my analyses. Since industry-quarter effects are still a concern, I normalize each variable within industry quarters.

Table 7.12 presents summary statistics from the estimation of equation (3.3) using continuous versions of each of the financial statement variables. All variables are significantly different from zero and each variable exhibits the same directional relation with volatility as in the indicator specifications. Table 7.12 reveals that volatility is significantly negatively related to size (coefficient = -0.107, t-statistic = -10.44), leverage (coefficient = -0.037, t-statistic = -3.55), and ROA (coefficient = -0.085, t-statistic = -9.00). Table 7.12 also reveals that future volatility is significantly and positively related to R&D expenditure (coefficient = 0.006, t-statistic = 5.83), cash flow volatility (coefficient = 0.235, t-statistic = 14.04), earnings opacity (coefficient = 0.111, t-statistic = 7.12), and equity book-to-market ratio (coefficient = 0.053, t-statistic = 2.89). In comparison to the results using indicators, each of these coefficients have lower coefficient magnitudes but the same level of statistical significance. Overall, the results from table 7.12 further support my first hypothesis that accounting-based fundamental information is associated with future equity volatility. The results also confirm that this observed relation is not specific to the indicator variable specification.

Table 7.13 presents coefficient estimates from equation (3.4) using continuous versions of each of the financial statement variables. Relative to the estimates using indicator variables, the results in table 7.13 provide even stronger support for my second hypothesis. All eight variables have coefficients that are significantly non-zero when measured in continuous form. The coefficients on size, leverage, ROA, and BTM ratio are negative, indicating that the market overestimates volatility for firms that are larger, more profitable, more levered, or more undervalued by the market than the industry median. Size has a coefficient of -0.007 (t-statistic = -2.66), ROA has a coefficient of -0.008 (t-statistic = -3.09), leverage has a coefficient of -0.004 (t-statistic = -1.97) and equity book-to-market ratio has a coefficient of -0.004 (t-statistic = -1.68). Sales growth has a coefficient of 0.004 (t-statistic = 3.35), cash flow volatility has a coefficient of 0.008 (t-statistic = 3.34), and earnings opacity has a coefficient of 0.004 (t-statistic = 2.531). In contrast to the indicator specification, the variable

R&D expenditure as a significantly positive coefficient of 0.006 (t-statistic = 2.60) when measured in continuous form. Again, the results of table 7.13 further confirm that my inferences are not specific to the indicator variable specification.

Since measurement error in implied volatility could cloud my inferences from any estimations made using Black-Scholes implied volatility, I also reestimate my tests of hypothesis 2 with model-free implied volatility using the continuous fundamental variable measurements. As in the previous specifications, using continuous versions of the fundamental variables in conjunction with model-free implied volatility results in relatively stronger empirical results than when using indicator variables. All six variables that had coefficients significantly different from zero continue to do so in this modified estimation. Three have coefficients that are significantly negative. Size has a coefficient of -0.015 (t-statistic = -3.32), leverage has a coefficient of -0.007 (t-statistic = -1.91) , and BTM has a coefficient of -0.020 (t-statistic = -4.28). Another three variables have significantly positive coefficients; R&D expenditure has a coefficient of 0.09 (t-statistic = 2.00), cash flow volatility has a coefficient of 0.015 (t-statistic = 3.10), and sales growth has a coefficient of 0.009 (t-statistic = 1.93). All coefficients also have the same sign in the continuous specification as they do in the indicator specification. Additionally, using continuous versions of the fundamental variables results in higher overall explanatory power when using all eight fundamental variables simultaneously (see column IX). In the continuous version summarized in table 7.14, the adjusted R-square in column IX is 71%, while in the indicator version summarized in table 7.8 it is 66%. Overall, the results of using continuous versions of fundamentals is consistent with the idea that the original use of indicator variables suppresses variation in the explanatory variables and biases my results against confirming my hypotheses. Using continuous versions instead of the indicator variables does not diminish the statistical significance of my results or materially change my inferences.

5.2 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. However, prior research identifies predictability in equity returns that is related to accounting-based fundamentals such as firm size, equity book-to-market ratio, and the level of accruals. If options traders use a price that adjusts for the predictability in equity returns, implied volatility will differ from the option market's expectation of future volatility and the gap between implied volatility and realized volatility will be a consequence of the observed mispricing in the first moment of equity returns. Moreover, if this adjustment is based on the same accounting-based fundamentals that I identify in my analyses, the observed relationship between these fundamental variables and the gap between implied and realized volatility would be a mechanical relation rather than an indication of informational inefficiencies.

However, this scenario is unlikely to be pervasive in my sample. Were there to be a systematic use of an equity price other than the current market price (such as one based on fundamentals), we would observe frequent and persistent violations of put-call parity. For European options on nondividend paying stocks, the no-arbitrage condition implies exact put-call parity. For American options, Merton (1973) shows that the puts will be more valuable because at every point in time there is a positive probability of early exercise. This early exercise premium (EEP) in put prices will create a gap between American option call and put prices, even in the absence of short sale constraints or microstructure effects. Empirical evidence on the relative pricing of puts and calls reveals that deviations from put-call parity are rare and typically temporary [Cremers and Weinbaum, 2010].

Nonetheless, I estimate the volatility spread for firm i on day t as follows:

$$VS_{it} = \sum_j w_j (IV^C_{i,j,t} - IV^P_{i,j,t})$$

In the above equation, w_j is a weighting for option pair j equal to the relative open interest on the option pair for firm i on day t. j is the number of option pairs (a pair consists of a call and put option on the same asset with the same maturity) for firm i on on day t. I include volatility spread in a re-estimation of equation (3.3) to examine the incremental informativeness of the eight fundamental variables. Volatility spread captures option market perceptions of equity mispricing. If option market perceptions of equity misplacing are driving my primary findings, then including a volatility spread control variable should eliminate the incremental statistical significance of the eight fundamental variables.

Panel B of table 7.15 presents the summary statistics from the modified equation (3.3). Consistent with my predictions, there are significant correlations between future realized volatility and seven of the eight fundamentals I examine, 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. The results of table 7.15 suggest that implied volatility is overstated for large, highly levered, profitable, and relatively undervalued firms, as the coefficients on size, leverage, return-on-assets, and book-to-market ratio are positive and significantly different from zero. Size has a coefficient of -0.020 (t-statistic = -2.770), leverage has a coefficient of -0.009 (t-statistic = -2.668), return on assets has a coefficient of -0.012 (t-statistic = -2.651), and book to market ratio has a coefficient of -0.013 (t-statistic =-2.264). Table 7.15 also reveals significantly positive coefficients on sales growth, earnings opacity, and cash flow volatility. Sales growth as a coefficient of 0.009 (t-statistic = 2.909), opacity has a coefficient of

0.009 (t-statistic = 2.607), and cash flow volatility has a coefficient of 0.018 (t-statistic = 3.038). These results suggest that implied volatility is understated for firms with high sales growth, opaque earnings, or volatile cash flows.

I also examine whether the eight fundamental variables can predict option volatility spread. Untabulated results reveal that they cannot; no single variable or combination of variables exhibiting more than 1 percent or lower adjusted R^2 values. The low adjusted R^2 values indicates that fundamentals explain very little of the variation in the volatility spread. If option traders were valuing options using an adjusted equity price, the basis for the adjustments would necessarily be predictive of the volatility spread. Since I find that it is not, it is unlikely that options traders are systematically using these fundamentals to adjust the prevailing market price of equity in their valuation of options.

5.3 Variance risk premia

An alternate explanation for the result that fundamentals are incremental to implied volatility in the prediction of future volatility is that fundamentals capture equity variance risk premia. Variance risk for an asset refers to uncertainty about the variance of the asset's returns. Option-implied volatility estimates are constructed under a risk-neutral measure, but options market investors are likely risk-averse. Like risk in the first moment, this uncertainty about the variance of equity could drive risk-averse investors to demand a premium for holding risky assets to the extent that uncertainty in the variance of the asset is correlated with uncertainty in the variance of the market portfolio. Since volatility is the only degree of freedom in most option-pricing models, this premium would be reflected as a gap between implied and future realized volatilities. A larger (smaller) implied volatility than the corresponding future realized volatility is consistent with a positive (negative) variance risk premium, which in turn implies a positive (negative) correlation between equity variance and market variance.

Prior research shows that, at the index level, implied volatility is consistently greater than future realized volatility. Similarly, roughly two-thirds (48,381) of observations in my sample of firm-level options feature implied volatility greater than future realized volatility. This empirical trend is consistent with a positive variance risk premium. However, in the remaining one third (29,653) of firm-quarters, implied volatility is less than future realized volatility. If the gap between implied and future realized volatility is driven by a variance risk premium, it must be the case that the firm's equity return variance and the market return variance are negatively correlated if the option-implied volatility is lower than the subsequent realized volatility.

Figure 7.4 presents time-series plots of the realized volatility of both the market return and equity returns for the firms in my sample. The realized volatility of S&P 500 index returns, which I use as a proxy for market portfolio returns, is represented by the green solid line. Consistent with prior research, the S&P 500 index exhibits less volatility than the individual firms in my sample. The red dotted (blue dashed) line plots the average future realized volatility of equity returns for each quarter for firms whose option-implied volatility is less than (greater than or equal to) the future realized volatility. Firms for which option-implied volatility is lower than realized volatility exhibit consistently higher realized volatility than firms for which option-implied volatility is higher than realized volatility. However, both sets of firms exhibit returns variances with strong positive co-movement with the market volatility series. In particular, the time series of realized volatility for firms whose implied volatility is less than realized volatility does not appear to move against the market volatility series. This suggests that the difference between implied volatility and future realized volatility for these firms is not entirely attributable to a variance risk premium.

The results from figure 7.4 imply that subsample of firms with lower option-implied volatility than future volatility offers a powerful setting in which to test my hypotheses with less concern about variance risk premia confounding the interpretation of results. Table 7.16 presents summary statistics from the estimation of equation (3.4) on the subsample of firms for which option-implied volatility is less than future realized volatility. Consistent with my predictions, there are significant correlations between future realized volatility and six of the eight fundamentals I examine, even in the subsample of firm-quarters where variance risk premia are less evident. The results of table 7.16 suggest that implied volatility is overstated for large, profitable, and relatively undervalued firms, as the coefficients on size, return-on-assets, and book-to-market ratio are positive and significantly different from zero. Size has a coefficient of -0.015 (t-statistic = -2.481), return on assets has a coefficient of -0.012 (t-statistic = -2.449), and book to market ratio has a coefficient of -0.011 (t-statistic =-2.380). In this subsample of firms, unlike the full sample, implied volatility also appears to be overstated for firms with high research and development expenditures. R&D expense has a coefficient of -0.34 (t-statistic = -4.86). Table 7.16 also reveals significantly positive coefficients on sales growth and cash flow volatility. Sales growth as a coefficient of 0.008 (t-statistic = 2.675) and cash flow volatility has a coefficient of 0.012 (t-statistic = 2.807). These results suggest that implied volatility is understated for firms with high sales growth and volatile cash flows. Overall, the results in table 7.16 confirm that the possible existence of a variance risk premium does not fully account for the incremental significance of fundamental variables relative to implied volatility in the prediction of future volatility.

Chapter 6

Conclusions

I provide evidence that accounting-based fundamental information is useful in the prediction of equity volatility incremental to option-implied volatility. Prior research establishes that implied volatility is a biased estimator of future realized volatility, but the precise cause of the bias has remained unclear. My study is the first to hypothesize and find evidence that information about fundamentals from financial statements is not fully incorporated in either past volatility or the market's expectation of future volatility.

I focus my analysis on accounting information that prior literature shows to be useful in identifying firms with growth opportunities, quantifying default and priced equity risks, and predicting extreme returns. From these literatures I identify the following eight variables: firm size, equity book-to-market ratio, cash flow volatility, earnings opacity, research and development expenditure, sales growth, return on assets, and leverage. Over 20% of total variation in observed equity volatility is explained by variation in these variables. As a benchmark forecast of future volatility, I use the expectation of volatility implied by options prices both under the Black-Scholes formula and the Britten-Jones and Neuberger [2000] model-free approach. Using either of these benchmarks, I show that the financial statement variables I identify can supplement implied volatility in predicting future volatility. 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 contribute to the literature studying variance risk premia. In supplementary analyses, I find that the observed differences between option-implied and realized equity volatilities cannot be fully explained by the existence of a variance risk premium. Specifically, I show that firms with a negative expected variance risk premium (that is, a lower option-implied expected volatility than the corresponding equity volatility realization) do not exhibit negative correlations with the market variance. Moreover, financial statement information can still supplement implied volatility in this subsample. These empirical facts suggest 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.

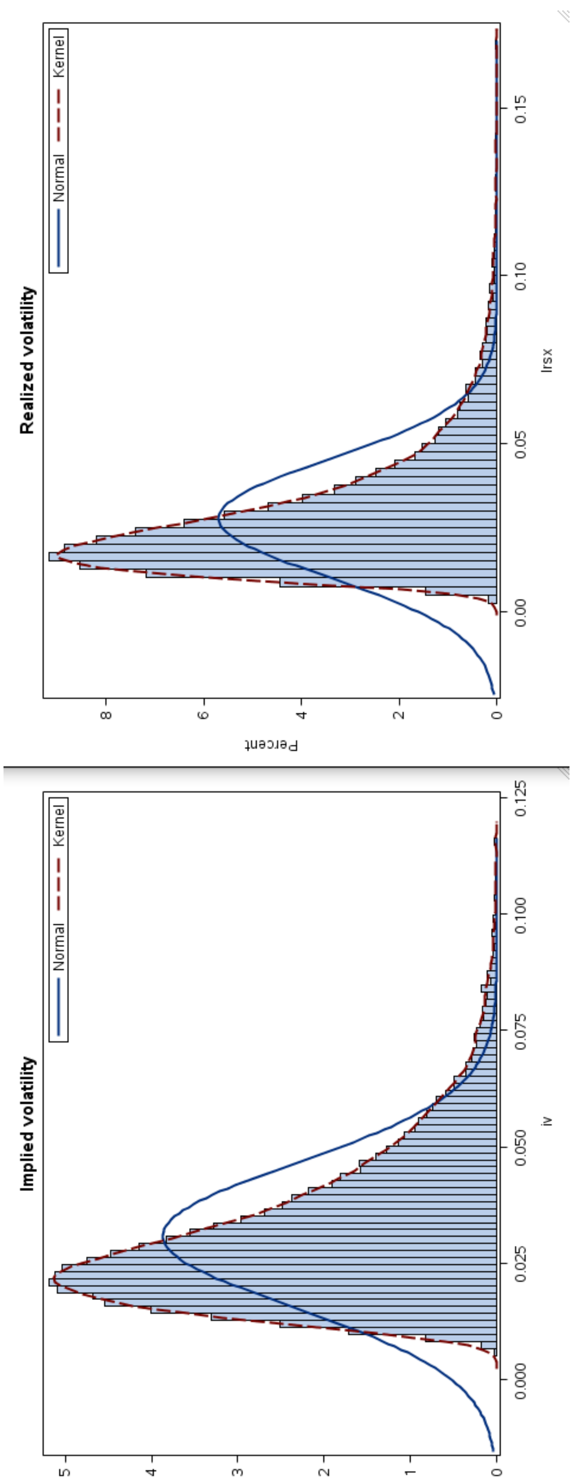
My results also have implications for options pricing, as I show that accounting-based fundamental information can be used to predict option returns. From the eight financial statement variables in my first stage analysis I construct a single summary metric, *VScore*. A trading strategy based on my *VScore* metric generates significantly positive straddle returns. The returns from this strategy are robust across quintiles formed using both book-to-market ratio and the difference between historical and 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 paper 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 bias in option-implied volatility, a long-standing empirical puzzle in the asset pricing literature. My research also improves our understanding of the relative informational efficiency of the options market relative to the equity market. Recent studies document the superior information discovery and processing capabilities of the options market. My results reveals that even with these improvements, options prices appear to reflect informational inefficiencies.

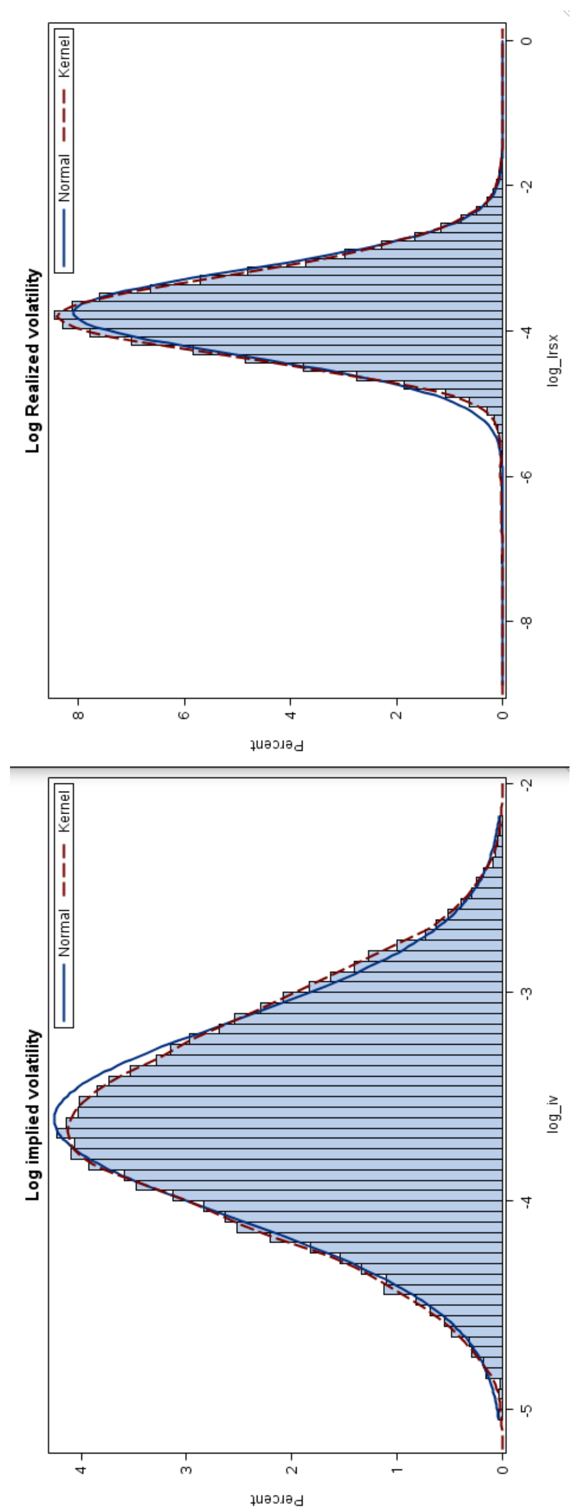
Chapter 7

Tables and figures

Figure 7.1: Density plots of implied and realized volatility



(a) Levels of implied and realized volatility



(b) Logarithms of implied and realized volatility

Figure 7.2: Timeline of events

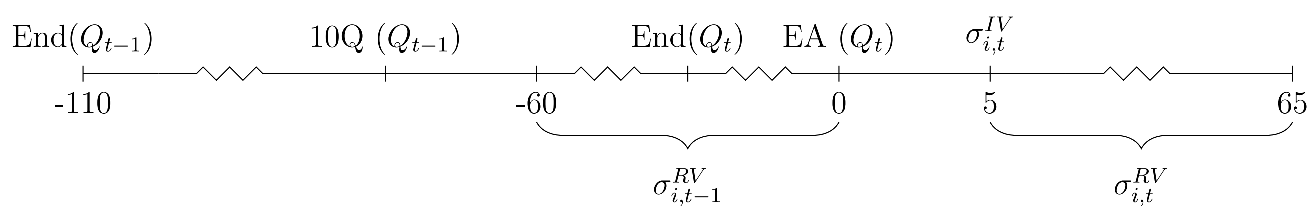


Table 7.1: Sample composition

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 (year); "Percent" indicates the percentage of the total sample attributable to each industry (year). Industries are defined using the 12 Fama-French classifications.

(a) Sample composition by industry

Industry	Observations	Percent
Consumer Nondurables	3948	5.06
Consumer Durables	1691	2.17
Manufacturing	8367	10.72
Energy	4170	5.34
Chemicals	2120	2.72
Business Equipment	17897	22.93
Telecommunications	2236	2.87
Utilities	2025	2.60
Wholesale and Retail Trade	9056	11.61
Healthcare	9160	11.74
Financial Services	6538	8.38
Other	10826	13.87
Total	78034	100.00

(b) Sample composition by year

Year	Observations	Percent
1996	1917	2.46
1997	3326	4.26
1998	3991	5.11
1999	4736	6.07
2000	4571	5.86
2001	4362	5.59
2002	5269	6.75
2003	5251	6.73
2004	5666	7.26
2005	6623	8.49
2006	7045	9.03
2007	7615	9.76
2008	7996	10.25
2009	7772	9.96
2010	1894	2.43
Total	78034	100.00

Table 7.2: Model free implied volatility sample composition

This table provides information on the composition of the subsample of observations for which I can calculate model free implied volatility. Panel A (Panel B) provides a breakdown of the subsample by industry (year). The column “Observations” indicates the number of firm-quarter observations for each industry (year); “Percent” indicates the percentage of the total subsample attributable to each industry (year). Industries are defined using the 12 Fama-French classifications.

(a) Sample composition by industry

Industry	Observations	Percent
Consumer Nondurables	349	4.80
Consumer durables	137	1.88
Manufacturing	932	12.81
Energy	416	5.72
Chemicals	225	3.09
Business Equipment	1854	25.48
Telecommunications	113	1.55
Utilities	67	0.92
Wholesale and Retail Trade	740	10.17
Healthcare	864	11.88
Financial Services	646	8.88
Other	932	12.81
Total	7275	100.00

(b) Sample composition by year

Year	Observations	Percent
1996	129	1.77
1997	415	5.70
1998	558	7.67
1999	733	10.08
2000	862	11.85
2001	499	6.86
2002	304	4.18
2003	311	4.27
2004	434	5.97
2005	508	6.98
2006	452	6.21
2007	626	8.60
2008	1112	15.29
2009	323	4.44
2010	9	0.12
Total	7275	100.00

Table 7.3: Descriptive statistics

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

	N	Mean	Std Dev	Minimum	Q1	Median	Q3	Maximum
$\log \sigma_t^{RV}$	78034	-3.65	0.54	-5.83	-4.02	-3.66	-3.29	-1.90
$\log \sigma_{t-1}^{RV}$	78034	-3.63	0.53	-5.10	-4.00	-3.65	-3.28	-1.99
$\log \sigma_t^{IV}$	78034	-3.58	0.47	-5.07	-3.91	-3.59	-3.26	-2.18
$\log \sigma^{MFIV}$	7275	-5.07	0.45	-7.69	-5.39	-5.09	-4.76	-3.49
Spread	78034	-1.95	0.51	-4.28	-2.29	-1.95	-1.63	0.38
ROA	78034	0.01	0.04	-0.44	0.00	0.01	0.02	0.13
log TA	78034	7.20	1.79	1.12	5.90	7.07	8.34	14.94
RND	78034	0.03	0.06	0.00	0.00	0.00	0.03	0.58
Lvg	78034	0.50	0.25	0.03	0.31	0.51	0.67	1.42
σ_t^{CF}	78034	0.02	0.04	0.00	0.01	0.01	0.02	0.48
Opacity	78034	-0.04	0.18	-0.81	-0.13	-0.04	0.05	0.89
log BTM	78034	-0.92	0.90	-13.77	-1.42	-0.89	-0.39	5.37
$VScore$	78034	3.72	1.59	0.00	3.00	4.00	5.00	8.00
$\mathbb{I}\{\sigma_{i,t+\tau}^{RV} < \sigma_{it}^{IV}\}$	78034	0.62	0.49	0.00	0.00	1.00	1.00	1.00

Table 7.4: Correlation matrices

This table correlation matrices for the main variables in my analysis. Pearson (Spearman) correlations are above (below) the diagonal. Section 3.1 contains a detailed explanation of each variable's construction.

	$\log\sigma_t^{RV}$	$\log\sigma_{t-1}^{RV}$	$\log\sigma_t^{IV}$	TA	Lvg	RND	σ_t^{CF}	Opacity	ROA	BTM	V-score	Spread
$\log\sigma_t^{RV}$	1	0.79	0.82	-0.4	-0.18	0.23	0.29	0.04	-0.24	-0.05	0.29	-0.04
$\log\sigma_{t-1}^{RV}$	0.79	1	0.89	-0.41	-0.19	0.23	0.3	0.03	-0.25	-0.08	0.29	-0.05
$\log\sigma_t^{IV}$	0.82	0.88	1	-0.5	-0.22	0.3	0.35	0.05	-0.31	-0.08	0.34	-0.01
TA	-0.42	-0.43	-0.51	1	0.52	-0.38	-0.34	0.02	0.18	-0.23	-0.52	-0.3
Lvg	-0.2	-0.21	-0.24	0.56	1	-0.24	-0.14	0.15	-0.06	-0.01	-0.34	0.05
RND	0.23	0.23	0.29	-0.4	-0.41	1	0.34	0.06	-0.37	0.26	0.39	-0.04
σ_t^{CF}	0.37	0.39	0.46	-0.49	-0.29	0.39	1	0.01	-0.37	0.15	0.37	-0.02
Opacity	0.03	0.02	0.04	0.08	0.2	-0.11	-0.03	1	-0.23	-0.13	0.15	0.09
ROA	-0.21	-0.21	-0.26	0.03	-0.17	-0.1	-0.18	-0.28	1	0.09	-0.26	-0.11
BTM	-0.05	-0.08	-0.09	-0.21	-0.06	0.25	0.15	-0.19	0.34	1	0.2	-0.29
$VScore$	0.29	0.29	0.34	-0.53	-0.34	0.35	0.49	0.12	-0.2	0.18	1	0.11
Spread	-0.02	-0.02	0.02	-0.29	0.03	-0.09	0	0.1	-0.2	-0.31	0.12	1

Table 7.5: Regressions of realized volatility on fundamental variables

This table provides the coefficient estimates from equation 3.3 below using each of the fundamental indicator variables discussed in section 3. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. Each t-statistic tests the hypothesis that the given coefficient is different from zero. *, **, and *** indicate significance at 10, 5, and 1 percent, respectively.

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_2 V\_FSV^{j}_{it} + \gamma Year_t + \epsilon_{it}$$

	Prediction	I	II	III	IV	V	VI	VII	VIII	IX
V_TA	-	-0.256***								-0.194***
		(-10.295)								(-11.507)
V_LVG	+/-		-0.076***							0.007
			(-4.098)							(0.620)
V_RND	+			0.231***						0.169***
				(5.826)						(5.126)
V_SGI	+				0.041***					0.046***
					(4.264)					(8.351)
σ_{CF}	+					0.235***				0.137***
						(14.037)				(13.158)
V_Opacity	+						0.111***			0.046***
							(7.124)			(4.959)
V_ROA	-							-0.155***		-0.118***
								(-6.340)		(-9.885)
V_BTM	+								0.053***	0.056***
									(2.892)	(4.002)
Obs		78034	78034	78034	78034	78034	78034	78034	78034	78034
Adj. R-square		0.309	0.259	0.256	0.300	0.259	0.275	0.257	0.260	0.370
Year fixed effects		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7.6: Regressions tests on fundamental variables and Black-Scholes implied volatility

This table provides the coefficient estimates from equation 3.4 below using each of the fundamental indicator variables discussed in section 3. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. *, **, and *** indicate significance at 10, 5, and 1 percent, 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 one.

	I	II	III	IV	V	VI	VII	VIII	IX
σ^{IV}	0.652†	0.659†	0.655†	0.659†	0.653†	0.659†	0.655†	0.661†	0.661†
	(-18.05)	(-17.84)	(-17.33)	(-17.81)	(-18.18)	(-17.91)	(-18.00)	(-17.81)	(-17.81)
σ^{RV}_{t-1}	0.247***	0.247***	0.248***	0.247***	0.247***	0.247***	0.248***	0.248***	0.247***
	(10.250)	(10.366)	(10.426)	(10.367)	(10.312)	(10.372)	(10.440)	(10.411)	(10.183)
Spread	-0.026***	-0.019**	-0.018**	-0.018**	-0.020***	-0.019**	-0.021***	-0.016**	-0.025***
	(-3.417)	(-2.580)	(-2.532)	(-2.495)	(-2.690)	(-2.528)	(-2.938)	(-2.421)	(-4.024)
V_TA	-0.022***								-0.016***
	(-3.163)								(-2.706)
V_LVG		-0.009**							-0.007***
		(-2.569)							(-3.014)
V_RND			0.013						0.008
			(1.483)						(0.900)
V_SGI				0.010***					0.008***
				(2.913)					(2.857)
σ_{CF}					0.018***				0.011**
					(3.068)				(2.262)
V_Opacity						0.010***			0.005
						(2.904)			(1.466)
V_ROA							-0.015***		-0.020***
							(-3.195)		(-4.703)
V_BTM								-0.011**	-0.011**
								(-2.075)	(-2.115)
Obs	78034	78034	78034	78034	78034	78034	78034	78034	78034
Adj. R-square	0.748	0.728	0.738	0.734	0.721	0.737	0.733	0.736	0.736
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7.7: Incremental informativeness of $VScore$

This table provides the coefficient estimates from equations 3.1, 3.3, and 3.4 using the V-score as a summary measure of financial statement information. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. *, **, and *** indicate significance at 10, 5, and 1 percent, 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 one.

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_1 \sigma^{IV}_{it} + \beta_2 VScore_{it} + \beta_3 Spread_{it} + \beta_4 \sigma^{RV}_{t-1} + \gamma Year_t + \epsilon_{it}$$

	I	II	III	IV
σ^{IV}	0.896†	0.655†	0.659†	0.644†
	(-5.80)	(-18.30)	(-17.78)	(-18.17)
σ^{RV}_{t-1}		0.251***	0.248***	0.247***
		(10.36)	(10.38)	(10.17)
$Spread$			-0.019**	-0.023***
			(-2.46)	(-3.08)
$VScore$				0.011***
				(3.95)
Obs	78034	78034	78034	78034
Adj. R-square	0.705	0.717	0.718	0.748
Year fixed-effects	Yes	Yes	Yes	Yes

Table 7.8: Regression tests using fundamental variables and model free implied volatility

This table provides the coefficient estimates from equation using each of the fundamental indicator variables discussed in section 3 and model-free implied volatility. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. *, **, and *** indicate significance at 10, 5, and 1 percent, 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 one.

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_1\sigma^{MFIV}_{it} + \beta_2 V\_FSV^j_{it} + \beta_3 Spread_{it} + \beta_4\sigma^{RV}_{t-1} + \gamma Year_t + \epsilon_{it}$$

	I	II	III	IV	V	VI	VII	VIII	IX	X
σ^{MFIV}	0.651†	0.635†	0.649†	0.646†	0.650†	0.644†	0.650†	0.651†	0.654†	0.631†
	(-10.28)	(-10.57)	(-10.21)	(-10.25)	(-10.27)	(-10.43)	(-10.27)	(-10.15)	(-10.23)	(-10.24)
σ^{RV}_{t-1}	0.273***	0.268***	0.272***	0.273***	0.273***	0.272***	0.273***	0.273***	0.272***	0.267***
	(8.204)	(7.853)	(8.152)	(8.216)	(8.177)	(8.068)	(8.192)	(8.198)	(8.218)	(7.836)
Spread	-0.016***	-0.025**	-0.015**	-0.017**	-0.010***	-0.029**	-0.012***	-0.016**	-0.025***	
	(-3.417)	(-2.370)	(-2.982)	(-2.385)	(-2.730)	(-2.258)	(-2.598)	(-2.381)	(-3.03)	
V_TA		-0.045***								-0.036***
		(-3.842)								(-3.069)
V_LVG			-0.014*							-0.003
			(-1.773)							(-0.315)
V_RND				0.020*						0.011**
				(1.767)						(1.970)
V_SGI					0.014*					0.010
					(1.692)					(1.137)
σ_{CF}						0.025***				0.015*
						(2.950)				(1.854)
V_Opacity							0.007			-0.004
							(0.795)			(-0.464)
V_ROA								-0.001		-0.010*
								(-0.088)		(-1.855)
V_BTM									-0.035***	-0.028***
									(-3.984)	(-3.421)
Obs	7275	7275	7275	7275	7275	7275	7275	7275	7275	7275
Adj. R-square	.601	0.648	0.637	0.658	0.644	0.647	0.653	0.642	0.621	0.660
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7.9: Straddle returns from a *VScore* strategy

This table provides 60-day returns from a long straddle portfolio on subsets of the sample divided by VScore. Panel A presents returns calculated assuming all trades occur at the midpoint of the closing bid and ask prices. Panel B presents returns calculated assuming investors buy options at the ask and sell options at the bid. The long straddle portfolio is formed by taking a long position in both an 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. *, **, and *** indicate significance at 10, 5, and 1 percent, respectively, of the hypothesis that the given mean is different from zero.

(a) *VScore* univariate sort

V0	V1	V2	V3	V4	V5	V6	V7	V8	V8 - V0
-0.079**	0.013	-0.005	-0.022*	0.008	-0.001	-0.026	0.077	0.050	0.129**

(b) *VScore* univariate sort with 100% effective spread

V0	V1	V2	V3	V4	V5	V6	V7	V8	V8 - V0
-0.03471**	-0.0303**	-0.0347*	-0.0291	0.0263	-0.0018**	-0.0016	0.0014	0.0006	0.0353**

Table 7.10: Straddle returns from a $VScore$ and $DiffVol$ strategy

This table provides 60-day returns from a long straddle portfolio on subsets of the sample divided by VScore and DiffVol. The long straddle portfolio is formed by taking a long position in both an 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. *, **, and *** indicate significance at 10, 5, and 1 percent, respectively, of the hypothesis that the given mean is different from zero.

(a) DiffVol univariate sort

D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D10 - D1
-0.131***	-0.042**	-0.009	0.011	0.020	0.047	0.034	-0.006	0.038	0.033	0.178* **

(b) $VScore$ and DiffVol double sort

	Low DiffVol								High DiffVol		
	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D10 - D1
$VScore = 0$	-0.200***	-0.050	-0.047	-0.128*	0.060	-0.159**	0.003	-0.055	0.052	0.054	0.254***
V1	-0.199***	-0.082	-0.082	-0.002	-0.086	0.204	0.033	0.002	-0.098	-0.063	0.137
V2	-0.047	-0.082	0.124	-0.111*	0.177	-0.124*	-0.065	-0.079	-0.084*	0.054	0.100
V3	-0.064	-0.157***	0.119	0.036	0.189	0.054	0.026	0.074	0.126	0.104	0.168
V4	-0.117**	-0.047	-0.017	-0.189***	0.127	0.118	-0.025	-0.093*	0.076	-0.043	0.074
V5	-0.151***	0.062	0.033	-0.071	-0.180***	0.070	0.117	0.115	0.059	0.037	0.189**
V6	-0.122**	-0.099*	-0.192***	0.164	-0.046	0.116	0.204	0.040	0.038	0.063	0.186**
V7	-0.158***	0.099	-0.018	0.002	0.048	0.082	-0.074	-0.041	0.199	-0.088*	0.071
$VScore = 8$	-0.052*	-0.107**	-0.018	0.152	0.054	0.001	0.033	-0.021	-0.006	0.056	0.107*
V8-V0	0.148**	-0.057	0.029	0.281**	-0.006	0.160**	0.030**	0.034	-0.058	0.001	

Figure 7.3: Hedge straddle returns to a *VScore* strategy from 1996 - 2010

This table provides 60-day returns from a long straddle portfolio averaged for each year in the sample. The long straddle portfolio is formed by taking a long position in both an 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.

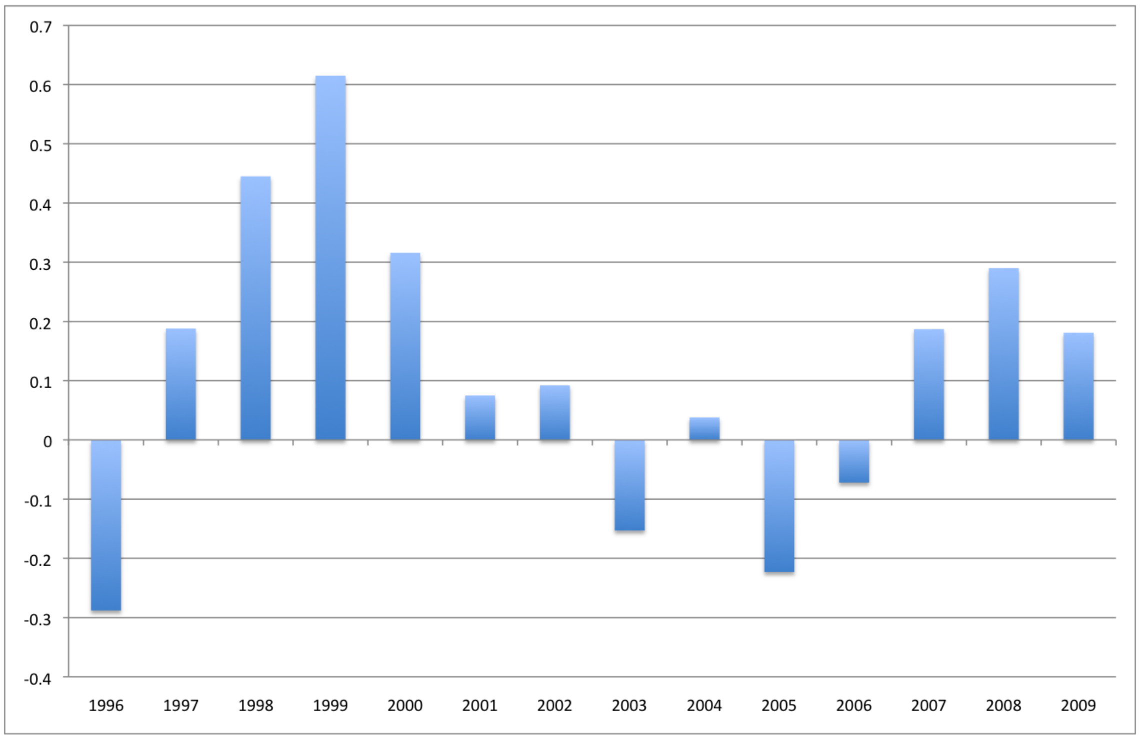


Table 7.11: Straddle returns regressions

This table provides coefficient estimates from equation 3.6. The dependent variable in each model is firm-level 60-day straddle portfolio returns. In parentheses below each coefficient are Fama MacBeth t-statistics from Newey-West corrected standard errors. *, **, and *** indicate significance at 10, 5, and 1 percent, respectively, of the hypothesis that the given coefficient is different from zero.

$$r^s_{i,t} = \alpha + \beta_1 \text{Rank}\sigma^{IV}_{i,t,\tau} + \beta_2 \text{Rank}\sigma^{RV}_{i,t-1,\tau} + \beta_3 VScore_{it} + \epsilon_{it}$$

	I	II	III
Intercept	0.079	0.082	0.039
	(1.51)	(1.47)	(0.73)
Rank σ^{IV}_t	-0.014***	-0.020***	-0.024***
	(-2.96)	(-4.32)	(-4.96)
Rank σ^{RV}_{t-1}		0.006	0.005
		(1.41)	(1.21)
VScore			0.021***
			(2.85)
No. of quarters	56	56	56
Obs per quarter	306	306	306

Table 7.12: Regressions of realized volatility on continuous fundamental variables

This table provides the coefficient estimates from equation 3.3 below using each of the fundamental indicator variables discussed in section 3. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. Each t-statistic tests the hypothesis that the given coefficient is different from zero. *, **, and *** indicate significance at 10, 5, and 1 percent, respectively.

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_2 V\_FSV^j_{it} + \gamma Year_t + \epsilon_{it}$$

	Dependent variable: realized volatility								
	I	II	III	IV	V	VI	VII	VIII	IX
TA	-0.107***								-0.087***
	(-10.444)								(-10.112)
LVG		-0.037***							-0.019**
		(-3.545)							(-2.111)
RND			0.076***						0.037***
			(5.371)						(4.656)
SGI				0.018***					0.018***
				(5.142)					(9.138)
σ_{CF}					0.107***				0.065***
					(18.373)				(13.627)
$Opacity$						0.071***			0.029***
						(9.607)			(5.550)
ROA							-0.085***		-0.062***
							(-9.001)		(-9.708)
BTM								0.031***	0.034***
								(2.982)	(4.321)
Obs	78034	78034	78034	78034	78034	78034	78034	78034	78034
R-square	0.29	0.26	0.27	0.26	0.29	0.27	0.28	0.26	0.34
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7.13: Regressions of realized volatility on continuous fundamental variables and implied volatility

This table provides the coefficient estimates from equation 3.4 below using each of the fundamental indicator variables discussed in section 3. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. *, **, and *** indicate significance at 10, 5, and 1 percent, 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 one.

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_1\sigma^{IV}_{it} + \beta_2 V\_FSV^j_{it} + \beta_3 Spread_{it} + \beta_4\sigma^{RV}_{t-1} + \gamma Year_t + \epsilon_{it}$$

	Dependent variable: realized volatility								
	I	II	III	IV	V	VI	VII	VIII	IX
σ^{IV}	0.655†	0.659†	0.657†	0.659†	0.655†	0.658†	0.654†	0.660†	0.646†
	(-17.82)	(-17.78)	(-18.07)	(-17.76)	(-17.94)	(-17.75)	(-17.86)	(-17.84)	(-18.10)
σ^{RV}_{t-1}	0.247***	0.247***	0.248***	0.248***	0.247***	0.248***	0.248***	0.248***	0.248***
	(10.323)	(10.367)	(10.419)	(10.395)	(10.341)	(10.381)	(10.478)	(10.418)	(10.317)
Spread	-0.023***	-0.019**	-0.018**	-0.019**	-0.019**	-0.019**	-0.021***	-0.017**	-0.024***
	(-3.050)	(-2.564)	(-2.507)	(-2.520)	(-2.654)	(-2.554)	(-2.913)	(-2.466)	(-3.507)
TA	-0.007**								-0.006**
	(-2.657)								(-2.305)
LVG		-0.004*							-0.005**
		(-1.972)							(-2.649)
RND			0.006**						0.003
			(2.603)						(1.278)
SGI				0.004***					0.003***
				(3.345)					(2.679)
σ_{CF}					0.008***				0.004**
					(3.344)				(2.518)
$Opacity$						0.004**			0.001
						(2.531)			(0.391)
ROA							-0.008***		-0.008***
							(-3.094)		(-3.649)
BTM								-0.004*	-0.004
								(-1.684)	(-1.569)
Obs	78034	78034	78034	78034	78034	78034	78034	78034	78034
R-square	0.724	0.718	0.734	0.717	0.723	0.712	0.724	0.736	0.754
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7.14: Regression tests of realized volatility on continuous fundamental variables and model free implied volatility

This table provides the coefficient estimates from equation 3.4 using each of the fundamental indicator variables discussed in section 3 and model-free implied volatility. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. *, **, and *** indicate significance at 10, 5, and 1 percent, 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 one.

$$\sigma^{RV}_{i,t+\tau} = \alpha + \beta_1\sigma^{MFIV}_{it} + \beta_2 V\_FSV^j_{it} + \beta_3 Spread_{it} + \beta_4\sigma^{RV}_{t-1} + \gamma Year_t + \epsilon_{it}$$

	Dependent variable: realized volatility								
	I	II	III	IV	V	VI	VII	VIII	IX
σ^{MFIV}	0.642†	0.649†	0.648†	0.650†	0.643†	0.652†	0.649†	0.654†	0.640†
	(-10.55)	(-10.27)	(-10.44)	(-10.28)	(-10.48)	(-10.28)	(-10.28)	(-10.20)	(-10.47)
σ^{RV}_{t-1}	0.271***	0.273***	0.273***	0.273***	0.273***	0.273***	0.274***	0.272***	0.269***
	(8.038)	(8.164)	(8.210)	(8.160)	(8.070)	(8.208)	(8.201)	(8.201)	(7.983)
Spread	-0.070***	-0.062***	-0.061***	-0.060***	-0.061***	-0.062***	-0.062***	-0.053***	-0.060***
	(-5.162)	(-4.841)	(-4.764)	(-4.718)	(-4.843)	(-4.859)	(-4.931)	(-4.604)	(-5.052)
TA	-0.015***								-0.009**
	(-3.324)								(-2.311)
LVG		-0.007*							-0.009**
		(-1.907)							(-2.071)
RND			0.009*						-0.001
			(2.002)						(-0.204)
SGI				0.009*					0.006
				(1.925)					(1.313)
σ_{CF}					0.015***				0.012**
					(3.099)				(2.280)
$Opacity$						-0.002			-0.012***
						(-0.588)			(-2.745)
ROA							-0.004		-0.009***
							(-1.196)		(-24.196)
BTM								-0.020***	-0.022***
								(-4.278)	(-4.636)
Obs	7275	7275	7275	7275	7275	7275	7275	7275	7275
R-square	0.684	0.678	0.694	0.677	0.683	0.672	0.684	0.696	0.714
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 7.15: Regressions of realized volatility on fundamental variables and volatility spread controls

This table provides the coefficient estimates from equation 3.4 using each of the fundamental indicator variables discussed in section 3 and a control for volatility spread defined in section 5.2. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. *, **, and *** indicate significance at 10, 5, and 1 percent, 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 one.

	Dependent variable: Realized volatility								
	I	II	III	IV	V	VI	VII	VIII	IX
σ^{IV}	0.679†	0.686†	0.682†	0.687†	0.681†	0.686†	0.684†	0.689†	0.670†
	(-17.82)	(-17.88)	(-17.49)	(-17.84)	(-18.22)	(-17.74)	(-18.05)	(-17.79)	(-17.53)
σ^{RV}_{t-1}	0.219***	0.219***	0.220***	0.219***	0.219***	0.219***	0.220***	0.219***	0.220***
	(9.848)	(9.923)	(9.996)	(9.930)	(9.905)	(9.948)	(10.033)	(9.979)	(9.850)
Spread	-0.026***	-0.019**	-0.018**	-0.018**	-0.020***	-0.019**	-0.021***	-0.016**	-0.023***
	(-3.396)	(-2.650)	(-2.614)	(-2.562)	(-2.767)	(-2.597)	(-2.882)	(-2.451)	(-3.831)
VolSpread	-0.004***	-0.005***	-0.005***	-0.005***	-0.004***	-0.005***	-0.005***	-0.005***	-0.004***
	(-9.379)	(-9.738)	(-9.622)	(-9.704)	(-9.570)	(-9.592)	(-9.655)	(-9.588)	(-9.217)
V_TA	-0.020***								-0.014**
	(-2.770)								(-2.134)
V_LVG		-0.009***							-0.007***
		(-2.668)							(-2.821)
V_RND			0.015						0.009
			(1.595)						(1.039)
V_SGI				0.009***					0.010***
				(2.909)					(3.588)
σ_{CF}					0.018***				0.012**
					(3.038)				(2.351)
V_Opacity						0.009**			0.003
						(2.607)			(0.991)
V_ROA							-0.012**		-0.018***
							(-2.652)		(-4.333)
V_BTM								-0.013**	-0.013**
								(-2.264)	(-2.254)
Obs	72834	72834	72834	72834	72834	72834	72834	72834	72834
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Figure 7.4: Market and firm-level equity volatility by quarter, 1996 - 2010

This figure plots the level of market and average firm-level equity volatilities in my sample by quarter. The red dotted (blue dashed) line plots the average volatility by quarter for firms whose option-implied volatility is less than (greater than or equal to) realized volatility. The green solid line plots the volatility of the S&P 500 index by quarter.

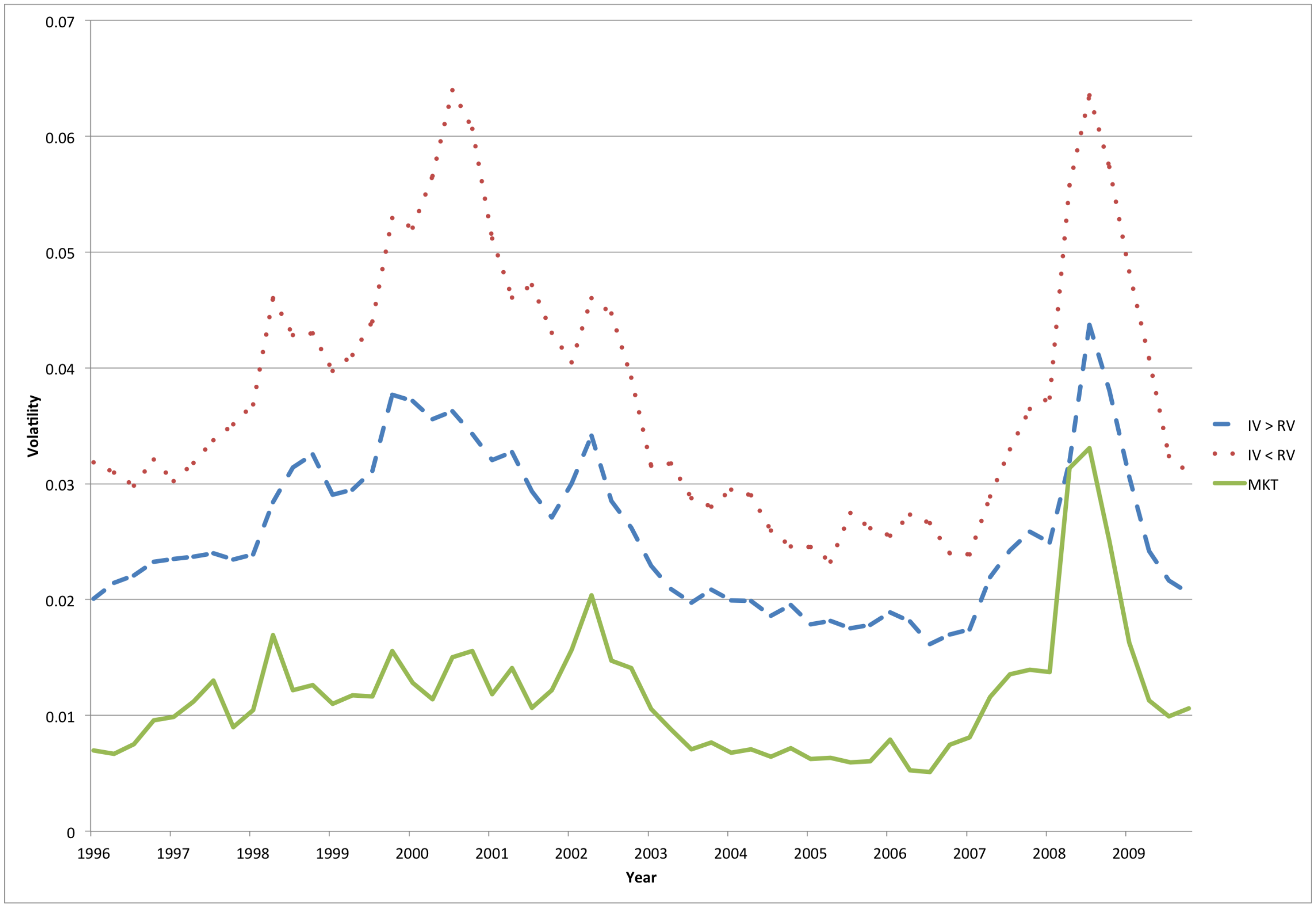


Table 7.16: Regressions of realized volatility on fundamentals and implied volatility: IV < RV subsample

This table provides the coefficient estimates from equation 3.4 using each of the fundamental indicator variables discussed in section 3 on the subsample of firm-quarters for which implied volatility is smaller that subsequent realized volatility. In parentheses below each coefficient are t-statistics from two-way industry and quarter clustered standard errors. *, **, and *** indicate significance at 10, 5, and 1 percent, respectively, of the hypothesis that the given coefficient is different from zero. †indicates significance at the 5 percent level of the hypothesis that the given coefficient is different from one.

	Dependent variable: realized volatility								
	I	II	III	IV	V	VI	VII	VIII	IX
σ^{IV}	0.942†	0.948†	0.949†	0.947†	0.944†	0.947†	0.944†	0.948†	
	(-2.54)	(-2.37)	(-2.35)	(-2.40)	(-2.44)	(-2.40)	(-2.46)	(-2.37)	
σ^{RV}_{t-1}	0.014	0.014	0.014	0.014	0.014	0.014	0.015	0.014	0.014
	(1.047)	(1.072)	(1.030)	(1.053)	(1.053)	(1.061)	(1.110)	(1.062)	(1.049)
Spread	0.017***	0.022***	0.022***	0.023***	0.022***	0.022***	0.020***	0.023***	0.014**
	(3.099)	(4.310)	(4.112)	(4.452)	(4.161)	(4.290)	(3.910)	(4.552)	(2.602)
V_TA	-0.015**								-0.018***
	(-2.481)								(-3.051)
V_LVG		0.003							0.006**
		(0.857)							(2.297)
V_RND			-0.034***						-0.005
			(-4.86)						(-0.790)
V_SGI				0.008***					0.007**
				(2.675)					(2.558)
σ_{CF}					0.012***				0.009**
					(2.807)				(2.113)
V_Opacity						0.002			-0.000
						(0.519)			(-0.104)
V_ROA							-0.012**		-0.012***
							(-2.449)		(-2.824)
V_BTM								-0.011**	-0.001
								(-2.380)	(-0.074)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Chapter 8

Appendices

8.1 Estimating model free implied volatility

As discussed in section 2.2, I calculate 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:

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

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 re-written 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 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

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

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

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:

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

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.

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

The above approximation assumes

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

where q is the annual dividend rate (assumed zero) and S_0 is the price of the underlying asset at $t = 0$. ΔK_i is defined as follows:

$$\Delta K_i = \begin{cases} K_2 - K_1 & if \quad i = 1 \\ \frac{K_{i+1} - K_{i-1}}{2} & if \quad 1 < 1 < M \\ K_M - K_{M-1} & if \quad i = M \end{cases}$$

8.2 Variable definitions

- $\sigma^{RV}_{i,t,\tau}$ is the logarithm of the observed standard deviation of firm i's equity returns over the period starting five days after quarter t's earnings announcement date and ending τ days later

- $\sigma^{IV}_{i,t,\tau}$ is the logarithm of 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. Implied volatility is either estimated using the Black and Scholes [1973] option pricing formula or the model-free estimation process described in appendix A.

- $Size_t$ is the level of total assets at the end of quarter $t - 1$.

- Lvg_t is the ratio of total liabilities to total assets reported at the end of quarter $t - 1$

- σ^{CF}_t, the firm's cash flow volatility as of quarter t, is the standard deviation of operating cash flows scaled by total assets over the 10 quarters prior to (and *not* including) the quarter t.

- ROA_t is the average of the ratio of earnings before interest divided by total assets for quarters $t - 1$ through $t - 4$.

- RND_t for quarter t is the ratio of research and development expense (assumed to be zero if not reported) to total assets, both measured at the end of quarter $t-1$.

- SGI_t for quarter t is the ratio of sales revenue in quarter $t-1$ to sales revenue in quarter $t-2$.

- BTM is the ratio of book value of equity to market value of equity. I estimate book value of equity as the difference between total assets and total liabilities at the end of quarter $t-1$. I measure market value of equity for firm i in quarter t using the closing stock price on the day before the quarter t earnings announcement date.

- $Opacity = |DAcc_{i,t-1}| + |DAcc_{i,t-2}| + |DAcc_{i,t-3}|$ where $DAcc_{it}$ denotes discretionary accruals, defined as the residual from estimating the following regression cross-sectionally for each quarter

$$\frac{TAcc_{it}}{AT_{it}} = \alpha + \beta_1 \frac{1}{AT_{it}} + \beta_2 \frac{\Delta Sales_{it}}{AT_{it}} + \beta_3 \frac{PPE_{it}}{AT_{it}} + \epsilon_{it}$$

  In the above equation, $TAcc_{it}$ is the difference between net income and operating cash flows for firm i at the end of quarter t, $\Delta Sales_{it}$ is the change in sales revenue from quarter $t-1$ to quarter t, AT_{it} is firm i's total assets at the beginning of quarter t, and PPE_{it} is firm i's net property plant and equipment at the beginning of quarter t.

- V_Size_{it} (V_Lvg_{it}, $V\_\sigma^{CF}_{it}$, V_ROA_{it}, V_RND_{it}, V_SGI, V_BTM_{it}, $V\_Opacity_{it}$) is an indicator variable equalling 1 if the $Size$ (Lvg, σ^{CF}_{it}, ROA, RND, SGI, BTM, $Opacity$) of firm i in quarter t exceeds that of its industry median.

- $VScore = V\_RND_{it} + V\_SGI_{it} + V\_\sigma^{CF}_{it} + V\_Opacity_{it} + V\_BTM_{it} + (1 -$

$V\_Size_{it}) + (1 - V\_ROA_{it}) + (1 - V\_Lvg_{it})$

- *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.

Bibliography

E. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 23(4):589–609, 1968.

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

P. P. Barros. Endogenous mergers and size asymmetry of merger participants. *Economics Letters*, 60(1):113 – 119, 1998.

M. Barth and E. So. Market volatility risk and risk premiums at earnings announcements. Working paper, 2013.

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

W. Beaver. Financial ratios as predictors of failure. *Journal of Accounting Research*, 4:71–111, 1966.

W. Beaver, M. McNichols, and J. W. Rhie. Have financial statements become less informative? evidence from the ability of financial ratios to predict bankruptcy. *Review of Accounting Studies*, 10:93–122, 2005.

M. Beneish, C. Lee, and R. Tarpley. Contextual Fundamental Analysis Through the Prediction of Extreme Returns. *Review of Accounting Studies*, 6:165–189, 2001.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, pages 637–654, 1973.

D. Breeden and R. Litzenberger. Prices of state-contingent claims implicit in option prices. *Journal of business*, pages 621–651, 1978.

M. Britten-Jones and A. Neuberger. Option prices, implied price processes, and stochastic volatility. *The Journal of Finance*, 55(2):839–866, 2000. ISSN 1540-6261.

J. Campbell, M. Lettau, B. Malkiel, and Y. Xu. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance*, 56(1):1–43, 2001. ISSN 1540-6261.

L. Canina and S. Figlewski. The informational content of implied volatility. *Review of Financial studies*, 6(3):659, 1993. ISSN 0893-9454.

L. Chan, J. Lakonishok, and T. Sougiannis. The Stock Market Valuation of Research and Development Expenditures. *Journal of Finance*, 56:2431–2456, 2001.

S. Chava and R. Jarrow. Bankruptcy prediction with industry effects, market versus accounting variables, and reduced form credit risk models. *Review of Finance*, 8 (4):537–569, 2004.

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

B. Christensen and N. Prabhala. The relation between implied and realized volatility. *Journal of Financial Economics*, 50(2):125–150, 1998. ISSN 0304-405X.

P. Christoffersen, R. Goyenko, K. Jacobs, and M. Karoui. Illiquidity premia in the equity options market. Working paper, 2011.

M. Cremers and D. Weinbaum. Deviations from put-call parity and stock return predictability. *Journal of Financial and Quantitative Analysis*, 45(02):335–367, 2010.

T. Day and C. Lewis. Stock Market Volatility and the Informational Content of Stock Index Options. *Journal of Econometrics*, 52(1), 1992.

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

P. Dechow, W. Ge, and C. Schrand. Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, 50:344–401, 2010.

P. Dennis and S. Mayhew. Risk-neutral skewness: Evidence from stock options. *Journal of Financial and Quantitative Analysis*, 37(03):471–493, 2002.

E. Fama and K. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.

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

E. Fama and J. MacBeth. Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy*, pages 607–636, 1973.

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

T. Goodman, M. Neamtiu, and F. Zhang. Fundamental analysis and option returns. Working paper, 2013.

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

A. Goyal and A. Saretto. Cross-Section of Option Returns and Volatilit. *Journal of Financial Economics*, 94(2):310–326, 2009.

S. Hillegeist, D. Cram, E. Keating, and K. Lundstedt. Assessing the probability of bankruptcy. *Review of Accounting Studies*, 9(1):5–34, 2004.

J. Hull. *Options, futures and other derivatives*. Pearson Prentice Hall, 2009.

A. Hutton, A. Marcus, and H. Tehranian. Opaque financial reports, r-square, and crash risk. *Journal of Financial Economics*, 94:67–86, 2009.

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

G. Jiang and Y. Tian. Extracting model-free volatility from option prices: An examination of the VIX index. *Journal of Derivatives*, 14(3):35, 2007. ISSN 1074-1240.

W. Jin, J. Livnat, and Y. Zhang. Options prices leading equity prices: Superior information discovery or superior information processing? *Journal of Accounting Research*, 2012.

R. S. Kaplan and G. Urwitz. Statistical models of bond ratings: A methodological inquity. *Journal of Business*, 52(2):231–261, 1979.

J. Lakonishok, R. Vishny, and A. Shleifer. Contrarian investment, extrapolation, and risk. *Journal of Finance*, 49(5):1541–1578, 1993.

C. Lamoureux and W. Lastrapes. Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. *Review of Financial Studies*, 6(2): 293, 1993. ISSN 0893-9454.

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

C. M. C. Lee. 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, 2001.

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

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

J. Ohlson. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18:109–131, 1980.

J. Pan. The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of financial economics*, 63(1):3–50, 2002.

L. Pastor and P. Veronesi. Stock valuation and learning about profitability. *Journal of Finance*, 58:1749–1789, 2003.

S. H. Penman. The articulation of price-earnings ratios and market-to-book ratios and the evaluation of growth. *Journal of Accounting Research*, 34(2):pp. 235–259, 1996.

J. D. Piotroski. Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, 38:1–41, 2000. ISSN 0021-8456.

V. Pool, H. Stoll, and R. Whaley. Failure to exercise call options: An anomaly and a trading game. *Journal of Financial Markets*, 11(1):1–35, 2008.

S. Poon and C. Granger. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2):478–539, 2003.

S. Rajgopal and M. Venkatachalam. Financial reporting quality and idiosyncratic return volatility. *Journal of Accounting and Economics*, 51:1–20, 2011.

R. Roll, E. Schwartz, and A. Subrahmanyam. O/s: The relative trading activity in options and stock. *Journal of Financial Economics*, 96(1):1–17, 2010.

S. Ross. The determination of financial structure: the incentive-signalling approach. *The Bell Journal of Economics*, 8(1):23–40, 1977. ISSN 0361-915X.

T. Shumway. Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business*, 74:101–124, 2001.