Attentive Options Traders: Textual Changes to 10-Ks and Option Volatility Smirk

Hua Cheng

Sun Yat-sen University International School of Business and Finance;
University of Texas at Austin
harrychenghua@utexas.edu

Steve Liu

University of Rhode Island College of Business
steve.liu@uri.edu

Zheng Qiao

Xi'an Jiaotong University School of Management
qiaozheng@xjtu.edu.cn (corresponding author)

Z. Jay Wang*

University of Oregon Lundquist College of Business
zhiw@uoregon.edu

Abstract

In contrast to the lazy prices phenomenon (Cohen, Malloy, and Nguyen, 2020) in the stock market, more 10-K textual changes lead to larger increases in volatility smirks—consistent with options traders buying more out-of-the-money put options based on negative information disclosed in textual changes. Moreover, the lazy-prices effect is mainly driven by stocks with tradable options, suggesting that limits to arbitrage lead to a delayed response of stock prices. Finally, the return predictability of textual changes is stronger for stocks with larger option volatility smirk changes. Sophisticated options traders therefore demonstrate superior skills at extracting relevant information from public filings.

Key words: annual report similarity; 10-K; option volatility smirk; informed trading; information processing; textual analysis
JEL codes: G12, G13, G14

*We are especially grateful to the editor, Thierry Foucault, and an anonymous referee for their insightful and constructive suggestions. We also thank seminar participants at UT-Austin, Nankai University, and Sun Yat-sen University, and conference participants at the FMA Europe, FMA Middle East, American Accounting Association, Eastern Finance Association, China International Conference in Finance, and Asian Finance Association conferences. All remaining errors are our own.

I Introduction

A growing body of research shows that options trading contains valuable information, with price discovery often occurring in the options market before appearing in the equity market (Black, 1975, Amin and Lee, 1997, Cao, Chen, and Griffin, 2005, Augustin, Brenner, and Subrahmanyam, 2019). In this paper, we examine options trading surrounding the release of 10-K filings—an important yet complex source of publicly available information. By focusing on a single type of information rather than various news releases, we can better pinpoint the nature of options traders' information advantages. Cohen et al. (2020) document a "lazy prices" phenomenon in the stock market: Despite changes to the language and construction of 10-Ks containing (mostly negative) information about future profitability and financial health, stock market investors appear to be inattentive to such changes and only slowly incorporate the new information into stock prices. We hypothesize that, compared to stock market investors, options traders are more sophisticated at parsing 10-K textual changes and incorporate it into options pricing more quickly after release. Our findings show that options traders do respond to these changes in a timely manner, especially among firms with higher information asymmetry and greater arbitrage restrictions. This suggests that advanced information-processing skills represent an important source of competitive advantage for options traders.

We use changes in option volatility smirks to proxy for options trading around 10-K releases. Bates (1991) argues that the set of index call and put option prices across all exercise prices reflects market participants' aggregate subjective distribution of future price realizations. Consequently, out-of-the-money (OTM) puts become unusually expensive compared to at-the-money (ATM) calls if traders expect large price declines, resulting in steep option volatility smirks. In an option pricing model incorporating both jump risk and volatility risk, Pan (2002) shows that investors' aversion toward negative price movements drives option volatility smirks as they tend to buy OTM puts to express their concerns about downside risk. Focusing on individual stock options, Xing, Zhang, and Zhao (2010) show that the shape of volatility smirk has significant cross-sectional predictive power on future equity returns. Stocks exhibiting the steepest smirks in their traded options underperform stocks with the least pronounced ones, and such predictability on future stock returns is persistent for at least 6 months. We therefore conjecture that attentive traders who derive insights from 10-K textual changes prefer to trade OTM put options, and that the newly revealed negative information is only slowly incorporated into equity prices due to limits to arbitrage.

To investigate options traders' response to the negative information contained in 10-K textual changes, we keep firm-year observations with volatility smirk information available within three months both before and after the 10-K release date.[1] Following Brown and Tucker (2011), Peterson, Schmardebeck, and Wilks (2015), and Cohen et al. (2020), we quantify textual changes in 10-Ks between two consecutive years using the cosine similarity. A lower cosine similarity score indicates more extensive textual

[1]In untabulated analysis, we document a similar but weaker effect of textual changes to 10-Qs filings, consistent with the fact that 10-Ks are much more important than 10-Qs. Therefore, we focus on 10-Ks throughout this paper.

changes in 10-Ks. Following Xing et al. (2010), we define volatility smirk as the difference in implied volatilities between OTM put options and ATM call options. To capture incremental options trading in response to new 10-K information, we focus on the changes in volatility smirk immediately after the 10-K release relative to at least 8 weeks before the release date.

We first sort firms into quintiles based on their 10-K textual similarity scores and compare changes in option volatility smirks. Firms with the most textual changes show a significant increase in volatility smirks after their 10-K release compared to firms with minimal changes. We then run Fama-MacBeth regressions (Fama and MacBeth, 1973) controlling for additional textual features—including tone optimism and file size—along with firm-level financial variables. We again find a highly significant and positive effect of textual changes on option volatility smirk changes. The economic magnitude is similar to the univariate analysis results. Our analysis of weekly options trading before and after 10-K releases shows that abnormal volatility smirk responses occur primarily in the weeks following the release. Our results are robust to different measures of annual report similarity as well as fixed effects regressions. Together, these results provide supporting evidence that sophisticated investors trade options in response to new information in 10-K releases, which in turn drives changes in option volatility smirks.

To better understand how textual changes relate to volatility smirk changes, we analyze the tone of 10-K textual changes. Empirical evidence in Cohen et al. (2020) suggests that textual changes on average contain negative information about future firm profitability. While not all textual changes convey negative news, our proposed mechanism suggests that the negative relationship between cosine measures and volatility smirk changes stems primarily from textual changes containing negative information. To test this, we construct a negative sentiment change measure based on the number of negative words in textual changes. Indeed, we find a much stronger effect in the subsample with larger increases in negative sentiment. We also show that the open interest of OTM put options is significantly higher for the same subsample. This result aligns with the conjecture that higher volatility smirks following textual changes with negative tones are driven by increased demand for OTM put options.

Next, we explore the effects of textual changes for different items in 10-Ks. We examine text similarity in product descriptions of firms' own and competitors' 10-Ks in neighboring years by utilizing two measures of product market fluidity from Hoberg, Phillips, and Prabhala (2014). Our findings using textual changes to product descriptions in firms' own 10-Ks are consistent with those using textual changes to the entire 10-Ks. By contrast, textual changes to product descriptions in competitors' 10-Ks have limited effect on a firm's option volatility smirk changes. Cohen et al. (2020) document that firms' reporting changes are largely concentrated in the Management's Discussion and Analysis (MD&A) section (Item 7), the Legal Proceedings section (Item 3), the Quantitative and Qualitative Disclosures about Market Risk section (Item 7a), and the Risk Factors section (Item 1a). These sections provide greater managerial discretion in reporting, which may account for the observed pattern. Our findings confirm that changes in these sections significantly impact option volatility smirk changes.

The evidence thus far is consistent with options traders demonstrating superior skills in processing negative information from annual report changes and profiting from trades in OTM put options. This information advantage is likely to be more pronounced among firms with greater information asymmetry and higher short-sale costs. Following prior studies, we measure a firm's information environment using abnormal idiosyncratic volatility (Yang, Zhang, and Zhang, 2020) and analyst coverage (Hu, 2014). We find that the sensitivity of volatility smirks to 10-K textual changes is concentrated among informationally asymmetric firms—those with high abnormal idiosyncratic volatility and low analyst oversight. Furthermore, since options are informationally more valuable for trading hard-to-short stocks, we find that volatility smirks respond more strongly to textual changes among firms with higher short-selling costs (Hu, 2014).

Finally, we investigate the return predictability of 10-K textual changes and option volatility smirk changes. While options trading offers several well-documented advantages, we acknowledge that sophisticated investors may also exploit their information advantage through stock trading. This is especially relevant for stocks without tradable options, as it affects how quickly stock prices incorporate new information. We therefore compare optionable and non-optionable stocks to assess the difference in the lazy-prices effect between the two groups. Our results show that the lazy-prices effect is primarily driven by optionable stocks, with no significant evidence among non-optionable stocks. This indicates that informed trading of non-optionable stocks could accelerate price adjustment, thereby diminishing the lazy-prices effect. In contrast, with optionable stocks, informed investors prefer trading options to leverage their information advantage. This leads to a delayed response of stock prices to 10-K textual changes, likely reflecting arbitrage limitations between the two markets.

We also examine whether changes in option volatility smirks have additional predictive power for future returns. Since not all textual changes contain price-relevant information, if options traders truly possess superior skills in extracting trading signals from textual changes, then combining cosine measures with option volatility smirk changes should yield stronger return predictability. To test this hypothesis, we double-sort stocks into quintile portfolios based on the cosine measures and also divide them into two groups based on the median volatility smirk change. We find that stocks with low cosine similarity and high smirk changes show negative alphas on average, underperforming all other groups—particularly those with high cosine similarity and low smirk changes. Our results suggest that the underperformance of stocks with greater 10-K textual changes is mainly driven by stocks exhibiting a steeper increase in volatility smirks.

Our paper contributes to three strands of literature. First, we contribute to the literature on informed options trading and particularly the sources of options traders' information advantage. There is a large literature on informed options trading around major corporate events and macroeconomic announcements (Augustin and Subrahmanyam, 2020). Empirical evidence strongly indicates that these pre-event trading patterns stem from private information. Other studies show that measures derived from option volumes, prices, and volatility strongly predict future stock returns and events in the equity market (Easley, O'Hara, and Srinivas, 1998, Pan, 2002, Chakravarty, Gulen, and Mayhew, 2004, Ofek, Richardson, and Whitelaw, 2004, Cao et al., 2005, Pan and Poteshman, 2006, Xing et al., 2010, Bali and Murray, 2013, Kim and Zhang, 2014). However, research on informed options trading based on public information remains limited. We provide evidence that trading around 10-K releases is primarily driven by options traders' ability to process complex information disclosed in these filings.

Second, our paper contributes to the growing literature on the use of public information by sophisticated investors. Cohen and Frazzini (2008) and Menzly and Ozbas (2010) argue that outside shareholders have limited capacity to fully comprehend the impact of public information across firms and industries, and show that this leads to return predictability across economically linked firms. Engelberg, Reed, and Ringgenberg (2012) find that a substantial portion of short sellers trading advantage comes from their ability to analyze publicly available information using a database of news releases. Alldredge and Cicero (2015) provide evidence that some profitable insider stock selling is motivated by public information on disclosed sales relationships. Dugast and Foucault (2025) develop a theoretical framework and show that improvements in information technology lead to lower information search costs, resulting in more precise performance predictions by quantitative investors. Regarding the use of EDGAR filings by investors, Drake, Roulstone, and Thornock (2015) and Gibbons, Iliev, and Kalodimos (2021) show that Forms 10-K, 10-Q, and 8-K are an important information source for sophisticated investors. Our study builds on this literature by demonstrating how options traders respond to and profit from new information revealed in 10-K textual changes through their superior information processing skills.

Third, our paper sheds new light on the "lazy prices" phenomenon documented in Cohen et al. (2020) and explores the interconnection between options and equity markets (Detemple and Selden, 1991, Duffie, Gârleanu, and Pedersen, 2002, Lamont and Thaler, 2003, Ofek et al., 2004). We show that the lazy-prices effect is mainly driven by optionable stocks and is insignificant for stocks without tradable options. This pattern suggests that informed investors prefer trading options when available, and the limited arbitrage between the two markets leads to a delayed response of stock prices to new information.

The remainder of this paper is organized as follows. Section II develops testable hypotheses. Section III presents the data sample and summary statistics. Section IV describes our research design and presents the main empirical results. Section V explores the heterogeneous effects of 10-K textual changes. Section VI investigates the return predictability of textual and option volatility smirk changes. Section VII provides concluding remarks.

II Hypothesis Development

Our research builds on two strands of literature on information acquisition via EDGAR filings and investor sophistication in options market. Recent research has highlighted the importance of information acquisition via EDGAR filings. Drake et al. (2015) document that a small group of users (likely sophisticated investors) frequently access a substantial amount of information from EDGAR, especially around major corporate announcements and after a period of poor stock performance. Moreover, EDGAR requests for filings related to earnings disclosure are positively associated with the speed at which the market impounds earnings news. Gibbons et al. (2021) focus on sell-side equity analysts and document that they actively acquire information via EDGAR prior to making earnings forecasts. Such information acquisition is related to a significant reduction in forecasting errors and a stronger market reaction following the recommendation. Both studies show that Forms 10-K, 10-Q, and 8-K are among the most highly accessed filings and contain information pertinent to major corporate events and firm profitability in general. Compared to 10-Q and 8-K filings, 10-K reports are generally much longer and more complex, requiring advanced information processing skills to fully understand their value implications. Our paper extends this strand of literature by examining options traders' information acquisition via textual changes in 10-K filings.

There is ample evidence on informed options trading around major corporate events (e.g., M&As, earnings news, and leveraged buyouts) and macroeconomic announcements (e.g., Federal Open Market Committee meetings). Augustin and Subrahmanyam (2020) provide a comprehensive survey of this literature. Another strand of research provides evidence consistent with informed options trading unconditional on major news announcements. These studies (Easley et al., 1998, Chakravarty et al., 2004, Pan and Poteshman, 2006, Xing et al., 2010, Ge, Lin, and Pearson, 2016) find that information-based measures derived from options trading volume and prices have strong predictability on future stock returns. These studies suggest that sophisticated investors prefer using options trading (especially OTM options) to exploit their information advantage, possibly due to several distinctive characteristics of options trading: high embedded leverage, low margin requirements, low implicit borrowing rates, low commissions, and limited downside risk (Black, 1975, Amin and Lee, 1997, Cao et al., 2005, Li, Lin, and Pearson, 2016).

Recently, Cohen et al. (2020) document the "lazy prices" phenomenon consistent with the inattentiveness of unsophisticated stock market investors. Although changes to the language and construction of 10-Ks predict future earnings, major news announcements, and bankruptcy risk, stock prices tend to react slowly to these changes. We hypothesize that options traders possess superior information processing skills and thus the comparative advantage over average stock market investors of extracting relevant information from complex 10-K filings. Building on Cohen et al. (2020), who find that textual changes in 10-K filings typically convey negative news about future firm profitability, we conjecture that this information prompts options traders to purchase more OTM put options, leading to an increase in volatility smirks. Furthermore, if options traders profit from extracting relevant information from these textual changes, the corresponding responses in option volatility smirks should predominantly occur after the 10-K release.

We formulate our first hypothesis regarding the relation between textual changes in 10-K filings and changes in option volatility smirks as follows:

Hypothesis H1: *Negative information in 10-K textual changes leads to an* *increase in option volatility smirks following 10-K releases.*

We next investigate the heterogeneous effects of 10-K textual changes on options trading. Since not all changes provide valuable trading signals, we expect traders to focus more on sections that reveal fundamental changes in the firm and where managers have greater discretion in their reporting. Hoberg et al. (2014) show that product market descriptions contain useful information about a firm's competitive position, in addition to the standard industry classification codes. The MD&A, Legal Proceedings, Quantitative and Qualitative Disclosures about Market Risk, and Risk Factors sections often provide subtle and soft information reflecting management's perceptions of business prospects and risk exposures. Cohen et al. (2020) find that firms' reporting changes are mainly concentrated in these sections. We therefore conjecture that textual changes in these segments have a disproportionate impact on changes in option volatility smirks.

A firm's information environment can also impact the speed of stock price adjustment to new information in 10-K textual changes. For firms with high information asymmetry, information dissemination tends to be noisy, leading stock investors to respond slowly to reporting changes. Sophisticated investors can more effectively leverage their information advantages by trading options on such firms. Cremers and Weinbaum (2010) find that option expensiveness is a significantly stronger predictor of future stock returns among firms with more opaque information environments. Wittenberg-Moerman (2008) also provides evidence that information asymmetry and financial reporting quality play a key role in bond trading. We therefore hypothesize that 10-K textual changes will have a stronger effect on changes in volatility smirks for firms with higher information asymmetry.

Short-sale constraints represent another factor that may influence the impact of 10-K textual changes on options trading. Although short selling provides an alternative avenue for sophisticated investors to act on negative information, options are often used in practice to circumvent short-sale constraints in the securities lending market, given the fluctuating costs of short selling over time and across firms. Even large institutional investors may find it difficult and costly to short certain stocks. Therefore, we expect option volatility smirks to react more strongly to 10-K reporting changes for firms facing higher short selling costs.

We summarize the heterogeneous effects of textual changes with the following set of hypotheses:

Hypothesis H2a: *Option volatility smirks respond more strongly to textual changes in 10-K sections that are more informative about firm fundamentals and subject to greater managerial discretion.*

Hypothesis H2b: *Option volatility smirks respond more strongly to 10-K textual changes among firms with greater information asymmetry and higher short selling costs.*

Finally, we examine whether the "lazy prices" phenomenon documented in Cohen et al. (2020) differs between stocks with and without tradable options. In developing the hypothesis, we assume that sophisticated investors prefer trading options to capitalize on price-relevant information in 10-K textual changes. While the advantages of options trading are well-documented, we acknowledge that sophisticated investors may also exploit their information advantage through stock trading. This is especially relevant for stocks without tradable options, where greater textual changes in 10-K filings may prompt more informed trading in the stock market. As a result, information may be incorporated into stock prices more quickly, leading to a weaker lazy-prices effect. However, the availability of tradable options does not necessarily explain the lazy-prices effect. Even if informed investors mainly use OTM puts to trade optionable firms, the speed of price discovery in the stock market depends on the degree of integration between the two markets. Under perfect integration, the no-arbitrage relation in put-call parity predicts that any information produced through options trading should be swiftly transmitted to the underlying stock prices. In contrast, with limited arbitrage, stock prices could deviate from the values implied by put-call parity for extended periods.

Limited arbitrage between the stock and options markets can arise from a combination of factors including market segmentation, transaction costs, and short sale constraints. First, market segmentation occurs when the two markets attract different types of investors. For example, sophisticated investors may prefer trading options, while inattentive investors mainly trade in the stock market. Although sophisticated investors may participate in both markets, price divergences can still emerge if the marginal investors differ. In such cases, there is often no immediate mechanism to correct the mispricing (Shleifer, 2000). Second, several factors may prevent investors from replicating the payoff of the underlying stock using bonds and call-put option combinations. Transaction costs, in particular, can be substantial in the options market—especially during periods of severe market dislocation—making such arbitrage strategies less feasible (Lamont and Thaler, 2003). Alternatively, there may be hidden values associated with owning shares (Duffie et al., 2002) or options (Detemple and Selden, 1991). Third, short-sale restrictions are among the most commonly cited frictions leading to the breakdown of put-call parity. When short selling becomes too expensive or unavailable, arbitrage opportunities that would normally drive overvalued stock prices toward their options-implied values vanish. For example, Ofek et al. (2004) show that violations of put-call parity are strongly related to short sales constraints as measured by the rebate rate spread.

For optionable stocks with limited arbitrage between the stock and options markets, the predictability of 10-K textual changes on future stock returns is likely to depend on the magnitude of changes in option volatility smirks. Cohen et al. (2020) show that the strong predictive power of 10-K textual changes on future stock returns persists more than six months following the filing's release. Xing et al. (2010) find that option volatility smirks predict stock returns up to six months. This implies that options traders can exploit the information revealed in textual changes before it is slowly incorporated into stock prices. However, it is unlikely that all textual changes carry the same amount of negative information. A more plausible scenario is that some textual changes are more informative than others. If sophisticated investors have the ability to identify and act on information-relevant textual changes, then the predictability on future stock returns should be stronger for changes that are accompanied by large shifts in option volatility smirks.

We state our third set of hypotheses about the return predictability of 10-K textual changes as follows:

Hypothesis H3a: *Compared to stocks without tradable options, the predictability of 10-K textual changes on future stock returns is stronger for optionable stocks due to limited arbitrage.*

Hypothesis H3b: *For optionable stocks, the predictability of 10-K textual changes on future stock returns is stronger when accompanied by larger changes in option volatility smirks.*

III Data and Summary Statistics

In this section, we describe our data sample, key variable definitions, and summary statistics.

III.A Data Sample

Our data sample includes all public firms with information on options trading available during the period from 1996 to 2017 as a result of combining different databases. We obtain options data from OptionMetrics provided by WRDS. For these firms, we download all 10-K filings during the same period from the EDGAR website maintained by the Securities Exchange Commission (SEC) and construct textual similarity measures. Finally, we construct firm-level financial variables, such as market capitalization, book-to-market ratio, stock return, return on assets (ROA), cash holdings, capital expenditure (Capex), and tangibility, based on items from Compustat. All variable definitions can be found in Table A1 of the Appendix. We keep firm-year observations with information on options trading available both before and after the 10-K release. Our final data sample consists of 4,604 unique firms and 28,704 firm-year observations.

III.B Main Variables

Our primary outcome variable is the change in volatility smirk ($\Delta Smirk$) before and after the 10-K release date. Motivated by prior research, we use the smirk change to capture options trading in response to new information disclosed in the most recent 10-K. Specifically, we first calculate daily volatility smirk as the difference in implied volatilities between OTM put options and ATM call options. Following Xing et al. (2010), we classify moneyness between 0.95 and 1.05 as ATM call options and moneyness between 0.80 and 0.95 as OTM put options. For stocks with multiple ATM (OTM) options trading on a given day, we use the average of implied volatilities to compute the difference.

Next, we compute changes in volatility smirk as follows. We keep firm-year observations with volatility smirk information available within three months both before and after the 10-K release date. We use the first available option volatility smirk, which is at least 8 weeks (i.e., 56 days) before the 10-K release date, as the "Baseline Smirk". The baseline smirk is defined sufficiently before the 10-K release to alleviate the concern that trading driven by proprietary information and/or other information events (e.g., earnings announcement) may cause the smirk change ahead of the release date. We use the first available option volatility smirk immediately after the 10-K release as the "Post-release Smirk". The difference between the two smirks reflects changes in options trading surrounding the 10-K release, i.e.,

$\Delta Smirk =$ "$Post - release\ Smirk$" $-$ "$Baseline\ Smirk$".[2] Alternatively, we use the "Post-release Smirk" ($Smirk$) as the outcome variable and obtain similar results.[3]

Our main explanatory variable is a textual similarity measure of 10-Ks between two consecutive fiscal years. Following Brown and Tucker (2011), Peterson et al. (2015), and Cohen et al. (2020), we measure textual changes based on the cosine similarity between the current year's 10-K and the prior year's 10-K. A lower such cosine similarity implies more year-to-year textual changes. Suppose that the two 10-Ks have n unique words,[4] each 10-K is then represented by an n-dimension vector—v_1 for the current year and v_2 for the previous year:

$$v_1 = (w_1, w_2, ..., w_n), \text{ and } v_2 = (z_1, z_2, ..., z_n),$$

where w_i and z_i are counts of each word $i \in [1, n]$. The cosine similarity score is computed as follows:

$$Cosine(\theta) = (v_1/||v_1||)(v_2/||v_2||) = (v_1 \cdot v_2)/(||v_1||||v_2||),$$

[2] For ease of exposition, we multiply $\Delta Smirk$ by 100 in all empirical analyses.

[3] See Table A5 of the Appendix.

[4] Every captured word starts with alphabets; that is, all preceding numbers and special characters are dropped (e.g. in "2-2_ab-1-c", only "ab-1-c" will be recognized as a word). All trailing special characters are removed. However, all special characters between the first alphabet character and last alphanumeric character are preserved (e.g. in "-ex-hi-bit-", "ex-hi-bit" will be captured). Words with numbers only, special characters only, numbers with special characters only, single alphabet, or single alphabet with special characters only (e.g. "a", "a_", "e–") are dropped.

where θ is the angle between v_1 and v_2, $(\cdot)$ is the dot product operator, $||v_1||$ is the vector length of v_1, and $||v_2||$ is the vector length of v_2. This score is bounded between 0 and 1 with a lower score indicating lower similarity.

We retrieve from WRDS SEC Analytic Suite the cosine measures constructed based on the entire 10-K documents.[5] We use the same method to construct cosine measures for individual items in 10-Ks. Other qualitative measures, such as log file size and tone optimism, are also constructed based on 10-Ks following the method of Loughran and McDonald (2011, 2014). Figure 1 shows the time pattern of average annual report cosine similarity for our sample firms. Year-to-year similarity in 10-K filings exhibits a clear upward trend, consistent with the findings in Brown and Tucker (2011) based on a smaller sample.[6]

Insert Figure 1 here

III.C Summary Statistics

Table 1 presents the descriptive statistics for the main variables. Consistent with prior studies (Xing et al., 2010), the post-release option volatility smirks are on average positive at around 5%. The year-to-year changes in option volatility smirks exhibit a slightly negative mean (-0.047%) and median (-0.149%), indicating a modest decrease on

[5]The 10-K filings used to construct the cosine similarity measures include Forms 10-K, 10-K40, 10KSB, and 10KSB40. See Hoberg and Phillips (2016) for a detailed description.

[6]Both Fama-MacBeth regressions (Fama and MacBeth, 1973) and fixed effects panel regressions account for such secular trends.

average after the release of 10-K filings. However, there are considerable variations in smirk changes across firm-years with a high standard deviation of 7.134%. Our primary explanatory variable is the annual report cosine similarity with a mean of 0.971 and a median of 0.988.[7]

Insert Table 1 here

We also report two other variables related to 10-K filings. The first is tone optimism, which represents the number of Loughran-McDonald positive words minus negative words scaled by the total number of words in a filing. The second variable is log file size, which denotes the logarithm of the total number of words in a 10-K filing. Overall, the file tone measure is slightly negative, indicating that managers tend to adopt a neutral or slightly pessimistic tone in their 10-K reports.

IV Empirical Design and Main Results

In this section, we present our research design, main empirical results, and robustness checks. We also conduct additional analyses to explore the mechanism driving our key results.

[7]These statistics are higher than those reported in Cohen et al. (2020) due to the requirement for options trading information for our sample firms and the trend of increasing similarity of 10-K filings in recent years. Approximately one third of the firm-year observations with 10-K release have the option smirk information required for this study.

IV.A Baseline Results

We use both sorting and Fama-MacBeth regressions to examine the response of option volatility smirks to 10-K textual changes. Sorting relies less on parametric assumptions, and Fama-MacBeth regressions allow us to control for various firm-level characteristics that may impact options trading.

We first sort all firms by the annual report cosine similarity into quintile groups and compare average changes in volatility smirks between the top and bottom quintiles. Note that Q1 (Q5) consists of firms with the most (least) textual changes to 10-Ks. Table 2 presents the sorting results. The average changes for Q1 are positive, suggesting an increase in volatility smirk for firms with the most textual changes. By contrast, we find the opposite for Q5. The equal- and value-weighted spreads in volatility smirk changes are -0.299 and -0.361, respectively. These are around 2-2.5 times the median volatility smirk change and are statistically significant for both equal- and value-weighted portfolios. Hence, these results are consistent with Hypothesis H1 that options traders buy more OTM put options on firms with more textual changes.

Insert Table 2 here

We next examine the effect of 10-K textual changes on volatility smirks using the following Fama-MacBeth regression model:

$$\Delta Smirk_{it} = \beta_0 + \beta_1 Cosine_{it} + \phi X_{it} + \varepsilon_{it}, \tag{1}$$

where i indexes a company and t indexes a certain fiscal year. $\Delta Smirk$ is the option volatility smirk right after the 10-K release minus the option volatility smirk at least 8 weeks before the release date.[8] *Cosine*, our main variable of interest, is the cosine similarity between the current and previous fiscal year's 10-Ks. As a lower such cosine similarity implies more textual changes to 10-Ks, which on average contain more negative information regarding future stock returns and firm fundamentals, we expect a negative sign of the estimated coefficient β_1.

The control variables, X, include log file size, tone optimism, log market cap, book-to-market, one-year stock return, ROA, cash, capex, and tangibility. All variables are measured at the previous fiscal year-end. Prior research shows that these firm characteristics could also induce options market reactions and are likely correlated with 10-K textual changes (Ertugrul, Lei, Qiu, and Wan, 2017, Cohen et al., 2020). Large and mature firms tend to have relatively stable operations, lower information asymmetry, and therefore lower downside risk. Firms with high book-to-market and tangible-to-total assets ratios are likely more transparent and easier to value. These firms are therefore less likely to experience widening of volatility smirks driven by investors with information advantages. Moreover, cash holdings and investment expenditures may also affect firm risk, though the direction likely depends on investment efficiency (Cheng, K. Gawande, and Qi, 2021).

Table 3 presents the Fama-MacBeth regression results. Column (1) includes only the cosine similarity measure, and Column (2) adds tone optimism and log file size to

[8] For a robustness check, we use the smirk on the day immediately before the 10-K release as the benchmark and find similar results. We report this result in Table A2 of the Appendix.

control for the effects of other textual information. Column (3) is our main specification and includes firm-level financial variables as additional controls. The estimated effect of textual changes is very similar across different specifications. The coefficient estimates of the cosine measure are all negative and statistically significant at the one percent level. This confirms that more textual changes (proxied by a lower cosine measure) are associated with larger increases in option volatility smirks. Regarding the economic magnitude, if the cosine measure decreases from the 90th percentile (0.998) to the 10th percentile (0.924), the coefficient in Column (3) would imply a twofold increase in the change of option volatility smirks relative to the sample median. This magnitude is quite similar to the one obtained in the sorting analysis.

As to control variables, most of the firm-level financial variables do not bear a significant relation with changes in option volatility smirks. Consistent with our conjecture, we find that firms with higher book-to-market ratios and more tangible assets tend to have significantly smaller smirk changes. Because the estimated effects of these control variables are very similar across regressions, we omit reporting them in the following tables for brevity.

Insert Table 3 here

We conduct several additional tests to ensure that our results are robust to alternative measures of textual changes and econometric specifications. Despite being widely used in measuring textual similarity, the cosine measure is subject to the criticism that it does not have the triangle inequality property and violates the coincidence axiom. To address such concerns, we use three alternative measures of annual report similarity: Jaccard similarity, modified Jaccard similarity, and minimum edit distance. More textual changes to 10-Ks are associated with lower Jaccard or modified Jaccard similarity but higher minimum edit distance. Table A3 of the Appendix presents the Fama-MacBeth regression results, which are consistent with the findings in Table 3. Option volatility smirk changes are negatively related to Jaccard or modified Jaccard similarity, and positively related to minimum edit distance. All of the estimated effects are statistically significant at the one percent level. Therefore, our baseline results are robust to alternative measures of textual similarity.

To ensure that our findings are not affected by unobserved industry and time factors, we estimate pooled regressions with industry (4-digit SIC code) and year fixed effects. We cluster standard errors at the firm level. Table A4 of the Appendix reports the estimation results. Consistent with the baseline results in Table 3, more textual changes to 10-Ks are associated with larger option volatility smirk changes. The coefficients of the cosine measure are statistically significant at the five percent level. Moreover, the magnitudes of estimated coefficients are similar to those obtained from Fama-MacBeth regressions.

IV.B Timing of Changes in Volatility Smirks

To provide further evidence on the timing of changes in volatility smirks, we conduct the Fama-MacBeth regression in equation (1) for each week during the eight-week period before and after the 10-K release. To reduce missing observations, we use the option volatility smirk on the day closest to the 10-K release date in a given week as the weekly measure ($Smirk_{wt}$). Then we compute the weekly change in option volatility smirk ($\Delta Smirk_{wt}$) as the difference between the weekly smirk ($Smirk_{wt}$) and the baseline smirk measured at least eight weeks before the 10-K release date ("*Baseline Smirk*"): $\Delta Smirk_{wt} = Smirk_{wt} - Baseline\ Smirk$.[9]

Insert Table 4 here

Table 4 presents the weekly Fama-MacBeth regression results. Panel A reports the responses in option volatility smirks to textual changes during weeks before and after the 10-K release, respectively. In general, the coefficients of cosine measure have smaller magnitude and lack statistical significance during pre-release weeks. By contrast, the coefficients are much larger in magnitude and statistically significant during each of the first seven weeks after the 10-K release. This suggests that the significant relation between option volatility smirks and 10-K textual changes is mainly driven by post-release trading.

Note that the response in option volatility smirks to 10-K textual changes in week[-2] is statistically significant and has a magnitude comparable to the post-release responses. One interpretation is that some options traders have access to proprietary information before the public release. Although we can not rule out the possibility of

[9] For robustness, we also use the average of all available option volatility smirks in a week to measure weekly smirk levels. The results are similar and available upon request.

informed trading based on private information, it is difficult to reconcile with why such trading occurs in week[-2] but not in week[-1] or other weeks. A more plausible explanation is that the spike in trading activity stems from a separate information event that coincides with the content later disclosed in the 10-K filing. This is likely because many firms disclose their fourth-quarter earnings approximately two weeks before releasing their 10-Ks. As extensively documented in the accounting and finance literature (e.g., Amin and Lee, 1997, Campbell, Ramadorai, and Schwartz, 2009, Truong and Corrado, 2014, Barron, Schneible, and Stevens, 2018), trading around earnings announcements is often heavy in both the stock and options markets, driven by the anticipation and arrival of new information about a firms profitability. The subsequent 10-K filing contains more detailed financial results and price-relevant information that can help investors better understand the earnings result. This drives additional trading during the post 10-K release weeks.

To test this explanation, we exclude firms reporting their fourth quarter earnings within two weeks before 10-K releases and re-estimate the weekly Fama-MacBeth regressions. As shown in Panel B of Table 4, the cosine coefficient for week[-2] is now much smaller in magnitude and no longer statistically significant. By contrast, the post-release cosine coefficients remain large and statistically significant for most weeks. This finding helps address concerns that option volatility smirk changes might be driven by pre-10-K release trading activity. The overall evidence thus supports our hypothesis that options investors gain a competitive advantage through their superior skill in processing public information.

IV.C Mechanism

Our hypothesis about the relation between textual changes in 10-K filings and changes in option volatility smirks relies on several implicit assumptions. These assumptions are grounded in existing literature on the informational content of 10-K textual changes and the determinants of option volatility smirks. In this section, we present additional evidence to support the proposed mechanism.

The key assumption is that 10-K textual changes typically signal negative information about a firm's profitability, leading sophisticated investors to increase their demand for OTM put options.[10] Cohen et al. (2020) provide evidence that greater textual changes predict negative news, including lower future earnings and a higher likelihood of bankruptcy. It is plausible, however, that not all textual changes convey negative news. If options traders truly possess superior information processing skills, then volatility smirk increases should primarily stem from textual changes containing new negative information.

Our first test examines whether it is the negative information that drives the relation between cosine measures and volatility smirk changes in Table 3. Following Cohen et al. (2020), we measure the negative sentiment of changes by calculating the ratio of negative words to total words in the changed text.[11] This measure serves as our proxy for negative news content in textual changes. After dividing our sample at the

[10]OTM options share lottery-like characteristicslow cost with potentially high payoffswhich may attract certain investor types (Cheng, Kong, Lin, Luo, and Zhou, 2025).

[11]Following Cohen et al. (2020)'s approach, we compute change as the number of additions, deletions, and other changes identified from the function "diff" in Unix/Linux. We draw sentiment category identifiers and word lists from Loughran and McDonald (2011).

median value of negative sentiment measure, we run the regression in Table 3 separately for each subsample. As reported in Table 5, the coefficient of cosine measure is much more negative and statistically significant only for the subsample with above-median changes in negative sentiment. In other words, greater textual changes predict increased volatility smirks only when 10-K filing reveals substantially negative information.

Insert Table 5 here

We next test whether more textual changes lead to higher demand for OTM put options. Following Lakonishok, Lee, Pearson, and Poteshman (2007), we compute delta-adjusted open interest by multiplying an options open interest by its delta and dividing the result by the number of outstanding shares of the underlying stock. This adjustment converts open interest into equivalent underlying shares and normalizes it by the total shares outstanding. We then average the delta-adjusted open interest across all OTM put options on the first day after the 10-K release. Table 6 reports the Fama-MacBeth regression results on the relation between the delta-adjusted open interest and the cosine measure. As shown in column (1) for the full sample, the coefficient of cosine is negative and statistically significant, confirming that greater textual changes are associated with higher open interest for OTM put options. Columns (2) and (3) report regression results for above- and below-median negative sentiment subsamples, respectively. The results indicate that the positive relation is primarily driven by textual changes containing substantially negative information. These findings align with our volatility smirk change results reported in Table 5.

Insert Table 6 here

The results in Tables 5 and 6 suggest that negative information disclosed through 10-K textual changes increases the demand for OTM puts, subsequently resulting in higher option prices (as reflected by implied volatility). However, if competitive market makers can perfectly hedge inventory risk, the option supply curves are essentially flat and option prices are determined by no-arbitrage conditions, irrespective of demand pressure. Upward-sloping supply curves will arise if market makers are risk averse and unable to hedge perfectly. Bollen and Whaley (2004) is the first paper to empirically link changes in demand to changes in option prices via limits to arbitrage. Increased demand for certain options results in higher prices to compensate market makers for the elevated hedging risk they bear. As market makers gradually rebalance their portfolios and/or demand pressure subsides, option prices revert at least partially to their previous levels.

Motivated by the findings in Bollen and Whaley (2004), Gârleanu, Pedersen, and Poteshman (2009) formally develop a demand-based option pricing model in which competitive, risk-averse dealers cannot hedge perfectly. The market incompleteness could be driven by various factors, including discrete trading, stochastic volatility, jump risk, and transaction costs. The model predicts that equilibrium option prices increase with demand pressure from end-users. Essentially, higher prices reflect the premium required by market makers to supply the volume of options contracts demanded by end-users. The model remains agnostic regarding the sources of end-users demand pressure—whether motivated by hedging or driven by information. The magnitude and duration of these demand-based price effects depend on the nature of market incompleteness, which in turn determines the speed at which market makers rebalance their inventory positions.[12]

The evidence of our paper is broadly consistent with Bollen and Whaley (2004) and Gârleanu et al. (2009). In Table 4, we examine changes in option volatility smirks during the 8-week period following the 10-K release. To examine the longer-term behavior of volatility smirks, we extend the window to 13 weeks in Table A7, covering approximately one quarter following the 10-K release. We limit the post-release horizon to one quarter to avoid any confounding information events, such as new earnings announcement, that may affect option prices. As shown in Tables 4 and A7, the coefficients for cosine measure are negative and statistically significant for the first 7 weeks and become statistically insignificant after week 8.[13] This suggests that the price impact on OTM puts, driven by informed options traders' demand, persists for several weeks before reverting to pre-10-K release levels. The evidence thus supports the limits-to-arbitrage interpretation, as it takes time for options market makers to rebalance their inventory and/or to attract more market makers to provide liquidity in response to increased demand pressure.

[12]Gârleanu et al. (2009) also presents empirical evidence consistent with that demand pressure from end-users helps explain the pricing levels and skew patterns of index and single-stock options.

[13]Our finding that the price impact on OTM put persists for 7 weeks appears to be relatively long and consistent with demand pressure from informed traders. This is relative to the finding in Bollen and Whaley (2004) that about 20% of change in OTM option volatility reverts in the next day, i.e., about one week for the change to fully revert.

Our final test examines how options trading responds to textual similarities in 10-K filings. If higher textual similarity (fewer textual changes) signals nonnegative or positive news, we expect stronger demand for bullish call options relative to protective puts. This, in turn, should lead to higher implied volatility for calls compared to puts. Similar to the baseline model, we measure the change in reverse volatility smirk from the first day after the 10-K release and the baseline day, which is set at least 8 weeks before the release. Table A6 of the Appendix reports the Fama-MacBeth regression results on the relation between reverse volatility smirks and cosine measures. The coefficients of the cosine measure are positive and statistically significant at the ten percent level.[14] This suggests that higher textual similarities in 10-K filings predict larger increases in reverse volatility smirks. The result contrasts with the changes in volatility smirks reported in Table 3. The evidence is thus consistent with informed investors buying OTM call options to trade on positive signals indicated by minimal or no textual changes in 10-K filings.

In summary, we find that negative information in 10-K textual changes leads to an increase in option volatility smirks. This reflects options traders buying more OTM put options in response to negative news extracted from 10-K textual changes. These results support our conjecture that sophisticated options traders have superior skills in processing public yet complex information disclosed in EDGAR filings.

[14]Compared to Table 3, the relation between reverse volatility smirks and cosine measures is weaker. This can be partly explained by sophisticated investors preferring to trade on positive information directly in the stock market.

V Heterogeneous Effects of Textual Changes

In this section, we explore whether the effects of 10-K textual changes on options trading vary by specific reporting items and across firms with differing information environments and arbitrage constraints. This analysis sheds light on how options traders respond to various types of textual information and diverse trading opportunities.

V.A Textual Changes by Reporting Items

We first explore whether textual changes in the product market description segment, usually displayed in Item 1 (Business) of 10-Ks, are related to options trading. Hoberg et al. (2014) show that product market descriptions contain useful information on the firm's competitive position in the industry, in addition to the standard industry classification codes. We utilize two measures of product market fluidity developed by Hoberg et al. (2014) to capture changes in own and competitors' product descriptions, respectively. To ensure consistent interpretation of textual change coefficient, we use one minus the product market fluidity measure in regressions, so that lower values reflect greater textual changes in product descriptions. Accordingly, a negative coefficient suggests that more such changes are associated with steeper volatility smirks. Table 7 reports the Fama-MacBeth regression results with the product description similarity measures. We find that greater textual changes to a firm's own product descriptions are associated with larger increases in option volatility smirks. In contrast, changes in competitors product descriptions have a limited impact on volatility smirk movements.

Insert Table 7 here

Next, we examine textual changes across different items in 10-K filings, acknowledging that managers have varying levels of discretion over specific sections. For example, Cohen et al. (2020) document that firms' reporting changes are mainly concentrated in Item 7 (Management's Discussion and Analysis), Item 3 (Legal Proceedings), Item 7a (Quantitative and Qualitative Disclosures about Market Risk), and Item 1a (Risk factors). We expect changes in these discretionary items to be more informative and, therefore, to have a greater impact on options trading. To test this, we construct cosine similarity measures for each item based solely on the textual content within the respective section.

Insert Table 8 here

Table 8 shows the coefficients of cosine measures estimated from equation (1) for four 10-K items: Item 7, Item 3, Item 7a, and Item 1a.[15] Consistent with the strong return predictability documented by Cohen et al. (2020), we find that option volatility smirks respond more strongly to textual changes in these discretionary 10-K items. Moreover, the average cosine coefficient for these four Items is significantly larger in magnitude compared to other sections of the 10-K. This difference is statistically

[15]Because the SEC mandates the disclosure of risk factors in Item 1a after the SOX (e.g., Dyer, Lang, and Stice-Lawrence, 2017), we use the post-SOX period for the analysis of Item 1a.

significant at the one percent level.[16] These findings support our Hypothesis H2a, suggesting that options traders actively respond to textual changes in parts of the 10-K that are both more informative about firms fundamentals and more subject to managerial discretion.

V.B Information Environments

Previous studies find that the predictive power of option expensiveness for future stock returns is particularly pronounced among firms with higher levels of information asymmetry (Hu, 2014). This is because option traders are more likely to profit from information advantages in firms with more diverse investor opinions. Therefore, we expect that option volatility smirks will react more strongly to 10-K textual changes for these firms.

Drawing from the existing literature, we use the following two measures to proxy for information asymmetry: abnormal idiosyncratic volatility (AIV) and analyst coverage. Morck, Yeung, and Yu (2000) and Durnev, Morck, and Yeung (2004) relate the idiosyncratic volatility (IV) of stock returns to firm-specific information flows. To isolate abnormal price variation stemming from informed trading, Yang et al. (2020) propose the AIV measure, defined as the IV prior to information events in excess of its normal levels. Specifically, we compute AIV as the difference between IV measured during pre-earnings-announcement windows and IV measured during non-earnings-announcement periods using daily returns for each fiscal year. Following

[16]This test result is available upon request.

Yang et al. (2020), we define the pre-earnings-announcement periods as the 5 business days preceding each of the 4 earnings announcement dates. The non-earnings-announcement periods comprise all days excluding the 11-day window surrounding each earnings announcement. Our next measure is analyst coverage (Hu, 2014), defined as the number of analysts following a firm, with data sourced from the Institutional Brokers' Estimate System (I/B/E/S). We compute these variables at the end of the previous fiscal year and partition the sample into high and low information asymmetry groups based on their respective median values.

Insert Table 9 here

Table 9 reports the regression results based on equation (1), estimated separately for the high versus low information asymmetry subsamples. The coefficients for the cosine measure are consistently negative, but statistically significant only for firms characterized by high AIV and low analyst coverage. Hence, an increase in textual changes leads to significantly higher volatility smirk only among firms with greater information asymmetry. This finding is consistent with the conjecture in Hypothesis H2b.

V.C Short-Sale Constraints

Options trading is not the only way sophisticated investors can exploit negative information in 10-K textual changes. Short selling offers informed traders another method to express bearish views. However, unlike options trading on centralized exchanges, the securities lending market is fragmented, and even large institutions often face high borrowing costs or may be unable to secure shares to borrow. Hence, informed investors are more likely to rely on options when short-sale constraints are binding. Consistent with this, Hu (2014) finds that options trading conveys more information about future stock prices for firms facing greater short-sale constraints. We conjecture in Hypothesis H2b that the effect of 10-K textual changes on option volatility smirks is stronger when short-sale constraints are more severe.

Following D'Avolio (2002), Asquith, Pathak, and Ritter (2005), and Hu (2014), we use institutional ownership as a proxy for short-selling costs. This is because lower institutional ownership is often associated with substantially higher costs for short sellers. We divide the sample into high-cost and low-cost groups based on the median of fiscal year-end institutional ownership and estimate equation (1) separately for each group. Table 10 presents the regression results. The coefficient estimates of the cosine similarity measure are negative and statistically significant for both groups. However, the effect of 10-K textual changes is approximately four times larger for the high-cost group compared to the low-cost group. This finding supports that informed investors rely more on OTM puts to trade on negative information for firms facing greater short-sale constraints.

Insert Table 10 here

VI Return Predictability of Textual Changes and Option Volatility Smirk Changes

Our tests thus far have focused exclusively on stocks with tradable options, assuming that informed investors prefer trading options. However, informed investors can also trade directly in the stock market, which is particularly relevant for stocks without tradable options. The presence of informed trading in either or both markets can have different implications for the lazy-prices effect between optionable and non-optionable stocks. In this section, we extend Cohen et al. (2020) by examining the "lazy prices" phenomenon separately for stocks with and without tradable options. We also expand the return predictability test to include option volatility smirk changes in addition to 10-K textual changes.

To examine whether the lazy-prices effect depends on the availability of tradable options, we follow a procedure similar to Flippou, Garcia-Ares, and Zapatero (2022) and match stocks with and without options based on industry, size, trading volume, and return volatility. For each optionable stock, we identify matched non-optionable stocks that have similar average market capitalization, trading volume, and return volatility during the 12-month period before the optionable stock begins option trading. Specifically, we require the distance between the optionable stock and the matched non-optionable stock to be less than 10% of the standard deviation of the predicted propensity scores.[17] We then require the matched non-optionable stocks to be in the

[17]In untabulated analysis, our results remain robust under alternative matching criteria.

same sector based on the Fama-French 48 industry classifications and retain the 3 stocks with the nearest propensity scores.

We construct quintile portfolios based on 10-K textual changes, separately for optionable and non-optionable stocks. Following Cohen et al. (2020), we compute monthly quintiles using the prior month's distribution of cosine similarity measures. Stocks enter the portfolio one month after their 10-K release and remain for nine months. Portfolios are rebalanced monthly, and value-weighted returns (multiplied by 100) are calculated. Portfolio performance is evaluated using excess returns, as well as alphas from Fama-French three-factor (market, size, and value), four-factor (market, size, value, and momentum), and five-factor (market, size, value, momentum, and liquidity) models.[18] Quintile 1 (Q1) contains firms with the lowest similarity between current and prior 10-Ks, while Quintile 5 (Q5) includes those with the highest similarity. "Q5 - Q1 represents a long-short portfolio that take a long position in Q5 and a short position in Q1.

The results in Table 11 show that, consistent with the prediction in H3a, the lazy-prices effect is primarily driven by optionable stocks and is insignificant for stocks without options. As shown in Panel A, the average risk-adjusted returns for Q1 are all negative and statistically significant at the one percent level. This suggests that more textual changes predict lower future returns for optionable stocks. The long-short portfolio generates on average 46 to 58 basis points abnormal return per month,

[18] Factors and portfolios are constructed following prior studies (Fama and French, 1993, Carhart, 1997, Pástor and Stambaugh, 2003). Monthly factors are downloaded from Kenneth French and Robert Stambaugh's websites.

statistically significant at the five percent level. In contrast, for non-optionable stocks in Panel B, the risk-adjusted returns for Q1 and the long-short portfolio are no longer statistically significant. The absence of a lazy-prices effect for stocks without options suggests that when options trading is unavailable, more informed trading based on textual changes occurs directly in the stock market. Such trading leads to a faster stock price adjustment to new information in textual changes and therefore attenuates the lazy-prices effect. The contrasting result for optionable stocks suggests that informed investors prefer to trade on negative information using OTM put options (Black, 1975, Easley et al., 1998), and that limited arbitrage contributes to a delayed stock price response to 10-K textual changes.

Insert Table 11 here

One remaining question is why sophisticated stock market investors appear to overlook the information embedded in option quotes. Notably, this information transfer does not rely on active arbitrage between the two markets; instead, it hinges on the attentive monitoring of option signals by stock market makers. We attribute this phenomenon to the lack of liquidity and price noise prevalent in many individual equity options. For example, Duarte, Jones, and Wang (2024) document that most individual equity options trade infrequently, with the top quintile stocks accounting for approximately 90% of option trading volume. Moreover, options tend to have wide closing bid-ask spreads—averaging 12-13% of the quote midpoint—compared to an average of 0.15% for the underlying stocks. Further, the spreads are particularly substantial for deep OTM options. Consequently, the excessive noise inherent in OTM put prices may impede the ability of stock market makers to extract informative signals from changes in volatility smirk.

We acknowledge that the extent of informational segmentation likely depends on the precision of price signals. Extant research shows that firm size is an important determinant of option liquidity (Christoffersen, Goyenko, Jacobs, and Karoui, 2018)—larger firms generally have more actively traded options and receive more intensive monitoring from market participants. We therefore conjecture that information dissemination from options to stock markets is more efficient for large-cap stocks. To test this prediction, we sort optionable stocks into large- and small-cap groups based on the median fiscal-year-end market capitalization and examine the lazy-prices effect for each subsample. In addition, we distinguish between S&P 500 and non-S&P 500 stocks, expecting the lazy-prices effect to be stronger for small-cap and non-S&P 500 stocks.

Table 12 presents the results. As shown in Panel A, there is no significant difference in returns between Q1 and Q5 for large-cap stocks. In contrast, Panel B shows that Q1 significantly underperforms Q5 by about 60 bps per month in risk-adjusted returns for small-cap stocks. A similar pattern manifests in Panels C and D, where the effect is significant only for non-S&P 500 stocks. These findings suggest that the lazy-prices effect is most evident in less transparent segments, where pervasive price noise and limited monitoring compromise price discovery.

Insert Table 12 here

Next, we examine whether the return predictability of 10-K textual changes are stronger for stocks with larger changes in option volatility smirks. Cohen et al. (2020) show that textual changes have strong predictive power on future stock returns up to 6 months after the 10-K release and the effect exhibits no reversal. Xing et al. (2010) find similar return predictability for option volatility smirks. Our baseline result suggests that more textual changes predict a significant increase in volatility smirk immediately after the 10-K release. If textual changes indeed convey negative information on firm fundamentals that motivates options trading, these changes should have stronger predictive power for future stock returns when accompanied by larger increases in option volatility smirks.

We double-sort stocks into portfolios based on the prior months distribution of similarity measures and changes in volatility smirks. Stocks are sorted into quintile groups based on the same procedure as in Table 11. In addition, we independently sort stocks into high-smirk (HS) and low-smirk (LS) change groups, where "high vs. low is relative to the median change of volatility smirk after the 10-K release. For each of the 2-by-5 portfolios, we report value-weighted portfolio excess returns and alphas estimated using various factor models across 1-, 3-, 6-, 9-, and 12-month horizons.

Table 13 presents the results of the double-sorted portfolio analysis. We observe that stocks with low volatility smirk changes tend to outperform those with high smirk changes. In particular, within the low cosine (Q1) group, HS stocks underperform LS stocks across all horizons. The difference becomes statistically significant for the 6-, 9-, and 12-month horizons. For example, HS stocks underperform LS stocks by 2-3% for the 12-month period. This finding aligns with Xing et al. (2010) that the shape of volatility smirk has significant cross-sectional predictive power for future stock returns.

We are particularly interested in the Q1-HS portfolio (high textual changes and high smirk changes), which exhibits significantly negative alphas and underperforms other portfolios. Specifically, the Q1-HS portfolio significantly underperforms the Q5-LS portfolio (low textual changes and low smirk changes) across all horizons. The magnitude of underperformance grows from 50-70 bps over one month to 6-7% over 12 months. By contrast, we find no significant difference between the Q1-LS and Q5-HS portfolios. This indicates that the underperformance of stocks with the most textual changes in 10-K filings is mainly driven by those that also exhibit the steepest increases in volatility smirks. Moreover, there is no evidence of return reversal, consistent with the conjecture that the combination of 10-K textual changes and higher volatility smirks contains new negative information regarding future firm performance.[19]

Insert Table 13 here

[19]To provide additional evidence of informed options trading in a different context, we examine whether changes in option volatility smirks can predict the post-earnings announcement drift—the tendency for stock prices to continue moving in the direction of earnings surprises for an extended period. Specifically, we employ a double-sorting portfolio approach to analyze the return predictability of volatility smirk changes following earnings announcements. We measure earnings surprises using standardized unexpected earnings (SUE) and sort stocks into quintile portfolios based on the prior month's SUE distribution. We also classify stocks into high-smirk (HS) and low-smirk (LS) change groups based on changes in volatility smirks post-announcement. Following the methodology used in Table 13, we calculate average monthly portfolio excess returns and alphas for each of the 2-by-5 portfolios. The results, presented in Table A8, show that firms with the most negative earnings surprises (Q1) and high-smirk changes (HS) generate significantly negative alphas and underperform those with the most positive earnings surprises (Q5) and low-smirk changes (LS).

Table 14 presents return predictability regressions that include both the cosine measure and the change in volatility smirk, along with an interaction term between the two. We examine return predictability across multiple time horizons. For each window, we present results first with only the cosine measure, then with the addition of the volatility smirk change and the interaction term.

The cosine measure yields positive and statistically significant coefficients (except for 1- and 3-month horizons), with the magnitude increasing over the predicative horizon up to 12 months following the 10-K release. This suggests that firms with more textual changes generally earn lower future returns. However, the statistical significance of the cosine measure disappears once we control for volatility smirk changes and the interaction term. In comparison, we find strong evidence that an increase in the volatility smirk predicts a significant decline in future returns. Moreover, the interaction term is positive and statistically significant, indicating that firms with both extensive textual changes AND larger increases in the volatility smirk exhibit the most pronounced declines in future returns. This finding aligns with the double-sorted portfolio results in Table 13. Consistent with hypothesis H3b, these results suggest that the joint signal from the cosine measure and volatility smirk changes provides a stronger indicator of negative news regarding future firm profitability.[20]

[20] We also estimate the return predictability regression separately for short-dated (8-60 days), intermediate-dated (61-180 days), and long-dated (181-360 days) options. Overall, the results are similar to those in Table 14 and remain consistent across different maturities. Notably, the predictive power of smirk change—and its interaction with the cosine measure—is strongest for intermediate-dated options. Options investors may favor these intermediate maturities, as they provide an ideal window for stock prices to incorporate new information in 10-K textual changes. In addition, we examine shifts in return predictability following the initial publication of Cohen et al. (2020) on SSRN in 2010. We estimate the

Insert Tables 14 here

In summary, our findings on return predictability are twofold. First, the "lazy prices" phenomenon documented by Cohen et al. (2020) appears to be primarily driven by stocks with tradable options. This suggests that informed investors prefer the options market to exploit their information advantage, while limited arbitrage delays the incorporation of this information into stock prices. Second, the predictive power of 10-K textual changes for future stock returns is especially pronounced when accompanied by substantial changes in option volatility smirks.

VII Conclusions

This paper investigates whether options traders demonstrate timely reactions to the negative public information contained in 10-K textual changes—in contrast to the "lazy prices" phenomenon documented for stock market investors in Cohen et al. (2020). We find that textual changes containing new negative information lead to a significant increase in option volatility smirks after the 10-K release, reflecting options traders' concerns about future negative price movement. By contrast, the evidence on pre-release volatility smirk changes is largely muted. Collectively, the evidence suggests that

return predictability regression separately for the pre- and post-2010 periods and find that the predictive power of cosine measures declines after 2010. This suggests that greater awareness of the information content contained in 10-K textual changes—following the SSRN publication—may have attenuated return predictability. These results are detailed in Tables A9 and A10 of the Appendix.

options traders possess superior skills in processing information from public filings and extracting profitable trading signals. Our results are robust to different measures of annual report similarity as well as fixed effects panel regressions.

We next examine whether the effect of textual changes varies in informational content and across firms with different characteristics. We find that option volatility smirks respond more strongly to textual changes that are more informative about firm fundamentals and subject to more managerial discretion, e.g., changes in sections on product market description, MD&A, Legal Proceedings, Quantitative and Qualitative Disclosures about Market Risk, and Risk Factors. Moreover, the effect is mainly concentrated among firms with higher information asymmetry and greater short-sale constraints.

Finally, we investigate the return predictability of 10-K textual changes and option volatility smirk changes. Our findings are twofold. First, the lazy-prices effect documented in Cohen et al. (2020) is mainly driven by optionable stocks. Our interpretation is that such stocks attract informed investors to the options market, which—coupled with limits to arbitrage—delays the incorporation of new information into stock prices. Second, we find that textual changes have a greater predictive power for future stock returns when accompanied by large increases in option volatility smirks. This suggests that textual changes combined with volatility smirk changes provide a more potent indicator of negative information regarding future firm profitability.

Taken together, our results demonstrate that sophisticated investors possess the ability to analyze complex 10-K textual changes and capitalize on these insights through options trading. This information advantage is particularly pronounced among firms with asymmetric information environments and high short-selling costs. Collectively, our study sheds light on the "lazy prices" phenomenon and highlights options traders' superior ability to process public information.

References

Alldredge, D. and D. Cicero. “Attentive insider trading.” *Journal of Financial Economics*, 115 (2015): 84–101.

Amin, K. and C. Lee. “Option trading, price discovery, and earnings news dissemination.” *Contemporary Accounting Research*, 14 (1997): 153–192.

Asquith, P., P. Pathak, and J. Ritter. “Short interest, institutional ownership, and stock returns.” *Journal of Financial Economics*, 78 (2005): 243–276.

Augustin, P., M. Brenner, and M.G. Subrahmanyam. “Informed options trading prior to takeover announcements: Insider trading?” *Management Science*, 65 (2019): 5449–5956.

Augustin, P. and M.G. Subrahmanyam. “Informed options trading before corporate events.” *Annual Review of Financial Economics*, 12 (2020): 5449–5956.

Bali, T. G. and S. Murray. “Does risk-neutral skewness predict the cross-section of equity option portfolio returns?” *Journal of Financial and Quantitative Analysis*, 48 (2013): 11–45.

Barron, O.E., R.A. Schneible, and D.E. Stevens. “The changing behavior of trading volume reactions to earnings announcements: Evidence of the increasing use of accounting earnings news by investors.” *Contemporary Accounting Research*, 35 (2018): 1651–1674.

Bates, D. S. “The crash of ’87: Was it expected? The evidence from options markets.” *Journal of Finance*, 46 (1991): 1009–1044.

Black, F. “Fact and fantasy in the use of options.” *Financial Analysts Journal*, 31 (1975): 36–41, 61–72.

Bollen, N.P.B. and R.E. Whaley. “Does net buying pressure affect the shape of implied volatility Functions?” *Journal of Finance*, 59 (2004): 711–753.

Brown, S. and J. Tucker. “Large-sample evidence on firms’ year-over-year MD&A modifications.” *Journal of Accounting Research*, 49 (2011): 309–346.

Campbell, J.Y., T. Ramadorai, and A. Schwartz. “Caught on tape: Institutional trading, stock returns, and earnings announcements.” *Journal of Financial Economics*, 92 (2009): 66–91.

Cao, C., Z. Chen, and J. M. Griffin. “Informational content of option volume prior to takeovers.” *Journal of Business*, 78 (2005): 1073–1109.

Carhart, M.M. “On persistence in mutual fund performance.” *Journal of Finance*, 52 (1997): 57–82.

Chakravarty, S., H. Gulen, and S. Mayhew. “Informed trading in stock and option markets." *Journal of Finance*, 59 (2004): 1235–1257.

Cheng, H., S. Ongena K. Gawande, and S. Qi. "Connected banks and economic policy uncertainty." *Journal of Financial Stability*, 56 (2021).

Cheng, H., L. Kong, T.-C. Lin, Y. Luo, and N. Zhou. "Lottery-like features and mutual fund performance-flow sensitivity." *Financial Management*, 54 (2025): 493–517.

Christoffersen, P., R. Goyenko, K. Jacobs, and M. Karoui. "Illiquidity premia in the equity options market." *Review of Financial Studies*, 31 (2018): 811–851.

Cohen, L. and A. Frazzini. "Economic links and predictable returns." *Journal of Finance*, 63 (2008): 1977–2011.

Cohen, L., C. Malloy, and Q. Nguyen. "Lazy prices." *Journal of Finance*, 75 (2020): 1371–1415.

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

D'Avolio, G. "The market for borrowing stock." *Journal of Financial Economics*, 66 (2002): 271–306.

Detemple, J. and L. Selden. "A general equilibrium analysis of option and stock market interactions." *International Economic Review*, 32 (1991): 279–303.

Drake, M.S., D.T. Roulstone, and J.R. Thornock. "The determinants and consequences of information acquisition via EDGAR." *Contemporary Accounting Research*, 32 (2015): 1128–1161.

Duarte, J., C.S. Jones, and J.L. Wang. "Very noisy option prices and inference regarding the volatility risk premium." *Journal of Finance*, 79 (2024): 3581–3621.

Duffie, D., N. Gârleanu, and L.H. Pedersen. "Securities lending, shorting, and pricing." *Journal of Financial Economics*, 66 (2002): 307–339.

Dugast, J. and T. Foucault. "Equilibrium data mining and data abundance." *Journal of Finance*, 80 (2025): 211–258.

Durnev, A., R. Morck, and B. Yeung. "Valueenhancing capital budgeting and firmspecific stock return variation." *Journal of Finance*, 59 (2004): 65–105.

Dyer, T., M. Lang, and L. Stice-Lawrence. "The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation." *Journal of Accounting and Economics*, 64 (2017): 221–245.

Easley, D., M. O'Hara, and P. S. Srinivas. "Option volume and stock prices: Evidence on where informed traders trade." *Journal of Finance*, 53 (1998): 431–465.

Engelberg, J., A. Reed, and M. Ringgenberg. "How are shorts informed? Short sellers, news, and information processing." *Journal of Financial Economics*, 105 (2012):

260–278.

Ertugrul, M., J. Lei, J. Qiu, and C. Wan. "Annual report readability, tone ambiguity, and the cost of borrowing." *Journal of Financial and Quantitative Analysis*, 502 (2017): 811–836.

Fama, E.F. and K.R. French. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, 33 (1993): 607–636.

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

Flippou, I., P.A. Garcia-Ares, and F. Zapatero. "Demand for lotteries: The choice between stocks and options." *Working paper*, (2022).

Gârleanu, N., L.H. Pedersen, and A.M. Poteshman. "Demand-based option pricing." *Review of Financial Studies*, 22 (2009): 4259–4299.

Ge, L., T. Lin, and N.D. Pearson. "Why does the option to stock volume ratio predict stock returns?" *Journal of Financial Economics*, 120 (2016): 601–622.

Gibbons, B., P. Iliev, and J. Kalodimos. "Analyst information acquisition via EDGAR." *Management Science*, 67 (2021): 811–836.

Hoberg, G. and G. Phillips. "Text-based network industries and endogenous product differentiation." *Journal of Political Economy*, 124 (2016): 1423–1465.

Hoberg, G., G. Phillips, and N. Prabhala. "Product market threats, payouts, and financial flexibility." *Journal of Finance*, 69 (2014): 293–324.

Hu, J. "Does option trading convey stock price information?" *Journal of Financial Economics*, 111 (2014): 625–645.

Kim, J. and L. Zhang. "Financial reporting opacity and expected crash risk: Evidence from implied volatility smirks." *Contemporary Accounting Research*, 31 (2014): 851–875.

Lakonishok, J., I. Lee, N.D. Pearson, and A.M. Poteshman. "Option market activity." *Review of Financial and Studies*, 20 (2007): 813–857.

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

Li, G., T. Lin, and N. Pearson. "Why does the option to stock volume ratio predict stock returns?" *Journal of Financial Economics*, 120 (2016): 601–622.

Loughran, T. and B. McDonald. "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks." *Journal of Finance*, 66 (2011): 35–65.

Loughran, T. and B. McDonald. "Measuring Readability in Financial Disclosures." *Journal of Finance*, 69 (2014): 1643–1671.

Menzly, L. and O. Ozbas. "Market segmentation and cross-predictability of returns." *Journal of Finance*, 65 (2010): 1555–1580.

Morck, R., B. Yeung, and W. Yu. "The information content of stock markets: Why do emerging markets have synchronous stock price movements?" *Journal of Financial Economics*, 58 (2000): 215–260.

Ofek, E., M. Richardson, and R.F. Whitelaw. "Limited arbitrage and short sales restrictions: Evidence from the options markets." *Journal of Financial Economics*, 74 (2004): 305–342.

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

Pan, J. and A. M. Poteshman. "The information in option volume for future stock prices." *Review of Financial Studies*, 19 (2006): 871–908.

Pástor, L. and R. F. Stambaugh. "Liquidity risk and expected stock returns." *Journal of Political Economy*, 111 (2003): 642–685.

Peterson, K., R. Schmardebeck, and T.J. Wilks. "The earnings quality and information processing effects of accounting consistency." *The Accounting Review*, 90 (2015): 2483–2514.

Shleifer, A. *Inefficient markets: An introduction to behavioral finance*. Oxford University Press, 2000.

Truong, H.A. and C.J. Corrado. "Options trading volume and stock price response to earnings announcements." *Review of Accounting Studies*, 19 (2014): 161–209.

Wittenberg-Moerman, R. "The role of information asymmetry and financial reporting quality in debt trading: Evidence from the secondary loan market." *Journal of Accounting and Economics*, 46 (2008): 240–260.

Xing, Y., X. Zhang, and R. Zhao. "What does the individual option volatility smirk tell us about future equity returns?" *Journal of Financial and Quantitative Analysis*, 45 (2010): 641–662.

Yang, Y. C., B. Zhang, and C. Zhang. "Is information risk priced? Evidence from abnormal idiosyncratic volatility." *Journal of Financial Economics*, 135 (2020): 528–554.

Figure 1: The Time Pattern of Average Annual Report Similarity

Note: This figure shows the time-series pattern of annual cosine similarity measures averaged across all sample firms. The cosine similarity measure for a firm quantifies textual changes between the current year's 10-K and the prior year's 10-K. A lower value implies more year-to-year textual changes.

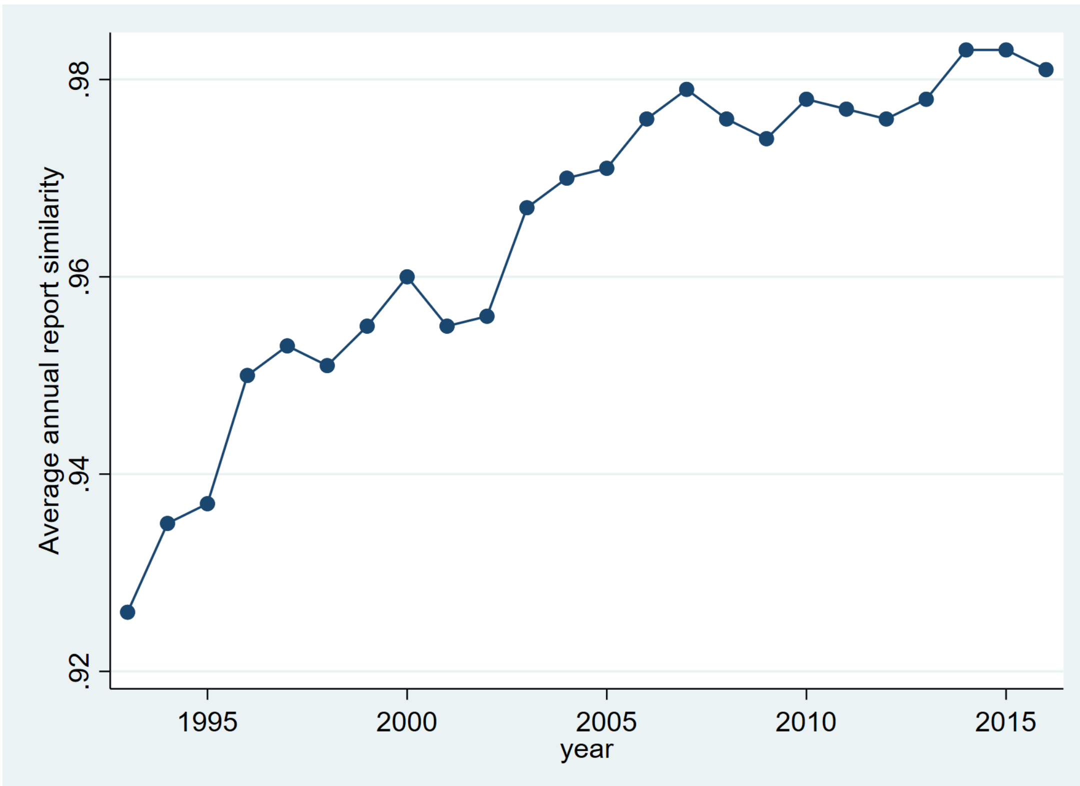

Table 1: Descriptive Statistics

Note: This table reports summary statistics for key variables including mean, standard deviation (stdev), median, 10th percentile (P10), and 90th percentile (P90). For each variable, summary statistics are computed based on time-series averages for all sample firms during the period from 1996 to 2017. Variable definitions are provided in Appendix Table A1.

	mean	stdev	median	P10	P90
Volatility smirk (%)	5.300	5.638	4.835	0.543	10.325
Volatility smirk change (%)	-0.047	7.134	-0.149	-6.047	6.120
Cosine	0.971	0.044	0.988	0.924	0.998
Tone optimism	-0.008	0.005	-0.008	-0.015	-0.002
Log file size	14.910	1.624	14.709	12.661	17.006
Market cap ($million)	16,752.719	51,428.039	4,616.951	803.243	37,760.941
Book-to-market	0.413	0.303	0.345	0.081	0.870
Return	0.266	0.830	0.142	-0.302	0.826
ROA	0.043	0.343	0.049	-0.039	0.139
Cash	0.180	0.204	0.099	0.011	0.483
Capex	0.052	0.062	0.034	0.003	0.118
Tangibility	0.247	0.239	0.162	0.016	0.645

Table 2: The Effect of Textual Changes on Option Volatility Smirk Changes: Cosine-Sorted Portfolios

Note: This table reports average option volatility smirk changes for portfolios sorted by the cosine similarity measure. For each fiscal year, firms are sorted into quintile portfolios based on their fiscal-year end cosine similarity measures. For each portfolio, equal- and value (market capitalization)-weighted averages are computed for each fiscal year and the time-series averages over the sample period are reported. The mean difference between Q5 and Q1 and the corresponding t-statistics are reported in the last row. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var.:	Equal-weighted	Value-weighted
Volatility smirk change	1	2
Q1	0.1367	0.0747
Q2	-0.0432	-0.1078
Q3	-0.0510	-0.1163
Q4	-0.1724	0.0293
Q5	-0.1622	-0.2867
Q5–Q1	-0.2990**	-0.3614***
	(-2.2772)	(-4.4682)

Table 3: The Effect of Textual Changes on Option Volatility Smirk Changes

Note: This table reports the Fama-MacBeth regression results on the effect of textual changes on option volatility smirk changes. The dependent variable is the option volatility smirk change. The key explanatory variable is the cosine similarity measure on 10-K textual changes. Control variables include log file size, tone optimism, log market cap, book-to-market, one-year stock return, ROA, cash, capex, and tangibility. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var: Volatility smirk change	1	2	3
Cosine	-2.964***	-3.200***	-2.944***
	(-3.13)	(-3.52)	(-3.11)
Tone optimism		8.771	12.034
		(1.52)	(1.67)
Log file size		-0.220***	-0.119
		(-3.63)	(-1.51)
Log Market cap			-0.031
			(-0.54)
Book-to-market			-0.498***
			(-3.44)
Return			0.026
			(0.26)
ROA			0.198
			(0.68)
Cash			-0.275
			(-1.01)
Capex			1.250
			(1.27)
Tangibility			-0.507*
			(-1.78)
N	28,704	28,704	28,704
Average R-squared	0.0009	0.0029	0.0156

Table 4: The Effect of 10-K Textual Changes on Option Volatility Smirk Changes Before and After Release

Note: This table reports the Fama-MacBeth regression results on the effect of textual changes on option volatility smirk changes for each week during the eight-week period before and after the 10-K release. Panel A includes the entire sample. Panel B excludes cases where earnings announcements occur within two weeks prior to the 10-K release date. The dependent variable is the weekly option volatility smirk change measured as the difference between volatility smirk measured on the day nearest to the 10-K release date and the baseline smirk measured at least eight weeks before the 10-K release date. The key explanatory variable is the cosine similarity measure on 10-K textual changes. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Panel A: Trading Surrounding the Release of 10-Ks								
	week[-8]	week[-7]	week[-6]	week[-5]	week[-4]	week[-3]	week[-2]	week[-1]
Cosine	-0.513	-1.460	-0.263	-1.301*	-0.606	-1.991*	-2.424***	-0.760
	(-1.13)	(-1.56)	(-0.40)	(-1.85)	(-0.70)	(-1.76)	(-3.33)	(-1.24)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	26,067	25,831	25,622	25,538	25,467	25,368	25,250	25,222
Average R-squared	0.0124	0.0117	0.0134	0.0138	0.0133	0.0129	0.0140	0.0141
	week[1]	week[2]	week[3]	week[4]	week[5]	week[6]	week[7]	week[8]
Cosine	-2.274**	-2.618**	-1.036*	-2.493***	-2.058**	-1.609**	-2.021**	-0.810
	(-2.60)	(-2.61)	(-1.93)	(-3.02)	(-2.76)	(-2.52)	(-2.66)	(-0.99)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	25,156	24,996	25,127	25,131	25,095	25,109	25,121	25,015
Average R-squared	0.0152	0.0178	0.0154	0.0161	0.0162	0.0163	0.0141	0.0142
Panel B: Trading Without Earnings Announcements								
	week[-8]	week[-7]	week[-6]	week[-5]	week[-4]	week[-3]	week[-2]	week[-1]
Cosine	-0.526	-1.890	0.337	-0.922	0.234	-1.263	-0.927	-0.420
	(-0.57)	(-1.29)	(0.53)	(-1.01)	(0.15)	(-1.25)	(-0.99)	(-0.59)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	21,221	21,034	20,899	20,784	20,721	20,657	20,481	20,437
Average R-squared	0.0185	0.0137	0.0151	0.0163	0.0167	0.0168	0.0203	0.0205
	week[1]	week[2]	week[3]	week[4]	week[5]	week[6]	week[7]	week[8]
Cosine	-2.605**	-3.057**	-0.803	-1.456**	-1.686**	-1.024*	-1.624*	0.288
	(-2.75)	(-2.15)	(-1.13)	(-2.13)	(-2.24)	(-1.83)	(-1.97)	(0.43)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	20,443	20,301	20,411	20,445	20,419	20,428	20,466	20,326
Average R-squared	0.0217	0.0248	0.0227	0.0253	0.0246	0.0234	0.0171	0.0166

Table 5: Analyses of Option Volatility Smirk Changes Based on Sentiment Subsamples

Note: This table reports the Fama-MacBeth regression results of option volatility smirk changes based on sentiment subsamples. Columns (1) and (2) are divided based on changes in negative sentiment, which is calculated as the ratio of negative words in textual changes to total words in textual changes (Cohen et al., 2020). Column (1) contains cases with negative sentiment greater than or equal to the median, while Column (2) contains cases with negative sentiment below the median. The dependent variable is the option volatility smirk change. The key explanatory variable is the cosine similarity measure on 10-K textual changes. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var: Volatility smirk change	High neg. sentiment	Low neg. sentiment
Cosine	-4.359**	-1.109
	(-2.68)	(-1.14)
Control variables	Yes	Yes
N	14,351	14,351
Average R-squared	0.0331	0.0273

Table 6: Analyses of Option Open Interest Based on Sentiment Subsamples

Note: This table reports the Fama-MacBeth regression results of option open interest based on sentiment subsamples. Column (1) includes the entire sample. Columns (2) and (3) are divided based on changes in negative sentiment, which is calculated as the ratio of negative words in textual changes to total words in textual changes (Cohen et al., 2020). Column (2) contains cases with negative sentiment greater than or equal to the median value, while Column (3) contains cases with negative sentiment below the median. The dependent variable is the delta-adjusted open interest of OTM put options on the first day of trade after the 10-K release. The delta-adjusted open interest is calculated by multiplying an option's daily open interest by its delta and dividing by the underlying stock's outstanding shares (Lakonishok et al., 2007). The key explanatory variable is the cosine similarity measure on 10-K textual changes. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var: Option open interest	Full sample	High neg. sentiment	Low neg. sentiment
Cosine	-0.013***	-0.022**	-0.006
	(-3.20)	(-2.47)	(-0.82)
Control variables	Yes	Yes	Yes
N	28,052	14,049	14,003
Average R-squared	0.1459	0.1569	0.1431

Table 7: The Effects of Product Market Description Changes on Option Volatility Smirk Changes

Note: This table reports the Fama-MacBeth regression results on the effects of textual changes in product market descriptions on option volatility smirk changes. The dependent variable is the option volatility smirk change. The key explanatory variables are two measures of product market fluidity from Hoberg et al. (2014): own product description and peer product description. Both measures are subtracted from 1 to make the coefficient interpretation consistent with the cosine similarity measure. Other control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var: Volatility smirk change	1	2
Self product des. Similarity	-0.520*	-0.529*
	(-1.72)	(-1.86)
Peer product des. similarity		-0.172
		(-0.11)
Control variables	Yes	Yes
N	27,942	27,942
Average R-squared	0.0170	0.0179

Table 8: The Effects of Textual Changes to Information Intensive and Discretionary Items

Note: This table reports the Fama-MacBeth regression results on the effects of textual changes in information intensive and discretionary items: Item 7 (Management's Discussion and Analysis), Item 3 (Legal Proceedings), Item 7a (Quantitative and Qualitative Disclosures about Market Risk), and Item 1a (Risk Factors). The dependent variable is the option volatility smirk change. The key explanatory variables are the cosine similarity measures constructed based on textual changes to each of these four items. Other control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var:	MD&A	Legal proceeding	Market risk	Risk factors
Volatility smirk change	1	2	3	4
Cosine for items	-0.267**	-0.254***	-0.339***	-0.377***
	(-2.74)	(-3.58)	(-3.79)	(-4.42)
Control variables	Yes	Yes	Yes	Yes
N	28,703	28,703	28,703	20,908
Average R-squared	0.0156	0.0155	0.0155	0.0167

Table 9: The Effect of Textual Changes on Option Volatility Smirk Changes by Information Environments

Note: This table presents the Fama-MacBeth regression results on how the impact of textual changes on option volatility smirk changes varies across firms' information environments. Information asymmetry is measured using two proxies: abnormal idiosyncratic volatility (Yang et al., 2020) and analyst coverage (Hu, 2014). Firms are classified into high and low information asymmetry groups based on the respective median values at the end of the previous fiscal year. The dependent variable is the option volatility smirk change, and the key explanatory variable is the cosine similarity measure. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. var.:	Low AIV	High AIV	High analyst	Low analyst
Volatility smirk change	1	2	3	4
Cosine	-2.401	-4.120***	-1.206	-3.732**
	(-1.45)	(-3.28)	(-1.11)	(-2.50)
Control variables	Yes	Yes	Yes	Yes
N	14,189	14,189	13,890	14,814
Average R-squared	0.0345	0.0204	0.0590	0.0227

Table 10: The Effect of Textual Changes on Option Volatility Smirk Changes by Short-Sale Constraints

Note: This table examines whether the effect of textual changes on option volatility smirk changes varies by short-sale constraints. Short selling costs are proxied with institutional ownership. Firms are sorted into the high-cost and the low-cost groups by the median value of institutional ownership measured at the end of the previous fiscal year. Fama-MacBeth regressions are estimated separately for each group. The dependent variable is the option volatility smirk change. The key explanatory variable is the cosine similarity measure. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. var.:	High cost	Low cost
Volatility smirk change	1	2
Cosine	-7.792**	-1.978**
	(-2.85)	(-2.18)
Control variables	Yes	Yes
N	8,504	8,503
Average R-squared	0.0292	0.0367

Table 11: Portfolio Returns of Optionable Stocks vs. Non-Optionable Stocks

Note: This table reports value-weighted portfolio returns based on textual changes for stocks with options and stocks without options. The dependent variable is monthly returns multiplied by 100. Using propensity score matching, we pair stocks with and without options based on their industry, size, trading volume, and return volatility (Flippou et al., 2022). Following Cohen et al. (2020), we calculate monthly quintiles based on the prior month's distribution of similarity scores. Stocks enter the portfolio one month after their 10-K release and remain for nine months, as 10-Ks are released on a yearly basis. Q1 includes firms with the lowest similarity between current and previous year's 10-Ks, while Q5 includes those with the highest similarity. Q5–Q1 denotes a long-short portfolio that goes long Q5 and short Q1 monthly. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Panel A: Stocks With Options	Q1	Q2	Q3	Q4	Q5	Q5–Q1
Excess return	0.234	0.649*	0.637**	0.507	0.729**	0.495**
	(0.73)	(1.87)	(2.02)	(1.35)	(2.30)	(2.32)
Three-factor alpha	-0.357**	0.008	0.045	-0.074	0.106	0.463**
	(-2.29)	(0.04)	(0.30)	(-0.36)	(0.64)	(2.09)
Four-factor alpha	-0.370**	0.064	0.050	-0.087	0.205	0.575**
	(-2.36)	(0.36)	(0.35)	(-0.46)	(1.22)	(2.38)
Five-factor alpha	-0.363**	0.125	0.062	-0.089	0.198	0.561**
	(-2.39)	(0.72)	(0.43)	(-0.50)	(1.17)	(2.37)
Panel B: Stocks Without Options						
Excess return	0.414	1.317***	0.416	0.476	0.570*	0.156
	(0.95)	(3.36)	(0.98)	(1.25)	(1.83)	(0.35)
Three-factor alpha	-0.169	0.652*	-0.248	-0.159	-0.036	0.133
	(-0.42)	(1.91)	(-0.58)	(-0.63)	(-0.16)	(0.30)
Four-factor alpha	-0.198	0.622*	-0.166	-0.192	0.046	0.245
	(-0.51)	(1.75)	(-0.34)	(-0.72)	(0.18)	(0.56)
Five-factor alpha	-0.178	0.625*	-0.196	-0.169	0.042	0.219
	(-0.46)	(1.71)	(-0.43)	(-0.62)	(0.17)	(0.52)

Table 12: Portfolio Returns by Firm Size

Note: This table reports value-weighted portfolio returns partitioned by firm size for stocks with traded options. Panels A and B report subsamples divided by firm market value, measured as fiscal-year-end market capitalization. Each firm-year is classified as either large- or small-cap based on the sample median, and only firms that remain consistently within the same size category are retained. Panel C presents results for firms included in the S&P 500 during the portfolio construction sample period, while Panel D covers firms that have never been part of the S&P 500 index. Firms are sorted into quintiles based on textual similarity between the current and prior years 10K filings, with Q1 representing the lowest similarity and Q5 the highest. Q5–Q1 denotes a long-short portfolio that goes long Q5 and short Q1 monthly. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Panel A: Large Cap	Q1	Q2	Q3	Q4	Q5	Q5–Q1
Excess return	0.556	0.471	0.684**	0.364	0.746**	0.190
	(1.64)	(1.34)	(2.37)	(0.97)	(2.17)	(1.09)
Three-factor alpha	-0.098	-0.201	0.085	-0.286*	0.092	0.190
	(-0.70)	(-1.17)	(0.60)	(-1.68)	(0.53)	(1.11)
Four-factor alpha	-0.070	-0.226	0.109	-0.260	0.126	0.196
	(-0.51)	(-1.26)	(0.85)	(-1.63)	(0.74)	(1.11)
Five-factor alpha	-0.057	-0.220	0.115	-0.254*	0.142	0.199
	(-0.44)	(-1.18)	(0.91)	(-1.74)	(0.86)	(1.11)
Panel B: Small Cap						
Excess return	-0.118	0.084	0.097	0.540	0.407	0.525**
	(-0.30)	(0.22)	(0.27)	(1.40)	(1.24)	(2.33)
Three-factor alpha	-0.907***	-0.851***	-0.662***	-0.269	-0.292	0.615***
	(-4.82)	(-4.28)	(-3.94)	(-1.35)	(-1.60)	(2.87)
Four-factor alpha	-0.823***	-0.738***	-0.610***	-0.259	-0.214	0.609***
	(-4.91)	(-3.89)	(-3.62)	(-1.24)	(-1.14)	(2.75)
Five-factor alpha	-0.830***	-0.733***	-0.627***	-0.278	-0.211	0.619***
	(-4.86)	(-4.05)	(-3.62)	(-1.31)	(-1.15)	(2.81)
Panel C: S&P500						
Excess return	0.451	0.494	0.592**	0.555	0.664*	0.213
	(1.36)	(1.30)	(2.11)	(1.54)	(1.91)	(1.06)
Three-factor alpha	-0.206	-0.226	0.022	-0.046	0.017	0.223
	(-1.56)	(-1.15)	(0.12)	(-0.20)	(0.10)	(1.09)
Four-factor alpha	-0.182	-0.255	0.118	0.014	0.071	0.253
	(-1.39)	(-1.25)	(0.67)	(0.06)	(0.43)	(1.26)
Five-factor alpha	-0.175	-0.250	0.134	0.015	0.086	0.260
	(-1.36)	(-1.20)	(0.76)	(0.07)	(0.53)	(1.30)
Panel D: Non-S&P500						
Excess return	0.339	0.444	0.576	0.514	0.635*	0.296
	(0.80)	(1.15)	(1.45)	(1.43)	(1.83)	(1.43)
Three-factor alpha	-0.502***	-0.329**	-0.268**	-0.292**	-0.143	0.359**
	(-3.36)	(-1.97)	(-2.10)	(-2.31)	(-1.34)	(2.01)
Four-factor alpha	-0.462***	-0.253	-0.234*	-0.227*	-0.100	0.362**
	(-3.46)	(-1.64)	(-1.80)	(-1.83)	(-0.92)	(2.18)
Five-factor alpha	-0.452***	-0.215	-0.246*	-0.233*	-0.150	0.302**
	(-3.54)	(-1.39)	(-1.87)	(-1.93)	(-1.42)	(2.02)

Table 13: Portfolio Analyses Based on Textual Changes and Option Volatility Smirk Changes

Note: This table reports value-weighted portfolio returns based on both textual changes and option smirk volatility changes. Panels A–E present 1-, 3-, 6-, 9-, and 12-month returns, respectively. The dependent variables are cumulative stock returns over different time horizons following the 10-K release. Following Cohen et al. (2020), stocks are sorted into 2-by-5 portfolios based on the distribution of similarity scores and the volatility smirk changes. Smirk changes below the median value are classified as Low-smirk change, while those above the median value are classified as High-smirk change. Low–High represents a long-short strategy that goes long the Low-smirk change portfolio and short the High-smirk change portfolio. Stocks enter the portfolio one month after their 10-K release and remain for nine months, as 10-Ks are released annually. Q1 includes firms with the lowest similarity between current and previous year's 10-Ks, while Q5 includes those with the highest similarity. Q5–Q1 denotes a long-short portfolio that goes long Q5 and short Q1. Cross Q5–Q1 in Low-smirk change group represents a portfolio that goes long Q5 of the High-smirk change group and short Q1 of the Low-smirk change group. Cross Q5–Q1 in High-smirk change group represents a similar long-short strategy with reversed smirk change groups. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Panel A: 1-Month Return

Excess return	Q1	Q2	Q3	Q4	Q5	Q5–Q1	Cross Q5–Q1
Low-smirk change	0.552	0.634*	0.723**	0.422	0.827**	0.276	0.165
	(1.58)	(1.94)	(2.37)	(1.12)	(2.42)	(1.21)	(0.73)
High-smirk change	0.303	0.360	0.728**	0.504	0.716**	0.414*	0.525**
	(0.87)	(1.10)	(2.14)	(1.39)	(2.30)	(1.82)	(2.40)
Low–High	0.249	0.274	-0.005	-0.083	0.111		
	(1.27)	(1.23)	(-0.02)	(-0.44)	(0.44)		
Three-factor alpha							
Low-smirk change	-0.080	0.048	0.135	-0.166	0.199	0.279	0.108
	(-0.51)	(0.23)	(0.80)	(-0.83)	(0.96)	(1.28)	(0.50)
High-smirk change	-0.380**	-0.254*	0.087	-0.111	0.028	0.408*	0.579**
	(-2.30)	(-1.69)	(0.53)	(-0.55)	(0.17)	(1.76)	(2.43)
Low–High	0.300	0.302	0.048	-0.055	0.171		
	(1.48)	(1.29)	(0.20)	(-0.29)	(0.65)		
Four-factor alpha							
Low-smirk change	-0.036	0.084	0.178	-0.134	0.301	0.337	0.071
	(-0.22)	(0.42)	(1.01)	(-0.70)	(1.49)	(1.57)	(0.33)
High-smirk change	-0.393**	-0.250	0.060	-0.028	0.036	0.429*	0.695***
	(-2.30)	(-1.59)	(0.41)	(-0.15)	(0.21)	(1.76)	(2.63)
Low–High	0.358	0.334	0.119	-0.106	0.266		
	(1.59)	(1.32)	(0.50)	(-0.59)	(1.02)		
Five-factor alpha							
Low-smirk change	-0.100	0.099	0.151	-0.134	0.284	0.383*	0.138
	(-0.65)	(0.49)	(0.88)	(-0.74)	(1.36)	(1.85)	(0.65)
High-smirk change	-0.373**	-0.221	0.063	-0.047	0.038	0.412*	0.657**
	(-2.30)	(-1.41)	(0.44)	(-0.27)	(0.23)	(1.74)	(2.46)
Low–High	0.274	0.321	0.088	-0.087	0.245		
	(1.27)	(1.24)	(0.38)	(-0.48)	(0.92)		

Panel B: 3-Month Return

Excess return	Q1	Q2	Q3	Q4	Q5	Q5–Q1	Cross Q5–Q1
Low-smirk change	1.636*	1.833**	2.207***	1.304	2.466**	0.830	0.506
	(1.74)	(2.04)	(2.64)	(1.25)	(2.56)	(1.31)	(0.83)
High-smirk change	0.870	1.086	2.157**	1.546	2.142**	1.272**	1.596***
	(0.91)	(1.19)	(2.29)	(1.54)	(2.52)	(2.03)	(2.69)
Low–High	0.767	0.746	0.049	-0.242	0.324		
	(1.49)	(1.23)	(0.07)	(-0.49)	(0.45)		
Three-factor alpha							
Low-smirk change	-0.341	0.180	0.408	-0.442	0.528	0.869	0.498
	(-0.78)	(0.32)	(0.95)	(-0.85)	(0.91)	(1.48)	(0.80)
High-smirk change	-1.088**	-0.752*	0.316	-0.254	0.157	1.245*	1.616***
	(-2.53)	(-1.86)	(0.76)	(-0.49)	(0.35)	(1.96)	(2.60)
Low–High	0.747	0.932	0.092	-0.188	0.371		
	(1.50)	(1.47)	(0.15)	(-0.41)	(0.51)		
Four-factor alpha							
Low-smirk change	-0.272	0.227	0.627	-0.279	0.709	0.981*	0.537
	(-0.66)	(0.39)	(1.38)	(-0.48)	(1.31)	(1.79)	(0.88)
High-smirk change	-1.070**	-0.539	0.178	-0.060	0.265	1.335**	1.779***
	(-2.43)	(-1.26)	(0.47)	(-0.12)	(0.58)	(1.99)	(2.89)
Low–High	0.798	0.766	0.449	-0.219	0.444		
	(1.53)	(1.11)	(0.72)	(-0.49)	(0.61)		
Five-factor alpha							
Low-smirk change	-0.385	0.320	0.552	-0.197	0.794	1.179**	0.631
	(-0.93)	(0.57)	(1.21)	(-0.38)	(1.53)	(2.38)	(1.07)
High-smirk change	-1.018**	-0.467	0.174	0.016	0.246	1.264*	1.812***
	(-2.43)	(-1.08)	(0.50)	(0.04)	(0.55)	(1.95)	(2.89)
Low–High	0.633	0.786	0.378	-0.213	0.548		
	(1.19)	(1.14)	(0.64)	(-0.47)	(0.79)		

Panel C: 6-Month Return

Excess return	Q1	Q2	Q3	Q4	Q5	Q5–Q1	Cross Q5–Q1
Low-smirk change	3.316**	3.652**	4.420***	2.720	4.962***	1.646	0.981
	(1.99)	(2.27)	(3.09)	(1.47)	(2.85)	(1.40)	(0.92)
High-smirk change	1.823	2.270	4.349**	3.158*	4.297***	2.474**	3.139***
	(1.06)	(1.38)	(2.58)	(1.76)	(2.90)	(2.29)	(3.07)
Low–High	1.493*	1.381	0.071	-0.439	0.665		
	(1.65)	(1.29)	(0.06)	(-0.54)	(0.51)		
Three-factor alpha							
Low-smirk change	-0.783	0.506	0.970	-0.774	1.061	1.844*	1.237
	(-1.14)	(0.49)	(1.40)	(-0.95)	(1.03)	(1.73)	(1.14)
High-smirk change	-2.312***	-1.768**	0.720	-0.500	0.454	2.766**	3.373***
	(-3.51)	(-2.46)	(1.06)	(-0.64)	(0.58)	(2.51)	(3.28)
Low–High	1.529*	2.274*	0.250	-0.274	0.607		
	(1.91)	(1.97)	(0.24)	(-0.38)	(0.43)		
Four-factor alpha							
Low-smirk change	-0.685	0.316	1.351*	-0.348	1.340	2.024**	1.432
	(-1.09)	(0.30)	(1.77)	(-0.38)	(1.38)	(2.01)	(1.32)
High-smirk change	-2.160***	-1.136	0.522	-0.294	0.747	2.907**	3.499***
	(-3.09)	(-1.54)	(0.84)	(-0.36)	(0.95)	(2.51)	(3.49)
Low–High	1.475*	1.452	0.829	-0.054	0.592		
	(1.77)	(1.22)	(0.76)	(-0.08)	(0.42)		
Five-factor alpha							
Low-smirk change	-0.791	0.687	1.079	-0.100	1.801**	2.592***	1.502
	(-1.27)	(0.63)	(1.40)	(-0.12)	(2.23)	(3.04)	(1.45)
High-smirk change	-1.975***	-0.881	0.375	0.002	0.712	2.686**	3.776***
	(-3.05)	(-1.17)	(0.66)	(0.00)	(0.95)	(2.50)	(3.90)
Low–High	1.184	1.568	0.704	-0.102	1.089		
	(1.43)	(1.29)	(0.67)	(-0.15)	(0.90)		

Panel D: 9-Month Return

Excess return	Q1	Q2	Q3	Q4	Q5	Q5–Q1	Cross Q5–Q1
Low-smirk change	5.032**	5.578**	6.543***	4.301*	7.354***	2.322	1.415
	(2.25)	(2.55)	(3.58)	(1.71)	(3.15)	(1.50)	(1.05)
High-smirk change	2.693	3.337	6.592***	4.789*	6.447***	3.754***	4.661***
	(1.16)	(1.51)	(2.90)	(1.97)	(3.24)	(2.68)	(3.66)
Low–High	2.339*	2.241	-0.049	-0.487	0.907		
	(1.95)	(1.55)	(-0.03)	(-0.47)	(0.53)		
Three-factor alpha							
Low-smirk change	-1.368	0.680	1.618*	-1.147	1.325	2.692*	1.826
	(-1.64)	(0.49)	(1.82)	(-1.05)	(0.91)	(1.73)	(1.26)
High-smirk change	-3.692***	-2.973***	1.429	-0.969	0.458	4.150***	5.016***
	(-4.53)	(-2.84)	(1.63)	(-0.98)	(0.45)	(2.89)	(3.82)
Low–High	2.324**	3.653**	0.189	-0.178	0.866		
	(2.34)	(2.30)	(0.14)	(-0.20)	(0.43)		
Four-factor alpha							
Low-smirk change	-1.096	0.405	2.248**	-0.364	1.732	2.829*	2.098
	(-1.43)	(0.28)	(2.33)	(-0.32)	(1.26)	(1.84)	(1.51)
High-smirk change	-3.363***	-1.800*	1.106	-0.843	1.002	4.365***	5.095***
	(-4.14)	(-1.79)	(1.41)	(-0.81)	(0.98)	(2.91)	(4.01)
Low–High	2.266**	2.204	1.141	0.479	0.731		
	(2.25)	(1.41)	(0.86)	(0.65)	(0.36)		
Five-factor alpha							
Low-smirk change	-1.224	0.941	1.731*	0.114	2.622**	3.847***	2.044
	(-1.57)	(0.67)	(1.87)	(0.12)	(2.52)	(3.14)	(1.55)
High-smirk change	-3.026***	-1.474	0.814	-0.335	0.819	3.846***	5.649***
	(-4.37)	(-1.47)	(1.18)	(-0.38)	(0.87)	(2.96)	(4.90)
Low–High	1.802*	2.415	0.917	0.449	1.803		
	(1.92)	(1.54)	(0.72)	(0.57)	(1.12)		

Panel E: 12-Month Return

Excess return	Q1	Q2	Q3	Q4	Q5	Q5–Q1	Cross Q5–Q1
Low-smirk change	6.733**	7.532***	8.670***	5.676*	9.627***	2.894	1.944
	(2.48)	(2.79)	(4.09)	(1.84)	(3.37)	(1.58)	(1.21)
High-smirk change	3.498	4.307	8.845***	6.316**	8.677***	5.179***	6.129***
	(1.22)	(1.61)	(3.16)	(2.11)	(3.53)	(3.08)	(4.19)
Low–High	3.235**	3.225*	-0.175	-0.640	0.950		
	(2.24)	(1.87)	(-0.08)	(-0.52)	(0.47)		
Three-factor alpha							
Low-smirk change	-1.807*	0.754	2.633**	-1.554	1.404	3.211	1.986
	(-1.84)	(0.46)	(2.36)	(-1.16)	(0.79)	(1.64)	(1.18)
High-smirk change	-4.828***	-4.043***	2.175*	-1.382	0.179	5.008***	6.232***
	(-4.87)	(-2.98)	(1.93)	(-1.16)	(0.15)	(3.05)	(4.02)
Low–High	3.022**	4.797**	0.457	-0.172	1.224		
	(2.45)	(2.51)	(0.28)	(-0.16)	(0.51)		
Four-factor alpha							
Low-smirk change	-1.603*	0.617	3.262***	-0.387	2.118	3.721*	2.519
	(-1.68)	(0.37)	(2.69)	(-0.28)	(1.25)	(1.93)	(1.60)
High-smirk change	-4.280***	-2.147*	1.795*	-1.247	0.916	5.196***	6.398***
	(-4.26)	(-1.69)	(1.73)	(-0.98)	(0.81)	(2.98)	(4.15)
Low–High	2.677**	2.765	1.467	0.860	1.202		
	(2.18)	(1.46)	(0.90)	(1.01)	(0.49)		
Five-factor alpha							
Low-smirk change	-1.847*	1.577	2.250**	0.266	3.669***	5.516***	2.256
	(-1.93)	(0.93)	(2.10)	(0.22)	(2.87)	(3.74)	(1.51)
High-smirk change	-3.707***	-1.683	1.334	-0.489	0.409	4.116***	7.376***
	(-4.51)	(-1.30)	(1.57)	(-0.43)	(0.40)	(2.95)	(5.12)
Low–High	1.860*	3.260*	0.916	0.755	3.260*		
	(1.77)	(1.66)	(0.62)	(0.79)	(1.72)		

Table 14: The Predictability of Textual Changes on Future Stock Returns

Note: This table presents the Fama-MacBeth regression results of return predictability based on cosine and option smirk volatility changes. The dependent variables are the cumulative stock returns over different time horizons after the 10-K release. Columns (1) and (2) include 1-month stock returns, columns (3) and (4) include 3-month returns, columns (5) and (6) include 6-month returns, columns (7) and (8) include 9-month returns, and columns (9) and (10) include 12-month returns. The key explanatory variables are the cosine similarity measure, High-smirk change, and their interaction. High-smirk change is an indicator determined by the median value of option volatility smirk changes. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var: Future return	1-mon return		3-mon return		6-mon return		9-mon return		12-mon return	
	1	2	3	4	5	6	7	8	9	10
Cosine	0.017	-0.019	0.036	-0.015	0.120***	0.006	0.151***	-0.011	0.185***	-0.021
	(1.28)	(-0.48)	(1.41)	(-0.35)	(3.58)	(0.08)	(5.27)	(-0.16)	(3.46)	(-0.21)
High-smirk change		-0.068		-0.095		-0.213**		-0.314***		-0.389**
		(-1.32)		(-1.43)		(-2.32)		(-2.96)		(-2.48)
Cosine × High-smirk change		0.070		0.094		0.215**		0.319***		0.398**
		(1.31)		(1.43)		(2.31)		(3.03)		(2.50)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	27,953	27,953	27,759	27,759	27,346	27,346	26,943	26,943	26,573	26,573
Average R-squared	0.0405	0.0439	0.0490	0.0528	0.0631	0.0667	0.0681	0.0718	0.0772	0.0796

Appendix

Table A1: Variable Definitions

Variable name	Definition	Source
Abnormal Idio. volatility	Idiosyncratic volatility prior to information events in excess of its normal levels (Yang et al., 2020).	Self calculation
Analyst coverage	The number of analysts covering the firm.	I/B/E/S
Book-to-market	Total book value of assets scaled by market capitalization plus total liabilities.	Compustat
Capex	Capital expenditure scaled by total assets.	Compustat
Cash	Cash holding scaled by total assets.	Compustat
Cosine	Cosine similarity of the current and previous fiscal years' 10-Ks.	SEC Analytic Suite
Cosine (by items)	Cosine similarity of the same item in the current and previous fiscal years' 10-Ks.	Self calculation
File size	The file size of the 10-Ks.	SEC Analytic Suite
Market cap	Number of shares outstanding multiplied by stock price.	Compustat
Negative sentiment	The ratio of negative words to total words in 10-K changes, (Cohen et al., 2020).	Self calculation
Open interest	Delta-adjusted open interest is calculated by multiplying an option's daily open interest by its delta, then dividing by the number of outstanding shares of the underlying stock (Lakonishok et al., 2007).	OptionMetrics
Return	Stock returns in the previous 12 months.	Compustat
Reverse smirk	Difference between the implied volatilities of out-of-the-money call options and implied volatilities of at-the-money put options. Reverse smirk change is computed in the same way as volatility smirk change.	OptionMetrics
ROA	Net income scaled by total assets.	Compustat
SUE	The difference between announced earnings and the latest consensus earnings forecast before the announcement, normalized by the standard deviation of the consensus earnings forecast.	I/B/E/S
Tangibility	Total tangible assets scaled by total assets.	Compustat
Tone optimism	Number of Loughran-McDonald positive words minus negative words scaled by the total words of 10-Ks.	SEC Analytic Suite
Volatility smirk	Difference between the implied volatilities of out-of-the-money put options and implied volatilities of at-the-money call options.	OptionMetrics
Volatility smirk change	Volatility smirk right after the 10-K release date minus the volatility smirk at least 8 weeks before the 10-K release date.	OptionMetrics

Table A2: Baseline Results Based on an Alternative Benchmark

Note: This table presents the Fama-MacBeth regression results using an alternative baseline for the volatility smirk measure. The dependent variable is the difference between the post-release smirk and the smirk on the day immediately before the 10-K release date. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

	1	2	3
Cosine	-1.402*	-1.465**	-1.678**
	(-1.91)	(-2.14)	(-2.13)
Tone optimism		3.197	0.853
		(1.04)	(0.19)
Log file size		-0.051	-0.041
		(-0.64)	(-0.69)
Control variables			Yes
N	28,036	28,036	28,036
Average R-squared	0.0006	0.0018	0.0150

Table A3: Baseline Results Based on Alternative Similarity Measures

Note: This table reports the Fama-MacBeth regression results on the effect of textual changes on option volatility smirk changes based on alternative measures of annual report similarity: Jaccard, modified Jaccard, and minimum edit distance. The dependent variable is the option volatility smirk change. The key explanatory variable is one of the three alternative measures replacing the cosine similarity measure. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var.: Volatility smirk change	1	2	3
Jaccard	-0.712***		
	(-4.06)		
Modified Jaccard		-1.224***	
		(-2.83)	
Min edit distance			0.597***
			(3.92)
Control variables	Yes	Yes	Yes
N	28,704	28,704	28,704
Average R-squared	0.0158	0.0156	0.0159

Table A4: Baseline Results Based on Fixed Effects Regressions

Note: This table reports the pooled regression results on the effect of textual changes on option volatility smirk changes. The dependent variable is the option volatility smirk change. The key explanatory variable is the cosine similarity measure on 10-K textual changes. Control variables are the same as in Table 3 but omitted from reporting for brevity. Industry and year fixed effects are included. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. var.: Volatility smirk change	1	2
Cosine	-2.241**	-2.376**
	(-2.24)	(-2.38)
Control variables	No	Yes
Industry and year FEs	Yes	Yes
N	28,696	28,696
Average R-squared	0.0291	0.0300

Table A5: The Effect of Textual Changes on Option Volatility Smirk Levels

Note: This table reports the Fama-MacBeth regression results on the effect of textual changes on option volatility smirk levels. The dependent variable is the level of option volatility smirk after the 10-K release. The key explanatory variable is the cosine similarity measure on 10-K textual changes. Control variables include log file size, tone optimism, log market cap, book-to-market, one-year stock return, ROA, cash, capex, and tangibility. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. var.: Volatility smirk	1	2
Cosine	-2.333***	-2.223***
	(-3.31)	(-3.15)
Control variables	No	Yes
N	32,473	32,473
Average R-squared	0.000653	0.019180

Table A6: The Effect of Textual Changes on Reverse Option Volatility Smirk Changes

Note: This table reports the Fama-MacBeth regression results of textual changes on reverse option volatility smirk changes. The dependent variable is the change of the difference between implied volatilities of out-of-the-money call options and implied volatilities of at-the-money put options. The key explanatory variable is the cosine similarity measure on 10-K textual changes. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var: Reverse volatility smirk change	1	2
Cosine	1.409*	1.221*
	(1.75)	(1.95)
Tone optimism	8.824	5.990
	(1.02)	(0.78)
Log file size	0.022	-0.141
	(0.52)	(-0.86)
Other control variables		Yes
N	30,802	30,802
Average R-squared	0.0031	0.0138

Table A7: The Effect of 10-K Textual Changes on Option Volatility Smirk Changes: An Extended Window

Note: This table reports the Fama-MacBeth regression results on the effect of textual changes on option volatility smirk changes during weeks 9 to 13 following the 10-K release. The dependent variable is the weekly option volatility smirk change measured as the difference between volatility smirk measured on the day nearest to the 10-K release date and the baseline smirk measured at least eight weeks before the 10-K release date. The key explanatory variable is the cosine similarity measure on 10-K textual changes. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

	week[9]	week[10]	week[11]	week[12]	week[13]
Cosine	-1.988	-1.975	-1.441	-0.995	-1.048
	(-1.58)	(-0.79)	(-0.78)	(-0.81)	(-1.21)
Control variables	Yes	Yes	Yes	Yes	Yes
N	24,882	24,730	24,488	24,368	24,268
Average R-squared	0.0141	0.0197	0.0201	0.0157	0.0191

Table A8: Portfolio Analyses Based on SUE and Option Volatility Smirk Changes

Note: This table reports value-weighted portfolio returns based on standardized unexpected earnings (SUE) and option smirk volatility changes. The dependent variable is monthly returns multiplied by 100. We sort stocks into quintiles based on their SUE, which is the difference between announced earnings and the latest consensus earnings forecast before the announcement scaled by the standard deviation of the consensus earnings forecast. Then we divide them by the volatility smirk changes. Smirk changes lower than the median value are classified as Low-smirk change. Stocks with smirk changes higher than the median value are classified as High-smirk change. Stocks enter the portfolio one month after their earnings announcements and remain for three months, as earnings are released on a quarterly basis. Q1 includes firms with the lowest SUE, while Q5 includes those with the highest SUE. Q5–Q1 denotes a long-short portfolio that goes long Q5 and short Q1 monthly. Cross Q5–Q1 in Low-smirk change group represents a portfolio that goes long Q5 of the High-smirk change group and short Q1 of the Low-smirk change group monthly. Cross Q5–Q1 in High-smirk change group represents a similar long-short strategy with reversed smirk change groups. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Panel A: Excess Return	Q1	Q2	Q3	Q4	Q5	Q5–Q1	Cross Q5–Q1
Low-smirk change	0.687*	0.757**	0.447	0.596**	0.763**	0.076	0.064
	(1.69)	(2.49)	(1.51)	(2.01)	(2.35)	(0.31)	(0.27)
High-smirk change	0.333	0.531*	0.653**	0.594*	0.752**	0.418*	0.430*
	(0.89)	(1.70)	(2.15)	(1.97)	(2.29)	(1.95)	(1.88)
Panel B: Three-Factor Alpha							
Low-smirk change	-0.108	0.092	-0.186*	-0.044	0.123	0.231	0.211
	(-0.52)	(0.63)	(-1.74)	(-0.41)	(1.02)	(0.94)	(0.87)
High-smirk change	-0.438***	-0.195*	0.034	-0.018	0.103	0.541***	0.561***
	(-2.85)	(-1.74)	(0.28)	(-0.18)	(0.99)	(2.81)	(2.77)
Panel C: Four-Factor Alpha							
Low-smirk change	-0.010	0.191	-0.159	-0.033	0.099	0.109	0.103
	(-0.04)	(1.34)	(-1.56)	(-0.31)	(0.90)	(0.42)	(0.38)
High-smirk change	-0.277*	-0.092	0.024	-0.021	0.093	0.370*	0.376**
	(-1.89)	(-0.88)	(0.19)	(-0.22)	(0.81)	(1.84)	(2.01)
Panel D: Five-Factor Alpha							
Low-smirk change	-0.005	0.179	-0.163	-0.047	0.099	0.103	0.087
	(-0.02)	(1.27)	(-1.57)	(-0.44)	(0.94)	(0.40)	(0.33)
High-smirk change	-0.288**	-0.088	0.006	-0.024	0.083	0.371*	0.387**
	(-2.00)	(-0.84)	(0.05)	(-0.25)	(0.73)	(1.86)	(2.08)

Table A9: The Predictability of Textual Changes on Future Returns by Options Maturity Groups

Note: This table presents the Fama-MacBeth regression results of return predictability based on cosine and option smirk volatility changes. Panels A, B, and C are based on options with maturities of 8-60 days (Panel A), 61-180 days (Panel B), and 181-360 days (Panel C). The dependent variables are the cumulative stock returns over different time horizons after the 10-K release. Columns (1) and (2) include 1-month stock returns, columns (3) and (4) include 3-month returns, columns (5) and (6) include 6-month returns, columns (7) and (8) include 9-month returns, and columns (9) and (10) include 12-month returns. The key explanatory variables are the cosine similarity measure, High-smirk change, and their interaction. High-smirk change is an indicator determined by the median value of option volatility smirk changes. Control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Panel A: 8-60 Days	1-mon return		3-mon return		6-mon return		9-mon return		12-mon return	
	1	2	3	4	5	6	7	8	9	10
Cosine	0.011	-0.013	0.033	0.008	0.117***	0.046	0.128***	0.039	0.174***	0.057
	(0.86)	(-0.42)	(1.13)	(0.21)	(2.98)	(0.80)	(3.55)	(0.97)	(3.17)	(1.13)
High-smirk change		-0.056		-0.045		-0.144*		-0.169*		-0.232*
		(-1.25)		(-1.03)		(-1.94)		(-1.85)		(-1.81)
Cosine × High-smirk change		0.056		0.047		0.145*		0.177*		0.240*
		(1.26)		(1.09)		(1.91)		(1.82)		(1.77)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	29,176	29,176	28,955	28,955	28,534	28,534	28,106	28,106	27,732	27,732
Average R-squared	0.0385	0.0412	0.0456	0.0471	0.0599	0.0612	0.0668	0.0683	0.0757	0.0776
Panel B: 61-180 Days										
Cosine	0.011	-0.030	0.029	-0.089	0.115***	-0.014	0.129***	-0.002	0.174***	-0.011
	(0.78)	(-1.11)	(1.08)	(-1.63)	(2.92)	(-0.18)	(3.54)	(-0.04)	(3.08)	(-0.17)
High-smirk change		-0.081**		-0.220**		-0.242**		-0.260***		-0.392**
		(-2.22)		(-2.69)		(-2.21)		(-3.12)		(-2.73)
Cosine × High-smirk change		0.083**		0.223**		0.250**		0.264***		0.403**
		(2.23)		(2.70)		(2.27)		(3.18)		(2.82)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	29,063	29,063	28,853	28,853	28,429	28,429	28,002	28,002	27,629	27,629
Average R-squared	0.0385	0.0412	0.0455	0.0483	0.0607	0.0632	0.0673	0.0696	0.0759	0.0782
Panel C: 181-360 Days										
Cosine	0.001	-0.023	0.012	-0.000	0.090**	0.008	0.116***	0.024	0.149**	0.020
	(0.05)	(-0.74)	(0.36)	(-0.00)	(2.40)	(0.14)	(3.95)	(0.48)	(2.28)	(0.31)
High-smirk change		-0.074		-0.023		-0.158		-0.207**		-0.261**
		(-1.26)		(-0.51)		(-1.50)		(-2.14)		(-2.37)
Cosine × High-smirk change		0.077		0.024		0.167		0.219**		0.277**
		(1.30)		(0.51)		(1.56)		(2.24)		(2.49)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	27,599	27,599	27,420	27,420	27,044	27,044	26,656	26,656	26,299	26,299
Average R-squared	0.0410	0.0457	0.0474	0.0503	0.0619	0.0640	0.0686	0.0704	0.0769	0.0783

Table A10: Predictability of Textual Changes Before and After 2010

Note: The table reports results of Fama-MacBeth regressions of stock returns on both cosine similarity and option smirk change. The dependent variable is the 6-month cumulative return, calculated using the method from Table 14. High-smirk change is an indicator determined by the median value of option volatility smirk changes. Columns (1) and (2) are based on the period before 2010, when Cohen et al. (2020) was first published on SSRN. Columns (3) and (4) are based on the period after 2010. We exclude the publication year of 2010 to avoid potential compounding effects. Other control variables are the same as in Table 3 but omitted from reporting for brevity. Variable definitions are provided in Appendix Table A1. T-statistics are reported in parentheses. Statistical significance of 1%, 5%, and 10% is indicated by ***, **, and *, respectively.

Dep. Var: Future return	Pre-2010		Post-2010	
Cosine	0.152***	0.033	0.016	-0.101
	(5.66)	(0.43)	(0.29)	(-1.53)
High-smirk change		-0.212*		-0.209**
		(-1.97)		(-2.42)
Cosine × High-smirk change		0.211*		0.204**
		(1.94)		(2.46)
Control variables	Yes	Yes	Yes	Yes
N	15,878	15,878	11,609	11,609
Average R-squared	0.0649	0.0677	0.0566	0.0588