The Cross-Section of Extrapolative Belief and the High-Volume Premium

Huaixin Wang *

This draft: November 10, 2022

Abstract

Stocks with abnormally high volumes are associated with high subsequent returns (Gervais et al., 2001). I show that return extrapolation bias contributes to the spike in trading volume for individual stocks. The high-volume premium (HVP) is more pronounced among firms with low extrapolative value, whereas the premium is mitigated among firms with high extrapolative value. The difference in the HVP between low- and high- extrapolative value firms can be predicted by DOX, the market-wide extrapolation level (Cassella and Gulen, 2018). I also provide evidence that the documented heterogeneity of the HVP is not driven by stock visibility or illiquidity. The results indicate that extrapolative expectation is an important contributor to cross-sectional volume-return relations.

JEL classification: G12, G41

Keywords: Extrapolation, Trading volume, Anomaly, Mispricing

*I'm grateful to Jianfeng Yu for his continuous guidance and encouragement. I thank Zhao Jin, Han Xiao (discussant), Hong Zhang, and conference participants at the 2022 China Behavioral and Experimental Economics Forum, 2022 AFR international conference of economics and finance (PhD student session) for valuable comments. I sincerely thank Zhiwei Su for sharing the DOX data. All remaining errors are my responsibility. Author affiliation/contact information: Huaixin Wang is from PBC School of Finance, Tsinghua University, Beijing 100083, China. E-mail Address: wanghx.19@pbcsf.tsinghua.edu.cn.

The Cross-Section of Extrapolative Belief and the High-Volume Premium

Abstract

Stocks with abnormally high volumes are associated with high subsequent returns (Gervais et al., 2001). I show that return extrapolation bias contributes to the spike in trading volume for individual stocks. The high-volume premium (HVP) is more pronounced among firms with low extrapolative value, whereas the premium is mitigated among firms with high extrapolative value. The difference in the HVP between low- and high- extrapolative value firms can be predicted by DOX, the market-wide extrapolation level (Cassella and Gulen, 2018). I also provide evidence that the documented heterogeneity of the HVP is not driven by stock visibility or illiquidity. The results indicate that extrapolative expectation is an important contributor to cross-sectional volume-return relations.

JEL classification: G12, G41

Keywords: Extrapolation, Trading volume, Anomaly, Mispricing

1 Introduction

Trading volume is one of the fundamental building blocks for analyzing market interactions. The variation of trading volume reflects investors' preferences, expectations, and the arrival of information. The seminal work of Gervais, Kaniel, and Mingelgrin (2001) documents a striking pattern of the intertemporal volume-return relation: stock with a substantial volume shock receives a significantly subsequent excess market-adjusted return. Such "high-volume premium" (HVP hereafter) is found to be an international phenomenon (Kaniel, Ozoguz, and Starks, 2012), has strong predictive power for future economic fundamentals (Wang, 2021), and is associated with real effects (Israeli, Kaniel, and Sridharan, 2021).

This paper extends the literature by providing evidence that a form of expectation bias, return extrapolation, is a influencing factor of HVP. Using a simple measure implied by survey evidence and theoretical models (Greenwood and Shleifer, 2014; Barberis, Greenwood, Jin, and Shleifer, 2015, 2018; Da, Huang, and Jin, 2021), I show that over-extrapolation bias contributes to the surge in trading volume of individual stocks, and the relationship between abnormal trading volume and future returns is conditional on the valuation from extrapolative belief. Specifically, the trading volume becomes abnormally high for both low- and high-extrapolative value stocks. The HVP is robustly strong among stocks with low extrapolative value, whereas the premium is substantially weakened among stocks with high extrapolative value. The difference in HVP between low- and high- extrapolative value stocks ranges from 0.51% to 0.60% per month, depending on specifications. The results still hold after controlling for a battery of firm characteristics in Fama and MacBeth (1973) regressions.

The analysis is motivated by the behavioral finance literature studying extrapolation and bubbles (DeFusco, Nathanson, and Zwick, 2017; Glaeser and Nathanson, 2017; Barberis et al., 2018; Liao, Peng, and Zhu, 2021). In historical financial bubbles, the large price run-up is usually accompanied by a sharp increase of trading volume and ends with a crash. This branch of research attributes the *coexistence* of asset mispricing and the surge of trading volume to the concept of return extrapolation, i.e., the formation of expected returns positively depends on recent realized returns. Extrapolators tend to purchase stocks that perform well recently, pushing the price further away from the fundamental value. Meanwhile, the trading volume also becomes abnormally high, either due to the interaction between extrapolators and rational investors, speculative trading, or a combination with other behavioral biases. [1]

[1]For example, Barberis et al. (2018) model extrapolator's demand function by introducing a "waver" between a fundamental value signal and an extrapolative growth signal. During bubbles, even a small degree of wavering can generate large portfolio adjustments, causing intense trading. Liao et al. (2021) introduce

Inspired by the theoretical arguments and empirical evidence, I conjecture that an abnormal increase in trading volume at short-horizon contains an extrapolation component, in addition to other factors such as information and market frictions. Hence, the high-volume premium should exhibit a cross-sectional heterogeneity with respect to extrapolative evaluation. The logic is as follows. First, as suggested by prior literature, it is expected that stocks with a high (low) extrapolative expectation to be overpriced (underpriced) and earn lower (higher) returns in the near future. Then, within high extrapolative value stocks, an abnormally high trading volume indicates a relatively higher degree of overpricing and subsequent reversal. Therefore, the HVP should be mitigated. In contrast, for stocks with low extrapolative value, an abnormally high trading volume implies a higher degree of underpricing and subsequent price correction, meaning the HVP would be more evident. Finally, following up on the previous hypothesis, the return reversal due to extrapolation would be more pronounced for stocks whose trading volumes become abnormally high.

I start the analysis by introducing a firm-level extrapolation proxy, the extrapolative value (EXVA). The construction of EXVA is motivated by empirical evidence (Greenwood and Shleifer, 2014; Da et al., 2021) and theoretical models (Barberis et al., 2015, 2018) about extrapolative belief, and is calculated as the average of historical daily returns with geometric-decay weighting. That is, recent return realizations receive higher weights than more distant ones. I follow previous studies and define a stock's abnormal trading volume as the percentile of the month-end volume as compared to a 50-day reference window of its trading history. Accordingly, I skip the last trading day each month when computing EXVA. By single-sorting, I find that EXVA portfolios share similar characteristics such as risk exposure, liquidity, firm size, and turnover ratio. Moreover, high EXVA stocks underperform low EXVA stocks, a long-short strategy by buying stocks in the bottom EXVA decile and selling stocks in the top EXVA decile generates a return of 92 basis points per month ($t = 5.58$).

Importantly, the level of abnormal trading is high for both low-EXVA stocks and high-EXVA stocks, suggesting a U-shape relation between extrapolative value and abnormal volume. Monthly cross-sectional regressions reveal that the square of EXVA strongly predicts the month-end abnormal volume. Based on this finding, I decompose the abnormal trading volume into a quadratic-extrapolation part and the residual part. In Fama-MacBeth regressions, I find the former is negatively priced on the cross-section whereas the latter is positively associated with future returns.

---

the realization utility into the framework of extrapolation. In their model, the "disposition extrapolators" tend to buy an asset with high recent returns, but sell it if the good performance continues. Therefore, investors during bubbles would switch between assets and hence generates the high-volume trading pattern.

Next, I investigate the interaction between extrapolative value and the high-volume premium using double-sorts and regression analysis. Specifically, individual stocks are first sorted into quintiles based on EXVA and then separated into quintile portfolios within each EXVA group, based on abnormal trading volume. The finding is twofold. First, high abnormal volume is associated with high next month return among firms with low EXVA and the return spreads of the two bottom EXVA groups are 0.61% and 0.63%, respectively. In sharp contrast, the abnormal volume return spreads are only 0.19% (t-value 1.43) and 0.09% (t-value 0.61), respectively, for the top two EXVA groups. Second, while low EXVA stocks outperform high EXVA stocks across all abnormal volume quintiles, such outperformance is stronger within high abnormal volume groups. In Fama and MacBeth (1973) regressions, I find that both abnormal volume and EXVA have strong predictive power for future returns, and their interaction term is significantly negative after controlling for firm characteristics such as past performance, book-to-market ratio, size, share turnover, risk measures, and illiquidity.

To test the economic mechanism, I first conduct time-series regressions using a market-wide extrapolation measure, DOX, following Cassella and Gulen (2018). DOX is a time-varying and recursively estimated series measuring the relative weight investors place on recent-versus-distant returns. If the extrapolative belief is indeed an important influencing factor of HVP, then it is expected that DOX positively predicts the difference between HVP among low-EXVA firms and high-EXVA firms (HVPDIFF). The regression shows that a one standard deviation increase in lagged DOX implies a 0.57% increase in HVPDIFF. The estimation is also robust to controlling for contemporaneous asset pricing factors and lagged investor sentiment. Second, I examine and control for the effect of stock visibility. Starting with Gervais et al. (2001), the prevailing explanation for the high-volume premium is Merton's (1987) visibility hypothesis. One would worry that my results are driven by such a "recognition" channel since a stock that recently performs poorly tends to become less visible to investors. I address this issue by adding stock visibility proxies into the main regression analysis. The result suggests that although visibility does play an important role in driving HVP, the effect of EXVA is hardly affected. Finally, I conduct placebo tests using equal-weighted past returns and show that the resulting HVP variation cannot be predicted by DOX, which highlights the necessarity of the decay structure in weighting historical returns.

In robustness tests, I first consider the weighting scheme used in bivariate portfolio analysis. Instead of using value-weighted portfolios, I repeat the analysis based on equal-weighted and gross return-weighted portfolios. I find that the HVP monotonically decreases with respect to EXVA quintiles in both cases and the difference in HVP between low- and high-EXVA stocks remains significant, ranging from 0.89% to 0.96% per month. Then, I use different parameter values to re-calculate EXVA and make use of an alternative extrapolation measure, the price-path convexity (Gulen and Woeppel, 2022). Across all specifications I find similar interactive patterns between extrapolative belief and HVP. To validate the plausibility of using EXVA to proxy for short-term extrapolation, I show that EXVA positively predicts month-end order imbalances. In addition, I also provide evidence that the main empirical findings remain valid after taking into account the popular short-term reversal effect. Finally, I discuss and compare my findings and the heterogeneous volume-return relation studied in Han et al. (2022). Using several mispricing proxies, Han et al. (2022) show that trading volume is positively (negatively) related to future returns among underpriced (overpriced) stocks. I discuss the differences in terms of the trading volume concept, the horizon of mispricing, and the theoretical motivation. I show that the pattern documented in this paper remains unchanged after controlling for additional mispricing measures.

This study relates to three streams of empirical asset pricing literature. First, it contributes to research on the intertemporal relationship between trading volume and stock return. The seminal work of Gervais et al. (2001) documents the HVP and attribute it to Merton's (1987) investor recognition hypothesis. Kaniel et al. (2012) further confirm the recognition hypothesis using cross-country data. Bali, Peng, Shen, and Tang (2014) relate HVP to liquidity shocks and investor underreaction. Regarding other factors, a large body of research builds on the concept of investor overconfidence. Statman, Thorley, and Vorkink (2006) find that share turnover positively predicts future return at the aggregate market. Huang, Heian, and Zhang (2011) show that the high-volume shock can be attributed to retail investor overconfidence and the HVP is weaker in Asian financial markets. On the other side, HVP is also associated with real economy. Wang (2021) finds that the volume premium contains information about economic fundamentals and has strong predictive power for various macroeconomic indicators. Israeli et al. (2021) investigate the real effect of HVP at the firm level and find that a positive volume shock is associated with an increase in future investment activities. This paper differs from previous studies in that I derive my hypothesis from the implications of return extrapolation. Instead of treating the abnormal trading volume as an exogenous shock, I show that extrapolative belief could contribute to individual stocks' volume spikes and examine the pricing effects.

Second, this paper contributes to the rapidly growing literature on the extrapolation bias. The concept of extrapolative belief has been widely applied to explain asset pricing dynamics (Amromin and Sharpe, 2014; Hirshleifer, Li, and Yu, 2015; Barberis et al., 2015, 2018; Choi and Mertens, 2019; Jin and Sui, 2021) at the aggregate market. Empirically, a vast body of work has been done focusing on survey data or the time-series relationship between extrapolation and market anomalies (Bacchetta, Mertens, and Van Wincoop, 2009; Greenwood and Shleifer, 2014; Koijen, Schmeling, and Vrugt, 2015; Cassella and Gulen, 2018; Bordalo, Gennaioli, La Porta, and Shleifer, 2020; He, Wang, and Yu, 2020). As of the cross-section, Gulen and Woeppel (2022) and Da et al. (2021) explore the pricing implications of extrapolative belief for individual stocks. I add to this branch of research by studying the cross-sectional interaction between extrapolative belief, trading volume, and expected returns. It should be emphasized that my focus is not to provide a complete explanation of the high-volume premium, nor to argue that EXVA is a "better" proxy for cross-sectional extrapolative expectation. Instead, I borrow from existing literature and simply measure the extrapolative value with an intuitive construction. I provide empirical evidence that the extrapolative belief is an important influencing factor of the volume-return relationship.

Finally, this paper belongs to a broad literature that investigates the conditional performance of asset pricing anomalies (Frazzini, 2006; DellaVigna and Pollet, 2009; Hirshleifer, Lim, and Teoh, 2011; Da, Gurun, and Warachka, 2014; Wang, Yan, and Yu, 2017; An, Wang, Wang, and Yu, 2020; Duan, Guo, Li, and Jun, 2020; Chen, He, Tao, and Yu, 2021). I extend the literature by investigating the variation of the high-volume premium with respect to extrapolation and show that extrapolative belief could have meaningful economic implications on the cross-section of expected stock returns.

The rest of the paper is organized as follows. Section 2 describes the data and key variables used in empirical tests. Section 3 presents the main empirical findings. Section 4 discusses and examines the economic mechanism. Additional robustness tests are reported in Section 5. Section 6 concludes.

2 Data and variables

2.1 Data source

The main sample used in this paper contains all common stocks (share codes 10 or 11) listed in NYSE/AMEX/NASDAQ, ranging from July 1965 to December 2021. The daily and monthly individual stock data is from the Center for Research in Security Prices (CRSP) and accounting information from COMPUSTAT. Analyst forecast data is available from the Institutional Brokers Estimate System (I/B/E/S). Quarterly data on institutional holdings (13F) is obtained from Thomson-Reuters starting from December 1980. The five-factor

(Fama and French, 2015) data is downloaded from Kenneth French's website. Finally, to construct the time-series of DOX (Cassella and Gulen, 2018), data on surveys of expectations of future stock market returns from American Association of Individual Investors Investor Sentiment Survey and the Investor Intelligence Survey is collected through Datastream.

I apply common filters that exclude stocks with a price of less than $5 per share or with trading days less than 10 in the portfolio formation month to mitigate the microstructure and illiquidity concerns associated with these securities. Since it has been documented that the properties of trading volume around firm public announcements could be very different from non-announcement trading days (Beaver, 1968; Bamber and Cheon, 1995; Kandel and Pearson, 1995; Garfinkel and Sokobin, 2006), I follow Gervais et al. (2001), Kaniel et al. (2012), and Wang (2021) that exclude stocks with a dividend or earnings announcement during a three-day window around the formation date. I also adjust stock returns for delisting using the procedure of Shumway (1997). [2]

2.2 Definitions of key variables

2.2.1 Abnormal trading volume

A stock's abnormal trading volume is a time-series concept, represented by the percentile position of the formation-day volume relative to the distribution of reference-period trading volumes. Following Gervais et al. (2001), Kaniel et al. (2012), and Wang (2021), for each stock-month, I calculate the percentile position of the trading volume on the last day of the month over a 50-day window volumes prior to the formation day (inclusive). I also use a discrete version of abnormal volume that classifies a stock as low-(high-) abnormal volume if its volume percentile is lower (higher) than 10% (90%). The remaining eight deciles can be similarly defined.

2.2.2 Extrapolative value

The key variable is the cross-sectional proxy for extrapolative expectation. Inspired by Greenwood and Shleifer (2014), Barberis et al. (2015), Da et al. (2021), and many others, the stock-level extrapolative value (hereafter EXVA) is defined as follows:

$$EXVA_{i,t} = \sum_{l=1}^{L} w_l R_{i,t-l}, \tag{1}$$

[2] Specifically, I use the delisting return from CRSP for each delisted stock. If the corresponding delisting return data is not available, I use -30% for cases where the delisting is coded as 500, 520, 551-573, 574, 580, or 584. Otherwise, the delisting return is assumed to be -100%

where $R_{i,t-l}$ is stock i's return on day $t-l$ and $w_l = \lambda^l/(\sum_{k=0}^{L-1}\lambda^k)$ is $R_{i,t-l}$'s weight in the computation of extrapolative value. The choice of using daily returns to calculate EXVA mainly comes from the requirement to clearly separate the abnormal volume formation day and the reference period. While there is no clear guidance for the choice of the decay parameter (λ) in a daily frequency, I refer to Da et al. (2021) and assign λ to be 0.75. [3] To be consistent with the high-volume premium literature, I use the same length of reference window as Wang (2021) and let L to be 49. That is, I skip the last trading day each month when calculating EXVA. [4]

In robustness tests, I also consider an alternative measure to proxy for extrapolative expectation, the price-path convexity of Gulen and Woeppel (2022). This variable is calculated by subtracting the average daily closing price from midpoint between the first day and last daily colsing price of the reference period, divided by the average closing price. I follow the same method of Gulen and Woeppel (2022) to construct the price-path convexity and leave the details in the robustness section.

Figure 1 depicts the timeline for variable calculation and portfolio construction. At the end of each month, I calculate firm-level extrapolation measures using historical returns and prices data spanning the reference period. The trading volume on the formation day is used to define the abnormal volume.

[Place Figure 1 about here]

3 Extrapolation and the high-volume premium

3.1 Summary statistics

This section reports summary statistics and results of single-sorted portfolios. Table 1 presents the distribution of variables used in my main analysis and their pair-wise correlations. The set of firm characteristics includes past return performance, logarithm of market value (Size), book-to-market ratio (BM), CAPM beta (Beta), idiosyncratic volatility (IVol), one-month turnover ratio, and illiquidity using Amihud (2002) measure (Illiq). Since

[3] The main empirical results do not rely heavily on the specification of λ and I get similar conclusions when taking λ as 0.6 or 0.9.

[4] I exclude formation-day returns data for two reasons. First, from the theoretical perspective, EXVA is a proxy for extrapolative *expectation*, which should be formed *before* real tradings happen. It follows that one needs to construct EXVA using only reference-period data because a stock's return is calculated using closing price, while the trading volume is a quantity over the entire trading day. Second, from empirical considerations, it would be inappropriate to include the formation day when calculating EXVA since there would be significant price pressure associated with abnormal trading volumes.

EXVA is a linear combination of historical returns, high-EXVA firm mechanically tends to have good past performance. This is reflected by the positive correlation coefficient (0.528) between EXVA and contemporaneous one-month return. In addition, EXVA is only weakly correlated with AbVol, the abnormal volume.

[Place Table 1 about here]

Table 2 reports summary statistics for portfolios sorted by EXVA. At the end of each month, I sort stocks into deciles using NYSE breakpoints. Each portfolio's excess return is calculated as the value-weighted one-month ahead returns of individual stocks minus the risk-free rate. Stocks are equally weighted in calculating other portfolio-level characteristics. Table 2 shows that there is no monotonic pattern of EXVA portfolios in terms of abnormal volume or risk exposure. In particular, high-EXVA stocks underperform low-EXVA stocks. The return spread between the bottom and top EXVA quintiles is 0.92% bps per month ($t = 5.58$), which is consistent with the hypothesis that high-EXVA stocks are overpriced. [5]

[Place Table 2 about here]

In addition, Table 2 suggests that both low-EXVA stocks and high-EXVA stocks could receive an abnormally high trading volume. Across decile portfolios, the average trading volume percentile first monotonically decreases with respect to EXVA, from 56.3% to 48.3%, and then increases to 63.3%. To illustrate this relation more clearly, Figure 2 presents 25 portfolios sorted by EXVA and their average volume percentiles. The trading volume spike is found to exist in both the low-EXVA group and the high-EXVA group, suggesting a U-shape relation between extrapolative value and abnormal volume. Note that EXVA does not include month-end returns, the relationship illustrated in Figure 2 should reflect the contemporaneous price pressure.

[Place Figure 2 about here]

3.2 Decomposing abnormal volume

Before presenting the main empirical findings, I first provide evidence that the extrapolative value contributes significantly to abnormal trading volume and the extrapolation component is negatively priced on the cross-section.

[5] In later robustness tests, I show that the pricing of EXVA is not driven by the short-term reversal effect. In particular, after orthogonalizing EXVA with the contemporaneous one-month return, the residual EXVA still negatively and significantly predicts future returns.

Figure 2 manifests that the relationship between abnormal volume and extrapolative value is non-linear and potentially follows a quadratic form. Therefore, I quantify this relation by running monthly cross-sectional regressions of abnormal volume against *contemporaneous* extrapolative value:

$$AbVol_{i,t} = a_t + b_t EXVA_{i,t} + c_t EXVA_{i,t}^2 + Controls_{i,t} + \varepsilon_{i,t}. \tag{2}$$

In each regression, I control for cumulative one-month return (REV), cumulative reference-period return (REFRET), firm size, book-to-market ratio, market beta, idiosyncratic volatility, and illiquidity. [6] Following Wang (2021), I also consider firm characteristics such as annual asset growth, return on equity (Hou, Xue, and Zhang, 2015), and coskewness (Harvey and Siddique, 2000). To make coefficients comparable, I normalize each independent variable to have zero mean and unit variance each month.

Table 3 shows the time-series averages of the estimated coefficients and adjusted R^2. While EXVA is strongly correlated with abnormal volume, adding a square-EXVA term significantly improves the model explanatory power (from 1.536% to 4.271%). The positive coefficients on EXVA and EXVA^2 confirms the pattern in Figure 2: both low-EXVA and high-EXVA contribute to the surge of trading volume. For other firm characteristics, I find asset growth (AG), profitability (ROE), illiquidity, and coskewness (COSK) have marginal contribution relative to EXVA. On average, the R^2 only improves 0.36% by separately adding these variables. Both firm size and valuation ratio generates incremental explanatory power for abnormal volume, after controlling for EXVA. In addition, the reference-period return is negatively correlated with abnormal volume, which consistent with the negative correlation between momentum and abnormal volume shown in Wang (2021).

[Place Table 3 about here]

Next, I project abnormal volume onto extrapolative value. The fact that trading volume is abnormally high for both low EXVA stocks and high EXVA stocks implies that the surge of trading may contain an extrapolation component, in addition to the arrival of new information, investor attention, and other factors. Therefore, I would expect the extrapolation component, i.e., abnormal volume from biased expectation, to be negatively priced in the cross-section. In contrast, the residual value should be positively priced, generating the high-volume premium.

Specifically, I run cross-sectional regression of abnormal volume against extrapolative

[6] When calculating REV, data on month-end days is deleted.

value each month and define projected abnormal volume (PROJ AbVol) as the predicted value. Consequently, the residual abnormal volume (RES AbVol) is the residual from the regression. Table 4 presents Fama-MacBeth regressions using PROJ AbVol and RES AbVol. I find the estimated coefficient on PROJ AbVol is significantly negative. whereas the slope of RES AbVol is significantly positive, which is consistent with my conjecture.

[Place Table 4 about here]

3.3 Double-sorts

Now I turn to the key tests of my hypothesis. At the end of each month, stocks are first sorted into quintiles by EXVA. Within each EXVA portfolio, I divide individual stocks into five groups based on abnormal volume percentile. Value-weighted portfolio returns in the next month are then calculated. In addition to conditional double-sorting, I also follow previous studies and consider an independent-sorting design. Specifically, I directly classify stocks into different abnormal volume groups. A stock is identified as high (low) volume if its volume percentile belongs to the highest (lowest) three deciles. The fourth to seventh deciles are classified as neutral. Figure 3 provides a visualization of the interaction between EXVA and HVP under the independent-sort setting. The HVP tends to decrease with respect to EXVA, and the return spread becomes insignificant or even reverses the sign slightly for the top EXVA quintile.

[Place Figure 3 about here]

The detail of the main result is presented in Table 5. Panel A and Panel B report the average returns and five-factor model (Fama and French, 2015) alphas of 5×5 conditional double-sorts portfolios. One sees the high-volume premium is more pronounced for low-EXVA stocks and the return spread is highly significant. Among high EXVA groups, however, high-abnormal volume stocks do not significantly outperform low-abnormal volume stocks. More importantly, the HVP of the low extrapolative value group is 52 bps higher than that of the high extrapolative value group. The difference remains significant after controlling for traditional asset pricing factors (51 bps with $t = 2.35$).

In Panel C and Panel D of Table 5, I use the abnormal volume classification, which is independently identified by each stock's own trading history, to construct 5×3 sorting portfolios. The independently double-sorted portfolios ensure that the abnormal volume spread is similar across each EXVA quintile. Again, I find the high-volume premium tends to decrease with respect to extrapolative value and the difference in HVP between low-EXVA and high-EXVA groups is highly significant (60 bps in return and 58 pbs in alpha).

I also find that the underperformance of high-EXVA stocks monotonically increases as the trading volume becomes abnormally high. For example, Table 5 shows that the return spread between low-EXVA stocks and high-EXVA stocks is around 0.45% per month for stocks identified as abnormally low volume, whereas the return spread is nearly 1% for stocks with abnormally high trading volumes.

It's worth mentioning that Wang (2021) also applies standard asset pricing method such as double-sorting to explore the cross-sectional heterogeneity of HVP. However, the evidence of Wang (2021) is mixed in the sense that either no significant variation is detected or the variation only holds for some special cases. [7]

In sum, the result suggests a cross-sectional heterogeneity of the high-volume premium. The hypothesis that a higher level of extrapolative expectation with an abnormally high trading volume mitigates the HVP and increases the degree of overpricing is supported.

[Place Table 5 about here]

3.4 Fama-MacBeth regressions

Although the portfolio-sorting approach has the advantage that it is intuitive and does not impose any functional form on the relation I seek to uncover, it cannot explicitly control for other variables that could predict returns. In this section, I perform a series of Fama-MacBeth (1973) regressions, which allow me to conveniently control for additional variables.

I control for a battery of return predictors in regressions, including past return performance, firm size, book-to-market ratio, CAPM beta, idiosyncratic volatility, share turnover, and illiquidity. Independent variables are winsorized at 1% and 99% percentiles on each cross-section. For ease of interpreting coefficients, I normalize independent variables each month to have zero mean and unit variance.

Columns (1) and (2) of Table 6 confirms the portfolio-sorting results that both abnormal volume and extrapolative value have very strong predictive power for future stock returns. A one standard deviation increase in abnormal volume percentile is associated with a 32 bps increase in future returns. The coefficient on EXVA has a similar magnitude but an opposite

[7] For example, Wang (2021) finds that the HVP is stronger among firms with low institutional ownership, but the difference is insignificant after adjustment using the five-factor model (Fama and French, 2015). It is also shown that the HVP can be absorbed by β^{VXO}, the market volatility beta, however, the result only hold for a much shorter sample period due to the availability of VIX data.

sign. In column (3), the intersection term of abnormal volume and EXVA is significantly negative, which is consistent with the double-sorting results that the HVP is mitigated as extrapolative value becomes high. Although the magnitude of the estimated coefficient is reduced from 0.085 to 0.060 after controlling for other firm characteristics, the t-statistics (-3.25) suggests that the interaction of EXVA and AbVol remains significant. On the other side, the negative sign of the coefficient also indicates that an abnormal increase of trading volume amplifies the negative pricing effect of EXVA, which is consistent with my hypothesis.

[Place Table 6 about here]

The fluctuation of trading volume is affected by many factors. For example, an increase in trading activity could come from investors' liquidity demand or reflect information flows. Therefore, I control for additional interaction terms and examine whether they absorb the effect of extrapolative value. I consider abnormal volume interacting with firm size, illiquidity, idiosyncratic volatility, shocks to illiquidity (UIlliq), and shocks to idiosyncratic volatility (UIVol). These variables are also studied in Wang (2021) through double-sorted portfolios. Table 7 reports the result. It is shown that controlling for additional predictors does not significantly influence the estimate of the slope of the EXVA-AbVol interaction term, as the magnitudes of coefficients do not deviate much from those in Table 6.

[Place Table 7 about here]

In sum, I find consistent evidence in support of the hypothesis that higher (lower) extrapolative value with an abnormally high trading volume implies a relatively higher degree of overpricing (underpricing), which reduces (strengthens) the high-volume premium.

4 Tests of economic mechanism

In this section, I examine the economic mechanism that the extrapolative belief drives the cross-sectional heterogeneity of HVP. In section 4.1, I use DOX, proposed by Cassella and Gulen (2018), to show that the conditional difference of HVP can be predicted by the degree of market-wide extrapolation level. In section 4.2, I investigate the role of Merton (1987) hypothesis and show that my result is not likely to be driven by stock visibility. Section 4.3 presents placebo tests using simple one-month return and reference-period return to highlight the importance of the decay structure in calculating extrapolative value.

4.1 Evidence from predictive regressions

Cassella and Gulen (2018) propose a recursive estimate of the degree of extrapolative weighting in investors' beliefs (DOX) using survey data on expectations of stock returns. The DOX time-series is measured as the relative loading of investors' expectations on recent-versus-distant returns and has a considerable variation over time. Since Cassella and Gulen (2018), a growing literature has noticed the role of DOX in financial markets. For example, DOX has been found to have strong predictive power for stock return synchronicity (Chue, Gul, and Mian, 2019), a broad set of overreaction-related anomalies (He et al., 2020), the low-risk anomalies (Liu, Su, Wang, and Yu, 2021), and equity term premium (Cassella, Golez, Gulen, and Kelly, 2021).

In this section, I examine the predictive power of DOX for the high-volume premium. Specifically, if the extrapolative expectation does lead to a cross-sectional variation of the HVP, then I would expect that the difference in HVP between low EXVA stocks and high EXVA stocks (hereafter I refer to this difference as HVPDIFF) to be larger following a high level of DOX. Specifically, I perform following time-series regressions:

$$\begin{aligned} &Y_t = a + b_1 DOX_{t-1} + b_2 Control_{t-1} + b_3 Factor_t + \varepsilon_t, \\ &Y_t \in \{HVP1_t, HVP5_t, HVPDIFF_t\}, \end{aligned} \tag{3}$$

in which DOX_{t-1} is one-month lagged DOX, constructed following the procedure of Cassella and Gulen (2018). $HVP1_t$ and $HVP5_t$ are the high-volume premium among stocks with low EXVA and stocks with high EXVA, respectively, and $HVPDIFF_t = HVP1_t - HVP5_t$. Since DOX tends to be correlated with investor sentiment (He et al., 2020), I include the investor sentiment index of Baker and Wurgler (2006) as a lagged control variable. [8] I also include traditional asset pricing factors (Fama and French, 2015) to control for the contemporaneous pricing effect.

In Panel A of Table 8, columns (1) to (3) present the baseline results. I find that DOX has strong predictive power for cross-sectional HVP variation and the predictability mainly comes from low-EXVA stocks (HVP1). Economically, a one standard deviation increase in DOX implies a 57 bps increase in HVPDIFF. Although the estimated coefficient on DOX from the high-EXVA group (HVP5) is not statistically different from 0, the negative sign of the estimate is consistent with previous analysis.

In columns (4) to (9) of Panel A, I control for traditional asset pricing factors such

[8] The investor sentiment index data is available from 1965 to 2018 and can be obtained in Jeffrey Wurgler's website (http://people.stern.nyu.edu/jwurgler/).

as market excess return (Mkt-Rf), small-minus-big (SMB), high-minus-low (HML), robust-minus-weak (RMW), and conservative-minus-aggressive (CMA). The result shows that the coefficient on lagged DOX is not affected by contemporaneous factors. Panel B of Table 8 reports regression results by adding lagged investor sentiment index as an additional predictor. It is shown that investor sentiment has only minor effect on the high-volume premium and the estimated coefficients on DOX are similar to those of Panel A. In Appendix A.2, I also report results using HVPDIFF constructed from independent-sorted portfolios. It turns out the estimation does not change substantially.

[Place Table 8 about here]

4.2 The role of stock visibility

The prevailing explanation of the high-volume premium builds on the Merton (1987) hypothesis. The intuition is that the arrival of a positive volume shock increases a stock's "visibility", which attracts potential investors' attention and leads to subsequent demand and price appreciation. An alternative interpretation of my finding is that past "losers" are more likely to have fallen out of investors' eyes and receive less attention. Therefore, a positive volume shock generates greater recognition toward these stocks and hence larger HVP. In this section, I directly test this explanation by applying stock visibility proxies and examining whether my results are dominated by this channel.

I consider four variables to proxy for the visibility, or the level of investor recognition, of a stock. The first one is analyst coverage (Aboody, Lehavy, and Trueman, 2010; Bali, Hirshleifer, Peng, and Tang, 2021), calculated as the number of analysts covering the firm. A high analyst coverage implies a high degree of investor recognition. The second measure is a firm's membership in Standard and Poors (S&P) 500 index (Israeli et al., 2021). If a stock is selected as a constituent of a major market index such as S&P 500 index, then it should be more visible to investors. Thirdly, I use the change in breadth of institutional ownership, defined as the percentage change in the number of Form 13F filers (Lehavy and Sloan, 2008; Israeli et al., 2021). [9] Lehavy and Sloan (2008) argue that the knowledge about a security is increasing in the number of investors that own the security. Thus, stock with a high change in breadth of institutional ownership tends to be more recognized. Finally, similar to the third proxy, I use the institutional ownership, which is the percentage of total shares outstanding held by 13F filers (Edelen, Ince, and Kadlec, 2016).

[9] Specifically, for each firm i, the variable is calculated as $\frac{Num_{i,t}-Num_{i,t-1}}{TNum_{i,t-1}}$, where $Num_{i,t}$ is the number of 13F filers holding stock i at quarter t and $TNum_{i,t-1}$ is the total number of 13F filers at quarter $t-1$.

In Table 9, I interact these measures with abnormal volume to examine the role of stock visibility. One can see that the intersection terms involving analyst coverage, S&P500 index indicator, and institutional ownership are all highly significant. The intersection term of the change in breadth of institutional ownership is also significant, although the effect is weaker compared with other proxies. The negative sign of the coefficient is consistent with the prediction of the Merton (1987) hypothesis: with less investor recognition, these stocks are ignored by a large fraction of the investors and an increase in visibility from abnormal treading volume attracts attention and leads to a higher subsequent return.

More importantly, Table 9 shows that the effect of extrapolative expectation on HVP is not a simple mimic of stock visibility. The interaction term of EXVA and AbVol remains significantly negative after controlling for the visibility channel and the economic magnitude hardly changes on average.

Overall, the analysis in this section suggests that the extrapolative expectation indeed generates a cross-sectional heterogeneity of the high-volume premium. The result is not a replication of stock visibility or investor recognition, although this channel does play an important role in driving the HVP.

[Place Table 9 about here]

4.3 Placebo tests using equal-weighted historical performance

The extrapolative value used in this paper is intuitive. However, an inevitable issue is its positive correlation with the simple arithmetic average of historical returns. Although it has been shown that the effect of EXVA on HVP is not driven by stock visibility, one could still worry that the exponential decay model is redundant. In fact, the seminal work of Gervais et al. (2001) conduct a "momentum subsample" analysis (Table 5 of Gervais et al. (2001)) that is related to my findings. Specifically, the authors split stocks into subsamples of "winners" and "losers" based on reference period returns and find that the high-volume premium is stronger for past losers. In this section, I show that a certain degree of decay in weighting past performance is necessary for interpreting my findings through an extrapolative belief explanation.

Specifically, I replace EXVA with one-month return and reference-period return, separately. [10] If my findings are mainly driven by over-extrapolation bias and EXVA better

[10] When calculating one-month return and reference-period return, I skip the last trading day of the month to be consistent and comparable with previous analysis. The calculation is equivalent to setting $\lambda = 1$ in equation (1).

captures short-horizon extrapolative expectation than equal-weighted performance, then using one-month return or reference-period return should not exert an effect on the HVP through a similar mechanism as EXVA.

First, I show that simply using equal-weighted performance variables can also generate significant cross-sectional variation of HVP. Table 10 presents the double-sorting results following the same procedure as in Table 5, replacing EXVA with equal-weighted performance variables. I find the difference in HVP between low past performance stocks and high past performance stocks is even larger relative to the results using EXVA, and the spread ranges from 67 bps to 104 bps.

[Place Table 10 about here]

However, when examining the predictive regressions as in section 4.1, no predictive power of DOX is detected. Table 11 reports the estimate of coefficients on predictors. While both one-month performance and reference-period performance can generate a greater HVP heterogeneity on the cross-section, such variation cannot be positively predicted by market-wide extrapolation level, if any, the effect is negative. For example, the coefficient on DOX in predicting HVP5 is positive, which is inconsistent with the extrapolation channel. As for the key prediction, HVPDIFF, the estimated slope of DOX is negative and the magnitude of t-statistics is generally less than 1. This finding thus highlights the special role of the decay structure in calculating EXVA, beyond its correlation with simple past performance.

[Place Table 11 about here]

5 Additional robustness checks

I conduct several robustness tests to assess the robustness of the main results. Particularly, I first examine the weighting scheme used in portfolio analysis and consider alternative extrapolation measures. Then, I validate EXVA as a return extrapolation proxy by examing its predictability for subsequent order imbalances and show that it is distinct from compensation for liquidity provision (Nagel, 2012). Finally, I provide evidence that the empirical findings are robust to controlling for other mispricing measures.

5.1 Weighting scheme in the portfolio analysis

In my baseline analysis, value-weighted portfolios are formed to alleviate the concern that the result is dominated by very small firms. However, the cross-sectional variation of HVP could also be driven by very large firms under value weighting. In addition, Asparouhova, Bessembinder, and Kalcheva (2013) propose the prior gross return weighting, which is designed to alleviate the liquidity bias and addresses the price noise concern. Therefore, I repeat the double-sorting analysis using equal-weighted and gross return-weighted portfolios.

Table 12 reports the results. The main difference between value-weighting and the two alternative weighting schemes is that the HVP is significant across all EXVA groups. However, the results in Table 12 confirm a strongly monotonic pattern of HVP along EXVA portfolios. Specifically, among low-EXVA firms, the HVP is highly significant with an average monthly premium over 1.30%; the HVP deceases monotonically with respect to EXVA and the HVP is around 0.45% among high-EXVA firms. On average, the difference of HVP between low-EXVA and high-EXVA groups is 95 bps in raw returns and 89 bps in five-factor alphas.

[Place Table 12 about here]

5.2 Choice of λ and alternative extrapolation measure

The decay parameter λ decides the relative weight that investors put on recent performance against distant performance when forming extrapolative expectation. The larger the λ, the slower the weight decreases into distant realizations. Since EXVA is calculated using daily stock returns data, it is reasonable to claim that λ should not be too small. This is because a tiny λ implies that investors only care about a few days performance, say one or two, when evaluating a stock, which is unrealistic. Nevertheless, I vary the benchmark value (0.75) and choose a smaller λ, 0.60, to re-calculate EXVA. I also consider a higher value (0.90) to repeat the calculation. [11] Table 13 reports the result using EXVA from different λ choices. It is shown that the interaction term of EXVA and abnormal volume remains negative and significant.

Gulen and Woeppel (2022) propose a convexity measure to proxy for return extrapolation of individual firms. Their construction is motivated by the "acceleration" concept of Greenwood, Shleifer, and You (2019) in measuring the price path of a recent run-up. Gulen and Woeppel (2022) show that price-path convexity does capture the effect of extrapolative expectations and has strong predictive power for future returns at both firm-level and the aggregate market-level.

[11] It should be mentioned that the choice of $\lambda = 0.90$ does not completely fit the extrapolation story, since in this case the resulting EXVA is highly correlated with the arithmetic average value.

For a given estimation window, a stock's price-path convexity at time t is given by:

$$Convexity_{i,t} = \frac{\frac{Price_{first,t} - Price_{last,t}}{2} - Price_{avg,t}}{Price_{avg,t}} \tag{4}$$

where $Price_{first,t}$ and $Price_{last,t}$ are respectively the first daily closing price and the last daily closing price of the estimation window, and $Price_{avg,t}$ is the average daily closing price over the window. I adjust for stock splits and stock dividends following the same procedure of Gulen and Woeppel (2022). In Table 13, one sees that the positive predictive power of abnormal volume is also significantly mitigated by price-path convexity, supporting the extrapolation hypothesis. [12] Importantly, I show in Appendix A.2 that DOX again positively predicts HVPDIFF using portfolios constructed by price-path convexity. This result manifests that EXVA and price-path convexity share similar mechanism in driving the cross-sectional variation of the high-volume premium, i.e., the short-term return extrapolation.

[Place Table 13 about here]

5.3 EXVA and order imbalances

If investors' trading behaviors are consistent with their subjective expectations and EXVA can capture the extrapolative belief, then it is expected that EXVA positively predicts the propensity to buy/sell a stock. In Appendix A.1, I use transaction data from TAQ to calculate month-end order imbalances and test this conjecture. I estimate three types of order imbalances, based on the number of trades, volume (in shares), and dollar value. In univariate analysis, Figure A1 shows that the order imbalances tend to increase with respect to EXVA. Among high-EXVA stocks, the month-end order imbalances are positive, showing a purchase demand for these stocks. Meanwhile, the order imbalances tend to decrease and turn to negative for low-EXVA stocks, suggesting a heightened likelihood of selling. In cross-sectional regressions, Table A1 shows that a one standard deviation increase in EXVA is found to be associated with a 0.007 increase in order imbalances, after controlling for a battery of firm characteristics.

[12] In this paper, I unify the calculation of extrapolation measure and thus use a 49-day window to compute price-path convexity. In their original paper, Gulen and Woeppel (2022) construct the price-path convexity with a one-month period. As a result, the pricing effect of Convexity I estimate is different from Gulen and Woeppel (2022), who find a very strong relationship between Convexity and future returns. However, this paper focuses on the interaction between extrapolation and abnormal volume, hence the main conclusions are unaffected by this difference.

5.4 EXVA and short-term reversal

Next, I provide a battery of analyses to show that the pricing pattern of EXVA is distinct from the short-term reversal effect. In Appendix A.3, I construct the residual extrapolative value (RES EXVA) by computing the residuals of monthly cross-sectional regressions of EXVA against REV, the contemporaneous cumulative one-month return. I also calculate a residual short-term reversal (RES REV) in a similar fashion by exchanging the dependent variable and the regressor.

By separately sorting stocks into decile portfolios based on RES EXVA and RES REV, I find completely different return patterns. Consistent with the original EXVA, high RES EXVA stocks underperform low RES EXVA stocks and a trading strategy by buying the bottom decile stocks and selling the top decile stocks generates model-adjusted returns above 1% per month. [13] In sharp contrast, the short-term reversal is completely absorbed by EXVA. Surprisingly, RES REV turns out to be positively associated with future returns, contradicting the negative pricing effect of REV. More importantly, I repeat Fama-MacBeth regressions (Table A7) and predictive regressions (Table A8) with RES EXVA and find that the main empirical finding of this paper still hold for the REV-orthogonalized extrapolative value. Therefore, the result suggests that the pricing implication of EXVA cannot be simply ascribed to returns from liquidity provision (Nagel, 2012).

5.5 Further control for mispricing measures

Finally, I discuss the difference between my findings and Han, Huang, Huang, and Zhou (2022, hereafter HHHZ). Using the mispring score of Stambaugh, Yu, and Yuan (2015), HHHZ explore the intertemporal volume-return relation and mispricing. They find that trading volume has an amplification effect on mispricing in that trading volume is positively (negatively) related to expected return among underpriced (overpriced) stocks. At first glance, this paper's finding seems to overlap with HHHZ since high-EXVA (low-EXVA) stocks are overpriced (underpriced), and the high-volume premium is mitigated (enhanced) among high-EXVA (low-EXVA) stocks. However, there are three distinguished differences between my findings and HHHZ. First, I focus on abnormal trading volume, which is determined by a *time-series* comparison of trading volumes for each individual stock, while the volume studied in HHHZ is a *cross-sectional* concept. A high level of AbVol does not

[13] I consider a broad set of asset pricing models such as CAPM, the three-factor model, the five-factor model (Fama and French, 2015), the momentum-augmented model, the short-term reversal-augmented model, the four-factor model of Hou et al. (2015), the five-factor model of Hou, Mo, Xue, and Zhang (2021), and the behavioral factor model of Daniel, Hirshleifer, and Sun (2020).

necessarily implies a high volume quantity. [14] Second, the mispricing measure used in HHHZ mainly captures the *long-term* overpricing/underpricing, whereas EXVA reflects *short-term* mispricing. Finally, while HHHZ interpret their findings through investor disagreement, this paper is motivated by return extrapolation bias.

That being said, whether my results hold under mispricing controls still remains an empirical question. In Appendix A.4, I follow HHHZ to construct mispricing measures and intersect the mispricing variables with abnormal volume. In Fama-MacBeth regressions, I find the negative slope of the EXVA-AbVol term remains economically and statistically significant, but the slope of the mispricing-AbVol term is positive and marginally significant. It thus suggests that the story of this paper is different from the volume-return relation studied in HHHZ.

6 Conclusion

The intertemporal relationship between trading volume and stock return is a fundamental issue in asset pricing. Extending prior literature studying the high-volume premium (Gervais et al., 2001), this paper documents that the abnormal trading volume contains an extrapolation component and the HVP varies substantially across portfolios with different levels of extrapolative evaluation. Using a simple extrapolative value measure (EXVA) implied by survey evidence and theoretical models, I find that the HVP is significant and strong among firms with a low level EXVA. By contrast, among firms with high EXVA, the HVP is weak.

This paper's analysis builds on theoretical work studying the role of extrapolation in generating volume spike during bubble times. Specifically, the extrapolation hypothesis suggests that an abnormally high trading volume implies a higher degree of overpricing (underpricing) for high (low) extrapolative value stocks, leading to a weaker (stronger) high-volume premium. I test the hypothesis through portfolio-level analysis and individual firm-level regressions. Further, it is shown that the difference between the HVP among low-EXVA firms and high-EXVA firms can be predicted by market-wide extrapolation level (DOX), even after controlling for investor sentiment and traditional asset pricing factors. In addition, I find the results are not dominated by a stock's visibility to investors, although it does play an important role in driving the high-volume premium. In placebo tests, I show that the decay structure in calculating EXVA is necessary.

My findings are robust to different parameter choice and alternative extrapolation proxy

[14] As reported in Table 1, the correlation coefficient between AbVol and Turnover is only 0.05.

such as price-path convexity (Gulen and Woeppel, 2022). The constructed EXVA is distinct from short-term reversal and the results are robust under mispricing controls. However, this paper is silent on why abnormal volume is associated with a positive premium unconditionally. It is still worth for future work to design a more accurate and solid variable to capture cross-sectional extrapolative expectation. While the empirical pattern documented in this paper is helpful for understanding the pricing of trading volume, it would be desirable to develop a formal dynamic asset pricing model. For example, Barberis et al. (2018) present an extrapolative model in which extrapolators' attention "waver" between fundamental signals and price signals. An extrapolation-based framework like Barberis et al. (2018) can help to jointly account for the pricing of abnormal volume, the predictive power of HVP for economic indicators, and the role of investor attention.

References

Aboody, D., R. Lehavy, and B. Trueman. 2010. Limited attention and the earnings announcement returns of past stock market winners. *Review of Accounting Studies* 15:317–344.

Amihud, Y. 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets* 5:31–56.

Amromin, G., and S. A. Sharpe. 2014. From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science* 60:845–866.

An, L., H. Wang, J. Wang, and J. Yu. 2020. Lottery-related anomalies: the role of reference-dependent preferences. *Management Science* 66:473–501.

Asparouhova, E., H. Bessembinder, and I. Kalcheva. 2013. Noisy prices and inference regarding returns. *The Journal of Finance* 68:665–714.

Bacchetta, P., E. Mertens, and E. Van Wincoop. 2009. Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance* 28:406–426.

Baker, M., and J. Wurgler. 2006. Investor sentiment and the cross-section of stock returns. *The journal of Finance* 61:1645–1680.

Bali, T. G., D. Hirshleifer, L. Peng, and Y. Tang. 2021. Attention, social interaction, and investor attraction to lottery stocks .

Bali, T. G., L. Peng, Y. Shen, and Y. Tang. 2014. Liquidity shocks and stock market reactions. *The Review of Financial Studies* 27:1434–1485.

Bamber, L. S., and Y. S. Cheon. 1995. Differential price and volume reactions to accounting earnings announcements. *Accounting Review* pp. 417–441.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer. 2015. X-CAPM: An extrapolative capital asset pricing model. *Journal of financial economics* 115:1–24.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer. 2018. Extrapolation and bubbles. *Journal of Financial Economics* 129:203–227.

Beaver, W. H. 1968. The information content of annual earnings announcements. *Journal of accounting research* pp. 67–92.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer. 2020. Expectations of fundamentals and stock market puzzles. *Working paper* .

Cassella, S., B. Golez, H. Gulen, and P. Kelly. 2021. Horizon bias and the term structure of equity returns. *Available at SSRN 3328970* .

Cassella, S., and H. Gulen. 2018. Extrapolation bias and the predictability of stock returns by price-scaled variables. *The Review of Financial Studies* 31:4345–4397.

Chen, X., W. He, L. Tao, and J. Yu. 2021. Media Coverage and Underreaction-Related Anomalies. *PBCSF-NIFR Research Paper* .

Choi, J. J., and T. M. Mertens. 2019. Extrapolative expectations and the equity premium. *Available at SSRN 3462056* .

Chue, T. K., F. A. Gul, and G. M. Mian. 2019. Aggregate investor sentiment and stock return synchronicity. *Journal of Banking & Finance* 108:105628.

Da, Z., U. G. Gurun, and M. Warachka. 2014. Frog in the pan: Continuous information and momentum. *The review of financial studies* 27:2171–2218.

Da, Z., X. Huang, and L. J. Jin. 2021. Extrapolative beliefs in the cross-section: What can we learn from the crowds? *Journal of Financial Economics* 140:175–196.

Daniel, K., D. Hirshleifer, and L. Sun. 2020. Short-and long-horizon behavioral factors. *The review of financial studies* 33:1673–1736.

DeFusco, A. A., C. G. Nathanson, and E. Zwick. 2017. Speculative dynamics of prices and volume .

DellaVigna, S., and J. M. Pollet. 2009. Investor inattention and Friday earnings announcements. *The Journal of Finance* 64:709–749.

Duan, X., L. Guo, F. W. Li, and T. Jun. 2020. Sentiment, limited attention and mispricing. *Working paper* .

Edelen, R. M., O. S. Ince, and G. B. Kadlec. 2016. Institutional investors and stock return anomalies. *Journal of Financial Economics* 119:472–488.

Fama, E. F., and K. R. French. 2015. A five-factor asset pricing model. *Journal of financial economics* 116:1–22.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of political economy* 81:607–636.

Frazzini, A. 2006. The disposition effect and underreaction to news. *The Journal of Finance* 61:2017–2046.

Garfinkel, J. A., and J. Sokobin. 2006. Volume, opinion divergence, and returns: A study of post–earnings announcement drift. *Journal of Accounting Research* 44:85–112.

Gervais, S., R. Kaniel, and D. H. Mingelgrin. 2001. The high-volume return premium. *The Journal of Finance* 56:877–919.

Glaeser, E. L., and C. G. Nathanson. 2017. An extrapolative model of house price dynamics. *Journal of Financial Economics* 126:147–170.

Greenwood, R., and A. Shleifer. 2014. Expectations of returns and expected returns. *The Review of Financial Studies* 27:714–746.

Greenwood, R., A. Shleifer, and Y. You. 2019. Bubbles for fama. *Journal of Financial Economics* 131:20–43.

Gulen, H., and M. Woeppel. 2022. Price-Path Convexity, Extrapolation, and Short-Horizon Return Predictability. *Available at SSRN 3800419* .

Han, Y., D. Huang, D. Huang, and G. Zhou. 2022. Expected return, volume, and mispricing. *Journal of Financial Economics* 143:1295–1315.

Harvey, C. R., and A. Siddique. 2000. Conditional skewness in asset pricing tests. *The Journal of finance* 55:1263–1295.

He, W., Y. Wang, and J. Yu. 2020. Time Variation in Extrapolation and Anomalies. *PBCSF-NIFR Research Paper* .

Hirshleifer, D., J. Li, and J. Yu. 2015. Asset pricing in production economies with extrapolative expectations. *Journal of Monetary Economics* 76:87–106.

Hirshleifer, D., S. S. Lim, and S. H. Teoh. 2011. Limited investor attention and stock market misreactions to accounting information. *The Review of Asset Pricing Studies* 1:35–73.

Hou, K., H. Mo, C. Xue, and L. Zhang. 2021. An augmented q-factor model with expected growth. *Review of Finance* 25:1–41.

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *The Review of Financial Studies* 28:650–705.

Huang, Z., J. B. Heian, and T. Zhang. 2011. Differences Of Opinion, Overconfidence, And The High-Volume Premium. *Journal of Financial Research* 34:1–25.

Israeli, D., R. Kaniel, and S. A. Sridharan. 2021. The real side of the high-volume return premium. *Management Science* .

Jin, L. J., and P. Sui. 2021. Asset pricing with return extrapolation. *Journal of Financial Economics* .

Kandel, E., and N. D. Pearson. 1995. Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy* 103:831–872.

Kaniel, R., A. Ozoguz, and L. Starks. 2012. The high volume return premium: Cross-country evidence. *Journal of Financial Economics* 103:255–279.

Koijen, R. S., M. Schmeling, and E. B. Vrugt. 2015. Survey expectations of returns and asset pricing puzzles. *Working paper* .

Lee, C. M., and M. J. Ready. 1991. Inferring trade direction from intraday data. *The Journal of Finance* 46:733–746.

Lehavy, R., and R. G. Sloan. 2008. Investor recognition and stock returns. *Review of Accounting Studies* 13:327–361.

Liao, J., C. Peng, and N. N. Zhu. 2021. Extrapolative bubbles and trading volume. *The Review of Financial Studies* forthcoming.

Liu, Q., Z. Su, H. Wang, and J. Yu. 2021. Extrapolation and Risk-Return Trade-offs. *Available at SSRN 4039296* .

Merton, R. C. 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance* 42:483–510.

Nagel, S. 2012. Evaporating liquidity. *The Review of Financial Studies* 25:2005–2039.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix. *Econometrica* 55:703–708.

Shumway, T. 1997. The delisting bias in CRSP data. *The Journal of Finance* 52:327–340.

Stambaugh, R. F., J. Yu, and Y. Yuan. 2015. Arbitrage asymmetry and the idiosyncratic volatility puzzle. *The Journal of Finance* 70:1903–1948.

Statman, M., S. Thorley, and K. Vorkink. 2006. Investor overconfidence and trading volume. *The Review of Financial Studies* 19:1531–1565.

Wang, H., J. Yan, and J. Yu. 2017. Reference-dependent preferences and the risk–return trade-off. *Journal of Financial Economics* 123:395–414.

Wang, Z. 2021. The high volume return premium and economic fundamentals. *Journal of Financial Economics* 140:325–345.

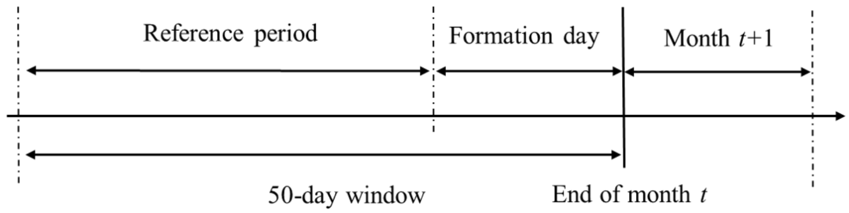

Figure 1. Time line for the calculation of key variables and construction of portfolios.
This figure depicts the time line for calculating extrapolative value (EXVA) and abnormal volume (AbVol), and constructing double-sorted portfolios. Stocks are sorted into quintiles using EXVA computed from reference-period returns. Within each EXVA quintile, stocks are further split into groups based on AbVol measured at the formation day. Portfolios by independent-sorts are formed similarly. Value-weighted portfolios are constructed and held for one month.

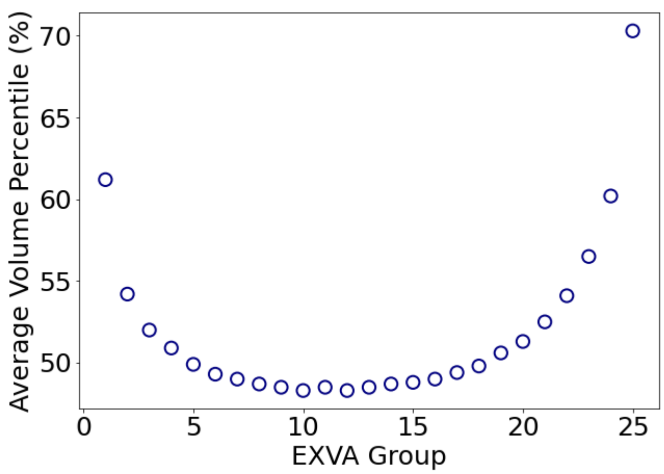

Figure 2. The U-shape relation between EXVA and abnormal volume.
The figure shows the average volume percentiles for 25 EXVA-sorted portfolios. At the end of each month, I sort stocks into 25 groups based on their extrapolative value. Equal-weighted volume percentile is then calculated for each EXVA portfolio. The sample period is from July 1965 to December 2021.

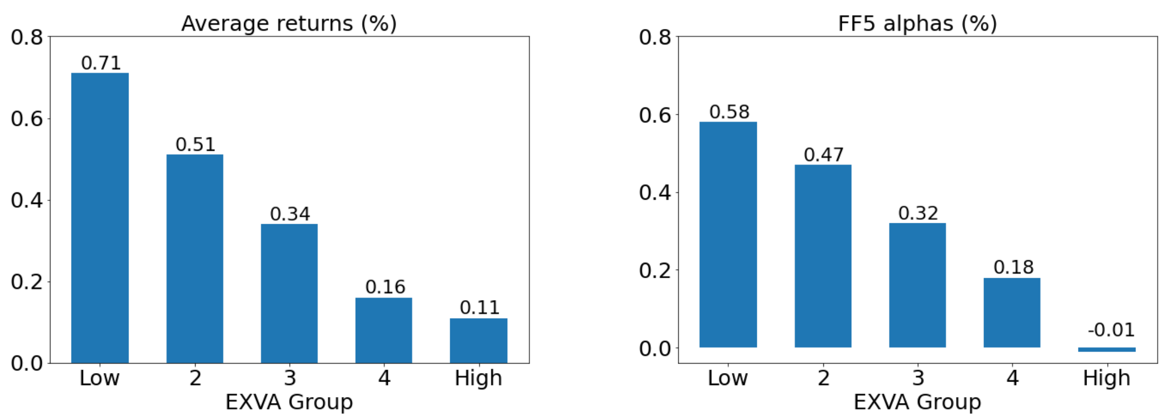

Figure 3. Cross-sectional variation of the high-volume premium.
This figure plots average return and Fama-French five factor alpha spreads (in percentages) between high abnormal volume and low abnormal volume stocks across extrapolative value (EXVA) quintiles. Value-weighted portfolios are constructed by independently sorting stocks based on EXVA and abnormal volume. The sample period is from July 1965 to December 2021.

Table 1. Summary statistics.

This table shows the summary statistics of the main variables used in the paper. Panel A presents the variable distributions and Panel B presents their time-series average of cross-sectional correlations. EXVA is a weighted-sum of past returns, defined as $EXVA_{i,t} = \sum_{l=1}^{L} w_l R_{i,t-l}$, where $w_l = \lambda^l / \sum_{k=0}^{L-1} \lambda^k$. We use daily stock return data to calculate EXVA and choose L and λ to be 49 and 0.75, respectively. AbVol is abnormal trading volume, calculated as the trading volume percentile at the end of each month over a stock's previous 50 trading days. Ret(-1) is return in the last month, Ret(-12,-1) is the cumulative return over the past year with a one-month gap, Ret(-36,-12) is the cumulative return over the past three years with a one-year gap. Size is the log of a firm's market equity. BM is the book-to-equity ratio. Beta is CAPM beta using monthly returns data with a five-year rolling window for regressions. IVol is idiosyncratic volatility defined as the standard deviation of the residuals from the Fama-French three-factor model using daily excess returns within a month with a minimum 15 nonmissing observations. Turnover is last month's turnover ratio and Illiq is the Amihud (2002) measure. All variables are winsorized at 1% and 99% levels at each month. The sample period is from July 1965 through December 2021. We use NYSE, Amex, and Nasdaq common stocks with a price of at least $5 and at least 10 trading days over the last month.

Panel A. Distributions

	EXVA	AbVol	Ret(-1)	Ret(-12,-1)	Ret(-36,-12)	Size	BM	Beta	Ivol	Turnover	Illiq
P10	-0.009	0.142	-0.104	-0.248	-0.297	3.267	0.214	0.407	0.008	0.010	0.001
P25	-0.004	0.297	-0.047	-0.078	-0.054	4.104	0.394	0.680	0.012	0.024	0.004
P50	0.000	0.532	0.008	0.113	0.244	5.197	0.680	1.027	0.018	0.055	0.020
P75	0.005	0.763	0.069	0.358	0.646	6.459	1.036	1.429	0.026	0.101	0.096
P90	0.011	0.908	0.147	0.713	1.268	7.733	1.460	1.878	0.036	0.176	0.324
Mean	0.001	0.528	0.017	0.199	0.416	5.361	0.797	1.094	0.020	0.080	0.149
Std	0.009	0.280	0.110	0.460	0.784	1.697	0.596	0.581	0.012	0.088	0.420

Panel B. Correlations

	EXVA	AbVol	Ret(-1)	Ret(-12,-1)	Ret(-36,-12)	Size	BM	Beta	Ivol	Turnover	Illiq
EXVA	1										
AbVol	0.078	1									
Ret(-1)	0.528	0.074	1								
Ret(-12,-1)	0.012	-0.02	0.004	1							
Ret(-36,-12)	-0.011	-0.008	-0.03	-0.062	1						
Size	0.002	0.004	0.008	0.035	0.085	1					
BM	0.01	0.002	0.025	0.032	-0.302	-0.231	1				
Beta	0.008	-0.013	0.009	0.019	0.05	0.029	-0.114	1			
Ivol	0.105	0.068	0.163	0.017	-0.002	-0.324	-0.015	0.266	1		
Turnover	0.05	0.034	0.113	0.149	0.091	0.152	-0.087	0.356	0.352	1	
Illiq	0.029	-0.009	0.049	0.03	-0.104	-0.492	0.195	-0.109	0.261	-0.16	1

Table 2. One-sort and portfolio characteristics.

This table presents future returns and characteristics of portfolios sorted by extrapolatiove value (EXVA). The definitions of variables are the same as in Table 1. At the end of each month, individual stocks are sorted into deciles based on EXVA and contemporaneous, equally weighted characteristics values are calculated. Mean returns of each portfolio are value-weighted returns in the next month. The sample period is from July 1965 through December 2021.

Group	Mean return	EXVA	AbVol	Ret(-1)	Ret(-12,-1)	Ret(-36,-12)	Size	BM	Beta	Ivol	Turnover	Illiq
Low	0.798	-0.015	0.563	-0.076	0.277	0.535	4.847	1.095	1.268	0.028	0.117	0.265
2	0.748	-0.007	0.504	-0.034	0.211	0.455	5.204	1.152	1.148	0.021	0.079	0.192
3	0.694	-0.004	0.489	-0.017	0.193	0.414	5.355	1.317	1.076	0.018	0.072	0.169
4	0.623	-0.002	0.483	-0.005	0.187	0.395	5.408	2.413	1.032	0.017	0.067	0.159
5	0.528	0	0.483	0.006	0.185	0.385	5.436	1.963	1.003	0.016	0.066	0.159
6	0.417	0.001	0.485	0.018	0.185	0.389	5.484	1.373	1.008	0.016	0.066	0.154
7	0.287	0.003	0.491	0.03	0.188	0.402	5.506	1.513	1.029	0.017	0.07	0.161
8	0.227	0.005	0.507	0.045	0.2	0.423	5.475	1.427	1.07	0.019	0.075	0.163
9	0.254	0.008	0.537	0.069	0.22	0.452	5.347	1.912	1.138	0.022	0.085	0.192
High	-0.119	0.019	0.636	0.149	0.279	0.478	4.914	1.072	1.257	0.032	0.139	0.315
High-Low	-0.917	0.034	0.073	0.225	0.002	-0.056	0.067	-0.022	-0.011	0.004	0.023	0.05

Table 3. Abnormal volume, extrapolative value, and other firm characteristics. This table reports time-series averages of coefficients from monthly cross-sectional regressions. Each month, we run cross-sectional regression of *contemporaneous* abnormal volume (AbVol) against extrapolative value (EXVA), quadratic EXVA, and other firm characteristics:

$$AbVol_{i,t} = a_t + b_t EXVA_{i,t} + c_t EXVA_{i,t}^2 + Controls + \varepsilon_{i,t}.$$

All independent variables are winsorized at 1% and 99% and standardized to have zero mean and unit variance. The sample period is from July 1965 through December 2021. We use NYSE, Amex, and Nasdaq common stocks with a price of at least $5 and at least 10 trading days over the last month. The t-statistics in parentheses are calculated based on Newey and West (1987) adjusted standard errors.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
EXVA	2.260	0.946	0.839	0.886	0.888	0.582	0.846	0.547
	(20.15)	(9.01)	(8.82)	(9.07)	(8.41)	(5.66)	(8.11)	(5.21)
EXVA^2		5.283	5.373	5.555	5.892	5.577	5.456	5.799
		(78.01)	(76.91)	(87.69)	(89.51)	(45.19)	(84.34)	(45.39)
REV			1.010					0.481
			(12.87)					(3.69)
REFRET			-1.296					-1.043
			(-13.65)					(-8.47)
LOGME				0.936				0.492
				(4.63)				(2.08)
LOGBM				0.645				0.554
				(14.04)				(6.17)
Beta					-1.255			-1.184
					(-17.87)			(-12.13)
IVol					-0.069			0.380
					(-0.51)			(3.61)
AG						-0.534		-0.503
						(-5.38)		(-4.57)
ROE						0.422		0.307
						(5.86)		(5.85)
Illiq							-0.954	-1.589
							(-9.10)	(-2.23)
COSK							0.504	0.324
							(9.27)	(6.39)
Adj.R^2(%)	1.536	4.271	5.132	6.770	5.138	4.278	4.983	8.079

Table 4. The pricing of projected abnormal volume and residual abnormal volume.
This table reports time-series averages of coefficients from monthly cross-sectional regressions. PROJ AbVol and RES AbVol are respectively the projected abnormal volume and residual abnormal volume, calculated from monthly cross-sectional regressions of abnormal volume against the quadratic-form extrapolative value. The sample period is from July 1965 through December 2021. We use NYSE, Amex, and Nasdaq common stocks with a price of at least $5 and at least 10 trading days over the last month. The t-statistics in parentheses are calculated based on Newey and West (1987) adjusted standard errors.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PROJ AbVol	**-0.169**		**-0.181**	**-0.201**	**-0.170**	**-0.190**	**-0.136**	**-0.140**
	(-3.57)		**(-3.85)**	**(-4.99)**	**(-4.92)**	**(-6.62)**	**(-6.19)**	**(-6.60)**
RES AbVol		**0.356**	**0.361**	**0.303**	**0.306**	**0.284**	**0.266**	**0.268**
		(14.19)	**(14.52)**	**(13.83)**	**(14.51)**	**(15.57)**	**(15.53)**	**(16.29)**
LOGME				-0.092	-0.077	-0.111	-0.217	-0.191
				(-1.62)	(-1.37)	(-2.06)	(-4.85)	(-4.22)
LOGBM				0.174	0.182	0.133	0.082	0.077
				(3.32)	(3.63)	(2.83)	(2.12)	(2.01)
Ret(-1)					-0.292	-0.360	-0.392	-0.395
					(-5.45)	(-7.89)	(-9.05)	(-9.42)
Ret(-12,-1)						0.325	0.326	0.305
						(6.07)	(6.92)	(6.81)
Ret(-36,-12)						-0.061	-0.038	-0.035
						(-1.79)	(-1.17)	(-1.09)
Beta							0.062	0.060
							(1.17)	(1.26)
Ivol							-0.279	-0.327
							(-5.86)	(-8.53)
Turnover								0.053
								(1.17)
Illiq								0.135
								(3.27)
Intercept	1.094	1.094	1.094	1.136	1.128	1.176	1.204	1.215
	(5.06)	(5.06)	(5.06)	(5.09)	(5.07)	(5.43)	(5.71)	(5.80)

Table 5. Double-sorted portfolio returns and alphas by extrapolative value and abnormal trading volume.

The table reports average returns and alphas for portfolios formed by sorting stocks on EXVA and abnormal trading volume. At the end of each month, we first sort individual stocks into quintiles using EXVA. Panel A shows the conditional sorts using abnormal volume percentiles: within each quintile, then, stocks are sorted into quintile portfolios based on abnormal volume. Panel B presents independent sorts using abnormal volume rank: a stock is classified into the high (low) volume group if its volume percentile belongs to the highest (lowest) three deciles. The remaining stocks are classified as neutral. We form value-weighted portfolios and hold for one month. The five-factor alphas are computed with respect to the five-factor model of (Fama and French, 2015). A stock's EXVA is a weighted-sum of its past returns, defined as $EXVA_{i,t} = \sum_{l=1}^{L} w_l R_{i,t-l}$, where $w_l = \lambda^l / \sum_{k=0}^{L-1} \lambda^k$. We use daily stock return data to calculate EXVA and choose L and λ to be 49 and 0.75, respectively. The sample period is from July 1965 through December 2021. The t-statistics are calculated based on Newey and West (1987) adjusted standard errors and reported in parentheses.

	Abnormal volume portfolios									
	Panel A. Average returns: percentile					Panel B. Five-factor alphas: percentile				
EXVA	Low	3	High	H-L	DiD	Low	3	High	H-L	DiD
Low	0.46	0.82	1.07	0.61		0.04	0.36	0.52	0.48	
	(1.50)	(2.86)	(4.01)	(3.94)		(0.35)	(2.92)	(3.56)	(3.16)	
2	0.36	0.67	0.98	0.63		-0.15	0.17	0.50	0.65	
	(1.35)	(2.72)	(3.97)	(4.52)		(-1.69)	(1.95)	(5.04)	(4.90)	
3	0.34	0.50	0.73	0.39		-0.13	0.05	0.25	0.38	
	(1.25)	(2.09)	(3.14)	(2.95)		(-1.54)	(0.60)	(2.69)	(3.05)	
4	0.18	0.36	0.37	0.19		-0.25	-0.09	-0.08	0.17	
	(0.70)	(1.48)	(1.52)	(1.43)		(-2.69)	(-1.13)	(-0.93)	(1.28)	
High	0.01	0.08	0.10	0.09	**-0.52**	-0.40	-0.32	-0.42	-0.03	**-0.51**
	(0.03)	(0.32)	(0.38)	(0.61)	**(-2.84)**	(-3.61)	(-3.12)	(-3.42)	(-0.19)	**(-2.35)**
H-L	-0.45	-0.73	-0.97			-0.44	-0.67	-0.95		
	(-2.72)	(-4.82)	(-6.23)			(-2.50)	(-3.87)	(-4.87)		
	Panel C. Average returns: rank					Panel D. Five-factor alphas: rank				
EXVA	Low	Neutral	High	H-L	DiD	Low	Neutral	High	H-L	DiD
Low	0.39	0.73	1.09	0.71		0.01	0.29	0.59	0.58	
	(1.27)	(2.55)	(4.08)	(4.62)		(0.09)	(2.58)	(4.24)	(3.84)	
2	0.35	0.48	0.86	0.51		-0.08	0.01	0.39	0.47	
	(1.38)	(1.91)	(3.48)	(4.58)		(-0.80)	(0.09)	(4.10)	(4.07)	
3	0.34	0.50	0.68	0.34		-0.15	0.08	0.17	0.32	
	(1.25)	(2.16)	(2.90)	(2.45)		(-1.71)	(1.06)	(1.83)	(2.47)	
4	0.11	0.43	0.27	0.16		-0.36	0.01	-0.18	0.18	
	(0.45)	(1.82)	(1.15)	(1.33)		(-4.22)	(0.12)	(-1.98)	(1.45)	
High	-0.05	0.20	0.06	0.11	**-0.60**	-0.43	-0.20	-0.44	-0.01	**-0.58**
	(-0.16)	(0.75)	(0.23)	(0.74)	**(-3.40)**	(-3.70)	(-2.03)	(-3.60)	(-0.04)	**(-2.77)**
H-L	-0.43	-0.54	-1.03			-0.44	-0.49	-1.03		
	(-2.66)	(-3.90)	(-6.50)			(-2.62)	(-2.94)	(-5.36)		

Table 6. Fama-MacBeth regressions.

This table presents time-series averages of coefficients from monthly cross-sectional regressions. EXVA is the extrapolative value. AbVol is abnormal trading volume, calculated as the trading volume percentile at the end of each month over a stock's previous 50 trading days. $Ret(-1)$ is return in the last month, $Ret(-12,-1)$ is the cumulative return over the past year with a one-month gap, $Ret(-36,-12)$ is the cumulative return over the past three years with a one-year gap. LOGME is the log of a firm's market equity. LOGBM is the log of book-to-equity ratio. Beta is CAPM beta using monthly returns data with a five-year rolling window for regressions. IVol is idiosyncratic volatility defined as the standard deviation of the residuals from the Fama-French three-factor model using daily excess returns within a month with a minimum 15 nonmissing observations. Turnover is last month's turnover ratio and Illiq is the Amihud (2002) measure. Each month, all independent variables are winsorized at 1% and 99% and standardized to mean zero and standard deviation. The sample period is from July 1965 through December 2021. We use NYSE, Amex, and Nasdaq common stocks with a price of at least \$5 and at least 10 trading days over the last month. The t-statistics in parentheses are calculated based on Newey and West (1987) adjusted standard errors.

	(1)	(2)	(3)	(4)	(5)	(6)
AbVol	0.319		0.348	0.248		0.262
	(11.89)		(13.59)	(14.95)		(15.53)
EXVA		-0.327	-0.360		-0.301	-0.324
		(-8.57)	(-9.05)		(-11.79)	(-13.11)
EXVA x AbVol			**-0.085**			**-0.060**
			(-4.81)			**(-3.25)**
Ret(-1)				-0.468	-0.303	-0.280
				(-9.97)	(-6.13)	(-5.68)
Ret(-12,-1)				0.300	0.300	0.308
				(6.36)	(6.37)	(6.56)
Ret(-36,-12)				-0.037	-0.039	-0.038
				(-1.26)	(-1.33)	(-1.31)
LOGME				-0.190	-0.141	-0.187
				(-4.34)	(-3.30)	(-4.24)
LOGBM				0.083	0.092	0.079
				(2.37)	(2.66)	(2.29)
Beta				0.054	0.055	0.060
				(0.93)	(0.98)	(1.06)
Ivol				-0.374	-0.298	-0.351
				(-10.96)	(-9.08)	(-10.51)
Turnover				0.051	0.053	0.044
				(1.31)	(1.38)	(1.13)
Illiq				0.131	0.068	0.130
				(3.64)	(2.48)	(3.67)
Intercept	1.097	1.082	1.100	1.223	1.175	1.223
	(4.96)	(5.20)	(5.00)	(5.77)	(5.64)	(5.76)

Table 7. Fama-MacBeth regressions: control for intersections with other predictors.

This table reports the time-series averages of the regression coefficients from Fama-MacBeth regressions with additional intersections. We intersect the abnormal volume with firm size (LOGME), idiosyncratic volatility (IVol), shocks to idiosyncratic volatility (UIVol), illiquidity (Illiq), and shocks to illiquidity (UIlliq), respectively. UIVol is defined as the difference between current-month idiosyncratic volatility and the average value of the previous 12 months. UIlliq is analogously defined. Every month, we run a cross-sectional regression of returns on lagged intersection terms and other control variables. Variable definitions and sample period are the same as in Table 6. The t-statistics are calculated based on Newey and West (1987) adjusted standard errors and reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
AbVol	0.308	0.287	0.274	0.274	0.273	0.278
	(17.34)	(16.08)	(14.97)	(13.57)	(13.51)	(12.19)
EXVA	-0.303	-0.301	-0.306	-0.301	-0.300	-0.306
	(-12.47)	(-12.36)	(-13.01)	(-12.45)	(-12.30)	(-12.89)
EXVA x AbVol	**-0.046**	**-0.056**	**-0.045**	**-0.047**	**-0.046**	**-0.043**
	(-2.77)	**(-3.37)**	**(-2.70)**	**(-2.80)**	**(-2.76)**	**(-2.64)**
LOGME x AbVol	-0.137					-0.120
	(-10.20)					(-7.14)
IVol x AbVol		0.107				0.038
		(5.84)				(2.04)
UIVol			0.114			0.112
			(2.27)			(2.16)
IVol x AbVol			-0.100			-0.032
			(-5.85)			(-1.68)
Illiq x AbVol				0.063		-0.085
				(2.26)		(-1.30)
UIlliq					-0.045	-0.040
					(-0.41)	(-0.34)
UIlliq x AbVol					-0.065	-0.014
					(-2.16)	(-0.22)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Intercept	1.236	1.235	1.247	1.242	1.260	1.264
	(5.84)	(5.85)	(5.97)	(5.85)	(5.95)	(6.02)

Table 8. Predictive regressions.
This table shows the predictive regressions of the high-volume premium (HVP) portfolios on lagged DOX. Panel A presents the results of regressions using DOX and controlling for contemporaneous asset pricing factors (Fama and French, 2015). Panel B presents the results that further controls for lagged investor sentiment index. HVP1 (HVP5) is the difference in future returns between low and high abnormal volume stocks within the lowest (highest) EXVA quintile and HVPDIFF is the difference between HVP1 and HVP5. DOX is standardized to have a standard deviation of one. The sample period is from June 1974 to December, 2019. All t-statistics are based on Newey and West (1987) adjusted standard errors.

	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Lagged DOX									
LDOX	**0.460**	-0.112	**0.572**	**0.407**	-0.147	**0.554**	**0.361**	-0.219	**0.580**
	(2.13)	(-0.56)	**(2.64)**	**(2.03)**	(-0.77)	**(2.53)**	**(1.99)**	(-1.25)	**(2.49)**
Mkt-RF				-0.148	-0.097	-0.051	-0.082	-0.045	-0.037
				(-2.85)	(-2.20)	(-1.07)	(-1.63)	(-0.98)	(-0.69)
SMB							0.038	0.109	-0.070
							(0.47)	(2.10)	(-0.85)
HML							0.013	-0.013	0.026
							(0.12)	(-0.16)	(0.22)
RMW							0.390	0.227	0.163
							(3.10)	(2.69)	(1.16)
CMA							0.122	0.256	-0.134
							(0.58)	(2.06)	(-0.53)
Intercept	0.006	-0.000	0.006	0.007	0.000	0.007	0.005	-0.002	0.007
	(3.52)	(-0.24)	(3.20)	(4.30)	(0.13)	(3.44)	(2.80)	(-0.97)	(2.69)
R^2 (%)	1.075	-0.063	1.287	3.527	1.558	1.345	7.172	4.409	1.912
N	547	547	547	547	547	547	547	547	547
Panel B. Control for investor sentiment									
LDOX	**0.415**	-0.149	**0.564**	**0.372**	-0.176	**0.548**	**0.365**	-0.231	**0.596**
	(2.11)	(-0.82)	**(2.77)**	**(1.98)**	(-0.99)	**(2.68)**	**(2.14)**	(-1.37)	**(2.82)**
LSENT	0.154	0.136	0.017	0.123	0.117	0.006	-0.054	0.043	-0.097
	(0.68)	(0.53)	(0.07)	(0.56)	(0.47)	(0.02)	(-0.23)	(0.19)	(-0.34)
Mkt-RF				-0.143	-0.089	-0.054	-0.076	-0.037	-0.039
				(-2.75)	(-2.07)	(-1.11)	(-1.47)	(-0.81)	(-0.71)
SMB							0.035	0.112	-0.077
							(0.43)	(2.12)	(-0.90)
HML							0.023	0.008	0.016
							(0.20)	(0.09)	(0.13)
RMW							0.392	0.225	0.167
							(2.90)	(2.76)	(1.14)
CMA							0.117	0.242	-0.125
							(0.55)	(1.93)	(-0.48)
Intercept	0.006	-0.000	0.006	0.007	0.000	0.007	0.005	-0.002	0.007
	(3.55)	(-0.23)	(3.21)	(4.34)	(0.11)	(3.47)	(2.83)	(-1.00)	(2.72)
R^2 (%)	0.928	-0.113	1.067	3.208	1.24	1.145	6.871	4.163	1.743
N	535	535	535	535	535	535	535	535	535

Table 9. Fama-MacBeth regressions: control for stock visibility.
This table reports the time-series averages of the regression coefficients from Fama-MacBeth regressions that control for stock visibility. We consider proxies of analyst coverage (AC), membership in Standard & Poor's 500 index (SP500), the change in breadth of institutional ownership (DB), and institutional ownership (INST). Every month, we run a cross-sectional regression of returns on lagged intersection terms and other control variables. Definitions of control variables and sample period are the same as in Table 6. The t-statistics are calculated based on Newey and West (1987) adjusted standard errors and reported in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
AbVol	0.301	0.360	0.294	0.289	0.268	0.292	0.224	0.242	0.260
	(14.00)	(14.98)	(13.33)	(12.79)	(14.51)	(16.43)	(12.69)	(13.65)	(12.73)
EXVA	-0.298	-0.356	-0.248	-0.252	-0.249	-0.300	-0.204	-0.208	-0.218
	(-7.04)	(-9.51)	(-5.68)	(-5.73)	(-9.97)	(-12.52)	(-7.40)	(-7.76)	(-8.33)
EXVA x AbVol	**-0.086**	**-0.074**	**-0.080**	**-0.083**	**-0.060**	**-0.048**	**-0.061**	**-0.063**	**-0.062**
	(-4.92)	**(-4.41)**	**(-3.98)**	**(-4.22)**	**(-3.54)**	**(-2.84)**	**(-3.20)**	**(-3.38)**	**(-3.37)**
AC	-0.015				0.112				0.123
	(-0.32)				(3.42)				(3.85)
AC x AbVol	-0.089				-0.115				-0.095
	(-5.67)				(-7.46)				(-4.71)
SP500		-0.047				0.039			0.065
		(-0.47)				(0.73)			(0.98)
SP500 x AbVol		-0.120				-0.113			0.001
		(-4.34)				(-4.92)			(0.04)
DB			0.143				0.042		0.039
			(3.50)				(1.89)		(1.80)
DB x AbVol			-0.020				-0.024		-0.009
			(-1.60)				(-2.02)		(-0.80)
INST				0.111				0.075	0.039
				(2.55)				(2.17)	(1.18)
INST x AbVol				-0.081				-0.084	-0.027
				(-5.52)				(-5.62)	(-1.77)
Controls	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Intercept	1.139	1.135	1.117	1.139	1.242	1.234	1.230	1.256	1.158
	(4.95)	(4.94)	(4.74)	(4.85)	(5.71)	(5.82)	(5.52)	(5.63)	(5.17)

Table 10. Double sort on past performance and abnormal volume.
The table reports average returns and alphas for portfolios formed by sorting stocks on equal-weighted historical returns and abnormal trading volume. Panel A shows the conditional sorts using one-month return. Panel B presents result using reference-period return. We form value-weighted portfolios and hold for one month. The five-factor alphas are computed with respect to the five-factor model of Fama and French (2015). The sample period is from July 1965 through December 2021. The t-statistics are calculated based on Newey and West (1987) adjusted standard errors and reported in parentheses.

	Abnormal volume quintiles									
Performance quintiles	Panel A. Average returns: one-month performance					Panel B. Five-factor alphas: one-month performance				
	Low	3	High	H-L	DiD	Low	3	High	H-L	DiD
Low	0.25	0.65	1.09	0.84		-0.45	-0.05	0.35	0.80	
	(0.90)	(2.72)	(4.65)	(4.59)		(-3.63)	(-0.39)	(1.91)	(4.15)	
2	0.38	0.73	1.07	0.69		-0.35	0.05	0.39	0.74	
	(1.80)	(3.76)	(5.28)	(4.93)		(-4.78)	(0.53)	(4.54)	(6.08)	
3	0.51	0.71	0.65	0.14		-0.06	0.09	-0.01	0.06	
	(2.68)	(4.10)	(3.41)	(1.08)		(-0.82)	(1.14)	(-0.08)	(0.46)	
4	0.52	0.55	0.68	0.16		-0.03	-0.03	0.08	0.11	
	(2.59)	(2.89)	(3.73)	(1.47)		(-0.44)	(-0.32)	(0.99)	(1.02)	
High	0.33	0.46	0.50	0.17	**-0.67**	-0.20	-0.11	-0.08	0.11	**-0.69**
	(1.36)	(2.12)	(2.70)	(1.07)	**(-3.37)**	(-1.92)	(-1.22)	(-0.83)	(0.81)	**(-2.86)**
H-L	0.08	-0.20	-0.60			0.26	-0.05	-0.43		
	(0.39)	(-1.25)	(-3.55)			(1.40)	(-0.31)	(-1.86)		

Performance quintiles	Panel C. Average returns: reference-period performance					Panel D. Five-factor alphas: reference-period performance				
	Low	3	High	H-L	DiD	Low	3	High	H-L	DiD
Low	0.10	0.58	1.01	0.91		-0.60	-0.07	0.31	0.91	
	(0.34)	(2.21)	(4.10)	(5.22)		(-3.76)	(-0.53)	(1.77)	(5.41)	
2	0.46	0.69	0.97	0.52		-0.21	0.01	0.22	0.43	
	(2.17)	(3.57)	(4.91)	(3.71)		(-2.18)	(0.09)	(1.89)	(3.67)	
3	0.46	0.66	0.85	0.39		-0.16	0.02	0.16	0.33	
	(2.33)	(3.63)	(4.61)	(3.41)		(-1.87)	(0.25)	(2.06)	(2.78)	
4	0.49	0.65	0.48	-0.01		-0.21	0.06	-0.17	0.03	
	(2.52)	(3.95)	(2.63)	(-0.05)		(-2.25)	(0.82)	(-2.10)	(0.28)	
High	0.66	0.55	0.55	-0.11	**-1.02**	0.11	-0.02	-0.02	-0.13	**-1.04**
	(2.83)	(2.55)	(2.95)	(-0.78)	**(-5.44)**	(0.73)	(-0.20)	(-0.21)	(-0.95)	**(-4.75)**
H-L	0.56	-0.04	-0.46			0.71	0.05	-0.33		
	(2.66)	(-0.20)	(-2.52)			(2.60)	(0.26)	(-1.36)		

Table 11. Predictive regressions using portfolios by equal-weighted performance variables.

This table shows the predictive regressions of the high-volume premium (HVP) portfolios on lagged DOX. HVP1 (HVP5) is the difference in future returns between low and high abnormal volume stocks within the lowest (highest) equal-weighted performance variable and HVPDIFF is the difference between HVP1 and HVP5. Panel A presents the results using one-month return and Panel B presents the results using reference-period return. DOX is standardized to have a standard deviation of one. The sample period is from June 1974 to December, 2019. All t-statistics are based on Newey and West (1987) adjusted standard errors.

	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. One-month performance									
LDOX	0.213	0.301	-0.088	0.166	0.242	-0.076	0.132	0.208	-0.076
	(0.93)	(1.50)	(-0.38)	(0.78)	(1.22)	(-0.31)	(0.63)	(1.09)	(-0.30)
LSENT				0.090	0.152	-0.062	0.066	0.128	-0.062
				(0.37)	(0.67)	(-0.26)	(0.28)	(0.59)	(-0.26)
Mkt-RF							-0.112	-0.112	0.000
							(-2.30)	(-2.54)	(0.00)
Intercept	0.008	0.002	0.006	0.008	0.002	0.006	0.009	0.003	0.006
	(4.49)	(1.02)	(3.00)	(4.71)	(1.17)	(2.98)	(5.27)	(1.67)	(2.98)
R^2 (%)	0.090	0.510	-0.155	-0.130	0.340	-0.336	1.220	2.107	-0.525
N	547	547	547	535	535	535	535	535	535
Panel B. Reference-period performance									
LDOX	0.072	0.103	-0.031	-0.013	0.133	-0.145	-0.033	0.103	-0.135
	(0.30)	(0.63)	(-0.14)	(-0.06)	(0.82)	(-0.65)	(-0.15)	(0.66)	(-0.61)
LSENT				0.247	-0.003	0.250	0.233	-0.023	0.257
				(1.03)	(-0.01)	(1.48)	(1.00)	(-0.13)	(1.52)
Mkt-RF							-0.066	-0.098	0.032
							(-1.45)	(-2.52)	(0.59)
Intercept	0.009	-0.001	0.009	0.009	-0.001	0.010	0.009	-0.000	0.010
	(4.76)	(-0.61)	(4.65)	(5.02)	(-0.82)	(5.12)	(5.29)	(-0.36)	(4.78)
R^2 (%)	-0.149	-0.087	-0.180	-0.074	-0.211	-0.149	0.325	1.447	-0.246
N	547	547	547	535	535	535	535	535	535

Table 12. Double-sorted portfolio: equal-weighted and gross-return weighted. The table reports average returns and alphas for portfolios based on EXVA and abnormal trading volume percentile. At the beginning of each month, we first separate stocks into five groups based on extrapolative value. Then, within each group, stocks are sorted into quintile portfolios using the abnormal volume percentile. We calculate average excess returns and five-factor (Fama and French, 2015) alphas. Panels A and B are results using equal-weighted portfolios and Panels C and D are the results using gross return-weighted potfolios. The sample period is from July 1965 through December 2021. The t-statistics are calculated based on Newey and West (1987) adjusted standard errors and reported in parentheses.

	Abnormal volume portfolios									
EXVA	Panel A. Average returns: equal-weighted					Panel A. Five-factor alphas: equal-weighted				
	Low	3	High	H-L	DiD	Low	3	High	H-L	DiD
Low	0.22	1.08	1.65	1.43		-0.40	0.45	0.96	1.36	
	(0.68)	(3.31)	(5.28)	(11.08)		(-3.90)	(4.96)	(8.65)	(11.12)	
2	0.25	0.75	1.30	1.05		-0.43	0.05	0.59	1.02	
	(0.87)	(2.68)	(4.70)	(10.63)		(-5.45)	(0.71)	(8.36)	(11.14)	
3	0.26	0.64	1.06	0.80		-0.39	-0.02	0.40	0.79	
	(0.93)	(2.40)	(4.02)	(8.55)		(-5.10)	(-0.34)	(5.75)	(9.77)	
4	0.09	0.51	0.89	0.80		-0.53	-0.13	0.20	0.73	
	(0.31)	(1.84)	(3.25)	(8.00)		(-7.05)	(-2.17)	(2.77)	(8.56)	
High	-0.32	0.09	0.17	0.49	**-0.94**	-0.90	-0.52	-0.43	0.47	**-0.89**
	(-1.02)	(0.28)	(0.59)	(5.72)	**(-7.33)**	(-9.93)	(-6.80)	(-4.93)	(5.33)	**(-6.48)**
H-L	-0.54	-1.00	-1.48			-0.50	-0.97	-1.39		
	(-4.66)	(-8.19)	(-10.55)			(-4.31)	(-7.88)	(-8.63)		
EXVA	Panel C. Average returns: gross return-weighted					Panel D. Five-factor alphas: gross return-weighted				
	Low	Neutral	High	H-L	DiD	Low	Neutral	High	H-L	DiD
Low	0.22	1.06	1.60	1.38		-0.39	0.45	0.91	1.31	
	(0.68)	(3.25)	(5.14)	(10.98)		(-4.17)	(5.00)	(8.62)	(11.23)	
2	0.25	0.74	1.27	1.03		-0.43	0.05	0.58	1.01	
	(0.88)	(2.65)	(4.64)	(10.48)		(-5.57)	(0.70)	(8.46)	(11.06)	
3	0.25	0.63	1.03	0.77		-0.39	-0.02	0.37	0.76	
	(0.91)	(2.38)	(3.89)	(8.38)		(-5.26)	(-0.36)	(5.50)	(9.65)	
4	0.08	0.51	0.86	0.79		-0.53	-0.13	0.19	0.72	
	(0.28)	(1.83)	(3.18)	(7.95)		(-7.20)	(-2.16)	(2.57)	(8.42)	
High	-0.34	0.07	0.08	0.42	**-0.96**	-0.91	-0.52	-0.49	0.41	**-0.89**
	(-1.08)	(0.22)	(0.27)	(4.43)	**(-7.49)**	(-10.18)	(-6.89)	(-5.00)	(4.49)	**(-6.61)**
H-L	-0.56	-1.00	-1.52			-0.51	-0.97	-1.41		
	(-4.84)	(-8.30)	(-10.76)			(-4.45)	(-7.92)	(-8.74)		

Table 13. Fama-MacBeth regressions using different λ values and alternative extrapolation measure.

This table presents the time-series averages of the regression coefficients from Fama-MacBeth regressions robustness tests. We recalculate EXVA using $\lambda = 0.60$ and $\lambda = 0.90$, respectively, and employ the price-path convexity (Convexity) from Gulen and Woeppel (2022) as an alternative extrapolation measure. The three extrapolation variables are calculated using a unified 49-day reference window as the EXVA in Table 6. Definitions of abnormal volume and other control variables, and sample period are the same as in Table 6. The t-statistics are calculated based on Newey and West (1987) adjusted standard errors and reported in parentheses.

	(1) $\lambda = 0.60$	(2) $\lambda = 0.90$	(3) Convexity
AbVol	0.265	0.248	0.251
	(15.77)	(14.68)	(14.85)
EXVA	-0.311	-0.244	-0.022
	(-13.04)	(-7.51)	(-0.79)
EXVA x AbVol	**-0.058**	**-0.094**	**-0.045**
	(-3.20)	**(-5.00)**	**(-2.73)**
Controls	Yes	Yes	Yes
Intercept	1.223	1.226	1.224
	(5.76)	(5.78)	(5.78)

Appendix to "The Cross-Section of Extrapolative Belief and the High-Volume Premium"

This appendix provides additional tests and discussions that pertain to the robustness of the main empirical results.

A.1. Extrapolative value and order imbalance

Order imbalance measures are calculated using data from TAQ. A buyer-initiated or seller-initiated trade is determined based on the Lee and Ready (1991) test. The resulting firm-day observations are provided by *Wharton Research Data Services* (WRDS) Millisecond Intraday Indicators.

At the last trading day of each month, three types of order imbalance are computed. Trade imbalance is calculated by the difference in the number of buys and number of sells divided by the total number of buys and sells. Volume imbalance is shares of buy trades minus shares of sell trades divided by the total volume of buys and sells. Dollar imbalance is the difference in the dollar value of buys and the dollar value of sells divided by the total dollar value of buys and sells. Note that the extrapolative value is calculated *without* the last trading day of the month, it thus can rule out the possibility that the effect of EXVA on order imbalances are driven by returns on the month-end day.

Figure A1 plots the order imbalances for stocks with different level of EXVA. At each month, I form 25 portfolios based on EXVA and calculate average order imbalances for each group. Order imbalances are positive for high-EXVA stocks, and decrease and turn to negative for low-EXVA stocks. Overall, the order imbalances tend to increase with respect to EXVA.

In regression analysis, I control for a list of firm characteristics. It consists of market value, book-to-market ratio, the cumulative return of the month, the cumulative return over the previous 11 months, the cumulative return over the past two years with a one-year gap, market beta, idiosyncratic volatility, asset growth, return on equity, the illiquidity measure of Amihud (2002), turnover ratio, and the coskewness of Harvey and Siddique (2000). Table A1 reports the time-series averages of slopes from monthly cross-sectional regressions. In the first three columns, I find a one standard deviation increase in EXVA is associated with a 0.012 increase in trade imbalance and a 0.013 increase in volume imbalance and dollar imbalance. After controlling for other firm characteristics, columns (4) to (6) of Table A1 imply the coefficient of EXVA reduce by about one-half compared to the univariate case, but the estimate is still highly significant.

A.2. Additional predictive regressions

In section 4 of the paper, I show that HVPDIFF, the difference of the high-volume premium between low EXVA stocks and high EXVA stocks, can be positively predicted by market-wide extrapolation level (DOX). While the HVPDIFF in section 4 is calculated via conditional sorts on EXVA and AbVol, I repeat the regression using independently sorted portfolios and portfolios constructed by price-path convexity.

Table A2 shows that the predictive power of DOX is similar to my main analysis when applying independent sorts, though the magnitudes of the estimated coefficients are slightly smaller than the conditional-sort case. Table A3 and Table A4 report the predictive regressions with portfolios constructed by price-path convexity. On average, a one-standard deviation increase in DOX is associated with an increase in HVPDIFF of about 40-45 pbs. Consistent with previous results, the predictive power of DOX is positive and stronger for HVP1, and negative and less pronounced for HVP5.

A.3. Orthorgonalization with short-term reversal

The construction of EXVA mechanically generates a positive correlation with the popular short-term reversal (REV hereafter) anomaly. In this section, I address this concern by showing that (1) the pricing of EXVA and REV is different and EXVA relates to future cross-sectional returns more robustly than REV; (2) after adjusting for REV, the main empirical findings with EXVA remain valid.

First, I compare EXVA and REV by running monthly cross-sectional regressions of the two variables

$$EXVA_{i,t} = a_t + b_t REV_{i,t} + \varepsilon_{i,t}^{EXVA},$$

$$REV_{i,t} = a_t + b_t EXVA_{i,t} + \varepsilon_{i,t}^{REV},$$

and then examine the pricing of $\varepsilon_{i,t}^{EXVA}$ and $\varepsilon_{i,t}^{REV}$, separately. Table A5 presents the average returns and alphas of decile portfolios sorted by residual EXVA. On average, a strategy by longing lowest $\varepsilon_{i,t}^{EXVA}$ decile stocks and shorting highest $\varepsilon_{i,t}^{REV}$ decile stocks earns a model-adjusted return above 1% per month. In shape contrast, after adjusting for EXVA, the short-term reversal disappears and the pricing effect of $\varepsilon_{i,t}^{REV}$ even becomes significantly positive, as reported in Table A6.

Then, I replace EXVA with $\varepsilon_{i,t}^{EXVA}$ and tests its interaction with abnormal volume. In Table A7, I perform Fama-MacBeth regressions using $\varepsilon_{i,t}^{EXVA}$. The estimated coefficient of the intersection term is smaller but comparable to that of unadjusted EXVA. In addition, Table A8 suggests a similar mechanism through which $\varepsilon_{i,t}^{EXVA}$ affects the high-volume premium: DOX can still positively predict HVPDIFF from portfolios sorted by $\varepsilon_{i,t}^{EXVA}$ and AbVol. The only difference is that the negative predictive power of DOX for HVP5 is a bit stronger, although the statistical significance is marginal.

A.4. Further control for mispricing measures

Following Han et al. (2022), I compute two mispricing measures. The first one is the mispricing score (MISP) of Stambaugh et al. (2015), constructed as the average rank percentile over 11 anomalies. [15] High MISP implies overpricing and low MISP implies underpricing. The second measure is CAPM alpha (Alpha), which is estimated by time-series regression for each stock-month using past two-year observations. I multiply the estimated intercepts by -1 so that high (low) Alpha means overpricing (underpricing).

Table A9 presents the Fama-MacBeth regressions using the two mispricing measures. It is shown that the main conclusion of this paper is unchanged since the sign and magnitude of the EXVA-AbVol intersection term remain similar. The mispricing measures negatively predict future returns, which is consistent with prior literature. However, the slopes of mispricing-AbVol terms are positive, suggesting that this paper's story is different from Han et al. (2022).

[15]The set of anomalies include net stock issues, composite equity issues, operating accruals, net operating assets, asset growth, investment-to-assets, distress risk, O-score, momentum, gross profitability, and return on assets. I follow the same procedures of Stambaugh et al. (2015) to calculate these anomaly variables.

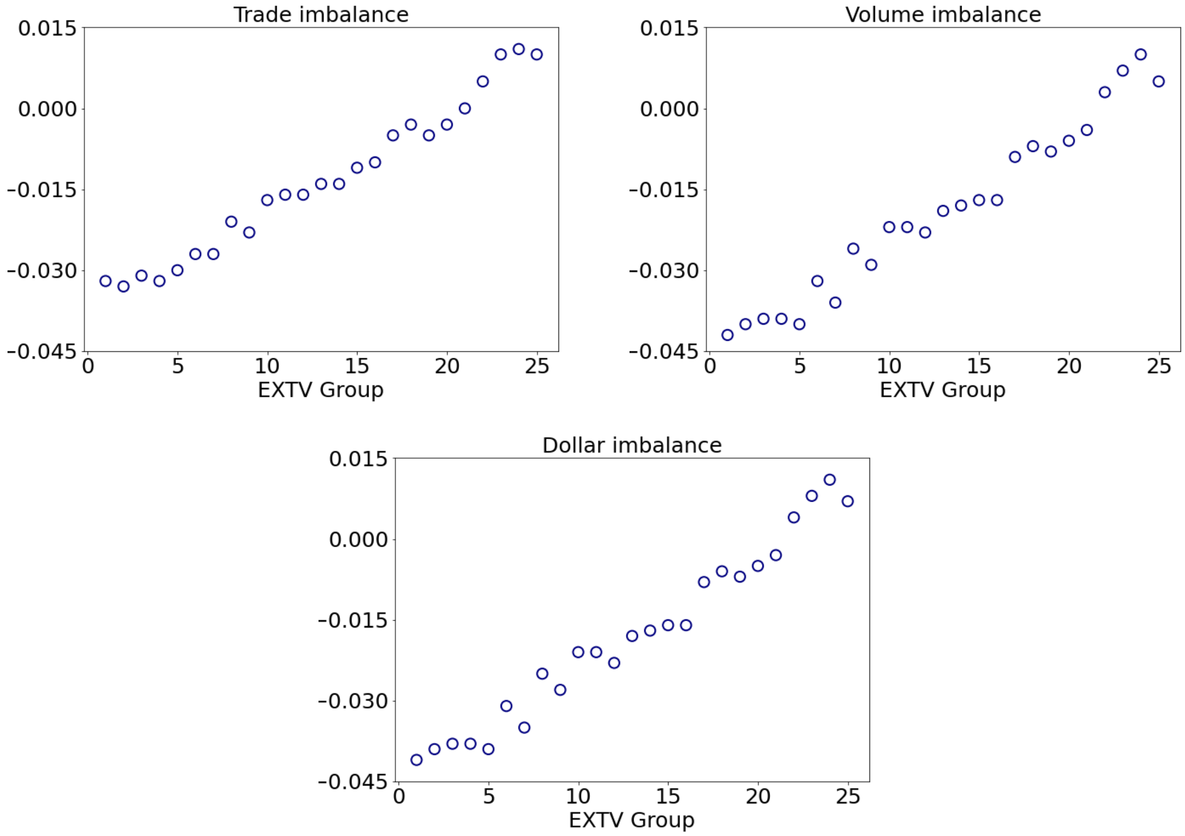

Figure A1. EXVA portfolios and order imbalances. This figure plots the average order imbalances for 25 EXVA-sorted portfolios. Each month, stocks are separated into 25 groups based on extrapolative value, then averaged month-end order imbalances are calculated for each portfolio.

Table A1. Extrapolative value and order imbalance.
This table presents time-series averages of coefficients from monthly cross-sectional regressions of order imbalance on extrapolative value (EXVA). Trade imbalance is calculated by the difference in the number of buys and number of sells divided by the total number of buys and sells. Volume imbalance is shares of buy trades minus shares of sell trades divided by the total volume of buys and sells. Dollar imbalance is the difference in dollar value of buys and dollar value of sells divided by the total dollar value of buys and sells. Order imbalance variables are measured for each month-end day. The control variables include firm size, book-to-market ratio, return in the current month, momentum ($Ret(-11,-1)$), long-term return performance ($Ret(-36,-12)$), CAPM beta, idiosyncratic volatility, asset growth, return on equity, illiquidity (Amihud, 2002), turnover ratio, and coskewness (Harvey and Siddique, 2000). Each month, all independent variables are winsorized at 1% and 99% and standardized to mean zero and standard deviation. Common stocks with a price of at least $5 and at least 10 trading days over the month are used. The t-statistics in parentheses are calculated based on Newey and West (1987) adjusted standard errors.

	Trade	Volume	Dollar	Trade	Volume	Dollar
	(1)	(2)	(3)	(4)	(5)	(6)
EXVA	0.012	0.013	0.013	0.006	0.007	0.007
	(19.02)	(18.36)	(18.40)	(7.16)	(8.09)	(8.09)
Controls	No	No	No	Yes	Yes	Yes
Intercept	-0.014	-0.019	-0.018	-0.014	-0.019	-0.018
	(-6.53)	(-10.37)	(-9.82)	(-6.85)	(-10.16)	(-9.61)
R^2(%)	0.374	0.341	0.342	2.599	2.040	2.015

Table A2. Predictive regressions: independent sorts by EXVA and AbVol.
This table shows the predictive regressions of the high-volume premium (HVP) portfolios on lagged DOX. HVP1 (HVP5) is the difference in future returns between low and high abnormal volume stocks within the lowest (highest) EXVA quintile and HVPDIFF is the difference between HVP1 and HVP5. Portfolios are constructed by independently sorting stocks based on EXVA convexity and abnormal volume. DOX is standardized to have a standard deviation of one. The sample period is from June 1974 to December, 2019. All t-statistics are based on Newey and West (1987) adjusted standard errors.

	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. DOX									
LDOX	**0.411**	-0.110	**0.521**	**0.374**	-0.137	**0.511**	**0.352**	-0.214	**0.566**
	(1.81)	(-0.60)	**(2.97)**	**(1.74)**	(-0.79)	**(2.88)**	**(1.80)**	(-1.40)	**(2.85)**
Mkt-RF				-0.104	-0.077	-0.027	-0.036	-0.027	-0.010
				(-1.95)	(-1.59)	(-0.57)	(-0.68)	(-0.58)	(-0.18)
SMB							-0.024	0.108	-0.131
							(-0.29)	(1.71)	(-1.73)
HML							0.013	0.107	-0.094
							(0.13)	(1.33)	(-1.04)
RMW							0.415	0.200	0.215
							(3.64)	(2.15)	(1.38)
CMA							0.050	0.146	-0.096
							(0.25)	(1.07)	(-0.39)
Intercept	0.007	0.000	0.007	0.008	0.001	0.008	0.006	-0.001	0.008
	(4.26)	(0.01)	(3.83)	(4.77)	(0.30)	(3.93)	(3.59)	(-0.76)	(3.18)
R^2 (%)	0.833	-0.073	1.038	1.962	0.825	0.924	6.821	4.038	3.035
N	547	547	547	547	547	547	547	547	547
Panel B. Control for investor sentiment									
LDOX	0.333	-0.140	**0.473**	0.304	-0.162	**0.466**	**0.327**	-0.218	**0.545**
	(1.67)	(-0.81)	**(2.71)**	(1.56)	(-0.96)	**(2.64)**	**(1.84)**	(-1.41)	**(2.93)**
LSENT	0.304	0.072	0.232	0.284	0.057	0.227	0.078	-0.020	0.098
	(1.21)	(0.29)	(0.96)	(1.17)	(0.23)	(0.94)	(0.30)	(-0.09)	(0.35)
Mkt-RF				-0.094	-0.071	-0.023	-0.026	-0.019	-0.007
				(-1.78)	(-1.48)	(-0.47)	(-0.48)	(-0.41)	(-0.13)
SMB							-0.026	0.107	-0.134
							(-0.33)	(1.71)	(-1.74)
HML							0.033	0.120	-0.087
							(0.30)	(1.46)	(-0.92)
RMW							0.410	0.204	0.206
							(3.35)	(2.18)	(1.24)
CMA							0.033	0.139	-0.106
							(0.16)	(1.01)	(-0.42)
Intercept	0.007	0.000	0.007	0.008	0.001	0.007	0.006	-0.001	0.008
	(4.28)	(0.09)	(3.73)	(4.81)	(0.35)	(3.80)	(3.63)	(-0.72)	(3.15)
R^2 (%)	0.996	-0.2	1.037	1.883	0.537	0.898	6.682	3.898	2.88
N	535	535	535	535	535	535	535	535	535

Table A3. Predictive regressions using portfolios by price-path convexity: conditional sorts.

This table shows the predictive regressions of the high-volume premium (HVP) portfolios on lagged DOX. HVP1 (HVP5) is the difference in future returns between low and high abnormal volume stocks within the lowest (highest) price-path convexity (Gulen and Woeppel, 2022) and HVPDIFF is the difference between HVP1 and HVP5. Portfolios are constructed by sequentially sorting stocks based on price-path convexity and abnormal volume. DOX is standardized to have a standard deviation of one. The sample period is from June 1974 to December, 2019. All t-statistics are based on Newey and West (1987) adjusted standard errors.

	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. DOX									
LDOX	**0.524**	0.095	**0.429**	**0.482**	0.058	**0.424**	**0.475**	-0.016	**0.490**
	(2.53)	(0.52)	**(2.01)**	**(2.41)**	(0.33)	**(1.96)**	**(2.35)**	(-0.10)	**(2.15)**
Mkt-RF				-0.119	-0.103	-0.015	-0.075	-0.062	-0.013
				(-2.32)	(-1.80)	(-0.25)	(-1.29)	(-1.09)	(-0.18)
SMB							-0.059	0.159	-0.218
							(-0.63)	(2.76)	(-2.32)
HML							-0.029	-0.020	-0.009
							(-0.22)	(-0.22)	(-0.06)
RMW							0.165	0.290	-0.125
							(1.49)	(3.17)	(-0.84)
CMA							0.106	0.203	-0.097
							(0.47)	(1.41)	(-0.35)
Intercept	0.006	0.001	0.005	0.006	0.002	0.005	0.005	-0.000	0.006
	(3.07)	(0.55)	(2.36)	(3.63)	(0.98)	(2.38)	(2.52)	(-0.20)	(2.33)
R^2 (%)	1.468	-0.119	0.559	2.992	1.238	0.395	3.648	3.709	1.058
N	547	547	547	547	547	547	547	547	547
Panel B. Control for investor sentiment									
LDOX	**0.528**	0.077	**0.452**	**0.492**	0.048	**0.445**	**0.503**	-0.004	**0.507**
	(2.56)	(0.39)	**(2.14)**	**(2.43)**	(0.25)	**(2.09)**	**(2.52)**	(-0.02)	**(2.30)**
LSENT	-0.003	-0.032	0.030	-0.028	-0.053	0.025	-0.133	-0.141	0.008
	(-0.01)	(-0.15)	(0.13)	(-0.09)	(-0.26)	(0.10)	(-0.43)	(-0.74)	(0.03)
Mkt-RF				-0.119	-0.096	-0.023	-0.072	-0.052	-0.020
				(-2.33)	(-1.68)	(-0.37)	(-1.22)	(-0.91)	(-0.29)
SMB							-0.052	0.154	-0.206
							(-0.56)	(2.65)	(-2.19)
HML							-0.014	0.002	-0.016
							(-0.11)	(0.02)	(-0.12)
RMW							0.182	0.295	-0.113
							(1.52)	(3.11)	(-0.72)
CMA							0.106	0.187	-0.081
							(0.47)	(1.29)	(-0.29)
Intercept	0.005	0.001	0.004	0.006	0.002	0.004	0.005	-0.000	0.005
	(2.97)	(0.72)	(2.18)	(3.53)	(1.13)	(2.23)	(2.47)	(-0.06)	(2.23)
R^2 (%)	1.322	-0.335	0.49	2.879	0.833	0.346	3.679	3.402	0.852
N	535	535	535	535	535	535	535	535	535

Table A4. Predictive regressions using portfolios by price-path convexity: independent sorts.

This table shows the predictive regressions of the high-volume premium (HVP) portfolios on lagged DOX. HVP1 (HVP5) is the difference in future returns between low and high abnormal volume stocks within the lowest (highest) price-path convexity (Gulen and Woeppel, 2022) and HVPDIFF is the difference between HVP1 and HVP5. Portfolios are constructed by independently sorting stocks based on price-path convexity and abnormal volume. DOX is standardized to have a standard deviation of one. The sample period is from June 1974 to December, 2019. All t-statistics are based on Newey and West (1987) adjusted standard errors.

	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. DOX									
LDOX	**0.387**	-0.033	**0.420**	**0.372**	-0.054	**0.426**	**0.356**	-0.081	**0.436**
	(2.17)	(-0.23)	**(2.19)**	**(2.11)**	(-0.38)	**(2.18)**	**(2.00)**	(-0.59)	**(2.08)**
Mkt-RF				-0.043	-0.059	0.017	-0.010	-0.038	0.028
				(-1.08)	(-1.47)	(0.33)	(-0.19)	(-0.88)	(0.47)
SMB							-0.026	0.048	-0.075
							(-0.34)	(0.89)	(-0.92)
HML							-0.013	-0.001	-0.013
							(-0.14)	(-0.01)	(-0.12)
RMW							0.082	0.144	-0.062
							(0.77)	(1.69)	(-0.50)
CMA							0.121	0.068	0.053
							(0.63)	(0.46)	(0.22)
Intercept	0.005	0.001	0.004	0.005	0.001	0.004	0.004	0.000	0.004
	(2.90)	(0.36)	(2.12)	(3.13)	(0.65)	(2.05)	(2.17)	(0.05)	(1.68)
R^2 (%)	0.898	-0.175	0.59	0.98	0.234	0.431	0.856	0.28	-0.099
N	547	547	547	547	547	547	547	547	547
Panel B. Control for investor sentiment									
LDOX	**0.341**	-0.043	**0.384**	**0.329**	-0.060	**0.388**	**0.324**	-0.072	**0.397**
	(2.00)	(-0.30)	**(2.07)**	**(1.92)**	(-0.42)	**(2.07)**	**(1.91)**	(-0.51)	**(2.03)**
LSENT	0.118	-0.028	0.146	0.110	-0.040	0.149	0.061	-0.097	0.157
	(0.46)	(-0.20)	(0.59)	(0.42)	(-0.29)	(0.60)	(0.23)	(-0.71)	(0.59)
Mkt-RF				-0.041	-0.054	0.013	-0.009	-0.031	0.022
				(-1.03)	(-1.35)	(0.26)	(-0.17)	(-0.72)	(0.37)
SMB							-0.033	0.040	-0.073
							(-0.42)	(0.73)	(-0.89)
HML							-0.006	0.020	-0.026
							(-0.07)	(0.26)	(-0.25)
RMW							0.071	0.144	-0.072
							(0.64)	(1.62)	(-0.55)
CMA							0.113	0.050	0.063
							(0.58)	(0.33)	(0.26)
Intercept	0.005	0.001	0.004	0.005	0.001	0.004	0.004	0.000	0.004
	(2.90)	(0.50)	(2.04)	(3.12)	(0.76)	(1.98)	(2.23)	(0.17)	(1.65)
R^2 (%)	0.676	-0.35	0.45	0.736	-0.033	0.278	0.566	0.045	-0.262
N	535	535	535	535	535	535	535	535	535

Table A5. Returns and alphas of portfolios sorted by residual extrapolative value.

This table presents the average returns and alphas on value-weighted decile portfolios of stocks sorted on residual extrapolative value. The residual extrapolative value is calculated from month-by-month cross-sectional regression:

$$EXVA_{i,t} = a_t + b_t REV_{i,t} + \varepsilon_{i,t}^{EXVA},$$

where EXVA is the extrapolative value defined in Table 1 and REV is the contemporaneous one-month stock return. Variables are winsorized at 5% and 95% level for each regression. At the end of each month t, stocks are sorted into decile portfolios by $\varepsilon_{i,t}^{EXVA}$ and model-adjusted alphas are calculated. We consider CAPM, three-factor model, five-factor model (Fama and French, 2015), the momentum-augmented model (FF5+MOM), short-term reversal-augmented model (FF5+REV), the four-factor model (Q4) of Hou et al. (2015), the five-factor model (Q5) of Hou et al. (2021), and the behavioral factor model (DHS) of Daniel et al. (2020). The t-statistics in parentheses are calculated based on Newey and West (1987) adjusted standard errors.

	Low	2	3	4	5	6	7	8	9	High	High-Low
Excess return	0.851	0.714	0.528	0.522	0.471	0.477	0.268	0.267	0.205	-0.096	**-0.948**
	(2.95)	(2.74)	(2.20)	(2.25)	(2.02)	(2.13)	(1.16)	(1.12)	(0.84)	(-0.34)	**(-5.50)**
CAPM	0.16	0.11	-0.03	0.02	-0.03	-0.01	-0.23	-0.27	-0.36	-0.77	**-0.93**
	(0.81)	(0.70)	(-0.21)	(0.13)	(-0.23)	(-0.08)	(-1.56)	(-1.76)	(-2.42)	(-4.41)	**(-5.29)**
FF3	0.15	0.12	-0.05	-0.01	-0.06	-0.04	-0.26	-0.27	-0.36	-0.78	**-0.93**
	(0.86)	(0.79)	(-0.39)	(-0.07)	(-0.43)	(-0.27)	(-1.91)	(-2.00)	(-2.67)	(-4.86)	**(-5.21)**
FF5	0.49	0.31	0.09	0.11	0.04	0.05	-0.20	-0.19	-0.24	-0.60	**-1.08**
	(4.25)	(3.70)	(1.31)	(1.64)	(0.73)	(0.93)	(-3.00)	(-2.85)	(-2.71)	(-4.82)	**(-5.72)**
FF5+MOM	0.53	0.31	0.08	0.08	0.03	0.02	-0.23	-0.22	-0.26	-0.58	**-1.11**
	(4.53)	(3.57)	(1.18)	(1.19)	(0.49)	(0.39)	(-3.62)	(-3.12)	(-3.14)	(-4.89)	**(-5.74)**
FF5+REV	0.49	0.32	0.10	0.09	0.02	0.03	-0.21	-0.20	-0.24	-0.62	**-1.11**
	(4.29)	(3.82)	(1.46)	(1.50)	(0.33)	(0.58)	(-3.05)	(-2.91)	(-2.64)	(-4.58)	**(-5.59)**
Q4	0.61	0.35	0.13	0.13	0.06	0.02	-0.19	-0.20	-0.24	-0.63	**-1.24**
	(4.79)	(3.53)	(1.71)	(1.48)	(0.95)	(0.35)	(-2.51)	(-2.52)	(-2.47)	(-4.56)	**(-5.95)**
Q5	0.64	0.29	0.08	0.06	0.01	-0.01	-0.25	-0.23	-0.24	-0.50	**-1.14**
	(4.70)	(2.98)	(0.95)	(0.63)	(0.11)	(-0.23)	(-3.27)	(-2.71)	(-2.68)	(-4.07)	**(-5.50)**
DHS	0.74	0.30	0.11	0.10	0.03	-0.00	-0.25	-0.24	-0.26	-0.40	**-1.14**
	(4.46)	(3.23)	(1.27)	(1.43)	(0.53)	(-0.07)	(-3.45)	(-2.97)	(-2.96)	(-2.92)	**(-4.59)**

Table A6. Returns and alphas of portfolios sorted by residual short-term reversal.

This table shows the average returns and alphas on value-weighted decile portfolios of stocks sorted on residual short-term reversal. The residual short-term reversal is calculated from month-by-month cross-sectional regression:

$$REV_{i,t} = a_t + b_t EXVA_{i,t} + \varepsilon_{i,t}^{REV},$$

where EXVA and REV follow the same definitions as in Table B1. Variables are winsorized at 5% and 95% level for each regression. At the end of each month t, stocks are sorted into decile portfolios by $\varepsilon_{i,t}^{REV}$ and model-adjusted alphas are calculated. We consider CAPM, three-factor model, five-factor model (Fama and French, 2015), the momentum-augmented model (FF5+MOM), short-term reversal-augmented model (FF5+REV), the four-factor model (Q4) of Hou et al. (2015), the five-factor model (Q5) of Hou et al. (2021), and the behavioral factor model (DHS) of Daniel et al. (2020). The t-statistics in parentheses are calculated based on Newey and West (1987) adjusted standard errors.

	Low	2	3	4	5	6	7	8	9	High	High-Low
Excess return	-0.049	0.433	0.570	0.474	0.521	0.433	0.434	0.402	0.499	0.513	**0.563**
	(-0.16)	(1.62)	(2.30)	(2.01)	(2.31)	(1.91)	(1.91)	(1.63)	(2.05)	(1.75)	**(2.86)**
CAPM	-0.80	-0.21	-0.00	-0.06	0.02	-0.07	-0.07	-0.13	-0.06	-0.10	**0.70**
	(-4.16)	(-1.30)	(-0.03)	(-0.41)	(0.15)	(-0.46)	(-0.46)	(-0.80)	(-0.41)	(-0.52)	**(3.62)**
FF3	-0.84	-0.25	-0.03	-0.09	-0.01	-0.10	-0.08	-0.14	-0.05	-0.08	**0.76**
	(-4.53)	(-1.66)	(-0.21)	(-0.64)	(-0.04)	(-0.75)	(-0.59)	(-0.95)	(-0.35)	(-0.44)	**(3.73)**
FF5	-0.60	-0.10	0.10	-0.01	0.07	0.01	0.03	0.04	0.12	0.17	**0.77**
	(-3.82)	(-1.10)	(1.20)	(-0.20)	(1.23)	(0.10)	(0.48)	(0.48)	(1.47)	(1.23)	**(2.91)**
FF5+MOM	-0.49	-0.06	0.09	-0.04	0.05	-0.01	-0.00	-0.01	0.05	0.06	**0.55**
	(-3.37)	(-0.71)	(1.15)	(-0.62)	(0.87)	(-0.19)	(-0.02)	(-0.18)	(0.62)	(0.46)	**(2.30)**
FF5+REV	-0.83	-0.29	-0.03	-0.10	0.04	-0.00	0.08	0.15	0.28	0.42	**1.25**
	(-7.25)	(-3.91)	(-0.45)	(-1.66)	(0.74)	(-0.08)	(1.32)	(2.00)	(4.05)	(3.78)	**(6.95)**
Q4	-0.57	-0.08	0.07	-0.02	0.08	0.03	0.03	0.01	0.11	0.19	**0.76**
	(-3.25)	(-0.72)	(0.82)	(-0.33)	(1.06)	(0.45)	(0.46)	(0.15)	(1.07)	(1.05)	**(2.43)**
Q5	-0.36	0.06	0.12	-0.03	0.04	0.01	-0.07	-0.02	-0.03	-0.04	0.32
	(-2.11)	(0.60)	(1.57)	(-0.44)	(0.46)	(0.14)	(-0.91)	(-0.23)	(-0.25)	(-0.24)	(1.10)
DHS	-0.14	0.20	0.23	0.10	0.06	-0.03	-0.04	-0.17	-0.04	-0.07	0.06
	(-0.84)	(1.78)	(2.70)	(1.27)	(0.89)	(-0.48)	(-0.61)	(-1.86)	(-0.43)	(-0.50)	(0.24)

Table A7. Fama-MacBeth regressions using residual extrapolative value.
This table presents time-series averages of coefficients from monthly cross-sectional regressions. RES EXVA is the residual extrapolative value, which is the residual of monthly cross-sectional regression of EXVA against REV. Other variables are identically defined as in Table 6. Each month, all independent variables are winsorized at 1% and 99% and standardized to mean zero and standard deviation. The sample period is from July 1965 through December 2021. We use NYSE, Amex, and Nasdaq common stocks with a price of at least $5 and at least 10 trading days over the last month. The t-statistics in parentheses are calculated based on Newey and West (1987) adjusted standard errors.

	(1)	(2)	(3)	(4)
AbVol		0.341		0.269
		(12.76)		(16.12)
RES EXVA	-0.276	-0.327	-0.315	-0.344
	(-7.48)	(-8.96)	(-14.22)	(-15.82)
RES EXVA x AbVol		**-0.098**		**-0.042**
		(-6.02)		**(-2.69)**
Ret(-1)			-0.446	-0.442
			(-9.39)	(-9.37)
Ret(-12,-1)			0.297	0.304
			(6.25)	(6.42)
Ret(-36,-12)			-0.039	-0.037
			(-1.32)	(-1.28)
LOGME			-0.140	-0.185
			(-3.29)	(-4.22)
LOGBM			0.091	0.078
			(2.63)	(2.25)
Beta			0.055	0.060
			(0.97)	(1.05)
Ivol			-0.298	-0.354
			(-9.13)	(-10.58)
Turnover			0.051	0.040
			(1.31)	(1.04)
Illiq			0.072	0.133
			(2.74)	(3.86)
Intercept	1.080	1.097	1.173	1.219
	(5.20)	(4.98)	(5.63)	(5.75)

Table A8. Predictive regressions using portfolios by residual extrapolative value. This table shows the predictive regressions of the high-volume premium (HVP) portfolios on lagged DOX. HVP1 (HVP5) is the difference in future returns between low and high abnormal volume stocks within the lowest (highest) residual extrapolative value and HVPDIFF is the difference between HVP1 and HVP5. The residual extrapolative value is computed as the residual of monthly cross-sectional regression of EXVA against REV. DOX is standardized to have a standard deviation of one. The sample period is from June 1974 to December, 2019. All t-statistics are based on Newey and West (1987) adjusted standard errors.

	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF	HVP1	HVP5	HVPDIFF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Lagged DOX									
LDOX	0.306	-0.194	**0.499**	0.281	-0.232	**0.513**	0.243	-0.313	**0.556**
	(1.44)	(-0.86)	**(1.93)**	(1.35)	(-1.07)	**(1.99)**	(1.18)	(-1.56)	**(2.10)**
Mkt-RF				-0.069	-0.108	0.039	-0.049	-0.011	-0.039
				(-1.49)	(-2.10)	(0.68)	(-0.96)	(-0.19)	(-0.56)
SMB							0.099	0.019	0.081
							(1.21)	(0.33)	(0.86)
HML							0.052	0.008	0.044
							(0.51)	(0.07)	(0.30)
RMW							0.216	0.293	-0.077
							(2.25)	(2.47)	(-0.49)
CMA							0.001	0.368	-0.367
							(0.01)	(2.33)	(-1.41)
Intercept	0.005	0.001	0.004	0.006	0.002	0.004	0.004	-0.001	0.005
	(3.10)	(0.52)	(1.93)	(3.56)	(0.86)	(1.73)	(2.37)	(-0.48)	(2.17)
R^2 (%)	0.418	0.095	0.792	0.849	1.661	0.731	1.538	7.003	1.729
N	547	547	547	547	547	547	547	547	547
Panel B. Control for investor sentiment									
LDOX	0.282	-0.269	**0.551**	0.262	-0.301	**0.563**	0.245	**-0.351**	**0.596**
	(1.38)	(-1.41)	**(2.34)**	(1.29)	(-1.62)	**(2.39)**	(1.25)	**(-2.01)**	**(2.54)**
LSENT	-0.019	0.329	-0.348	-0.033	0.307	-0.340	-0.109	0.170	-0.278
	(-0.07)	(1.21)	(-0.97)	(-0.12)	(1.14)	(-0.94)	(-0.37)	(0.68)	(-0.75)
Mkt-RF				-0.066	-0.106	0.040	-0.046	-0.009	-0.036
				(-1.42)	(-2.08)	(0.70)	(-0.88)	(-0.17)	(-0.52)
SMB							0.094	0.023	0.071
							(1.13)	(0.40)	(0.74)
HML							0.058	0.011	0.047
							(0.56)	(0.10)	(0.31)
RMW							0.222	0.278	-0.056
							(2.09)	(2.36)	(-0.35)
CMA							-0.013	0.378	-0.390
							(-0.06)	(2.37)	(-1.47)
Intercept	0.005	0.001	0.005	0.006	0.002	0.004	0.005	-0.001	0.006
	(3.23)	(0.47)	(2.14)	(3.69)	(0.81)	(1.92)	(2.49)	(-0.52)	(2.30)
R^2 (%)	0.124	0.481	0.842	0.502	1.989	0.784	1.21	7.115	1.795
N	535	535	535	535	535	535	535	535	535

Table A9. Fama-MacBeth regressions: controlling for mispricing measures. This table presents time-series averages of coefficients from monthly cross-sectional regressions. MISP is the mispricing score of Stambaugh et al. (2015), Alpha is the estimated CAPM alpha using a two-year rolling window for each stock multiplied by -1. Other variables are identically defined as in Table 6. Each month, all independent variables are winsorized at 1% and 99% and standardized to mean zero and standard deviation. The sample period is from July 1965 through December 2021. We use NYSE, Amex, and Nasdaq common stocks with a price of at least \$5 and at least 10 trading days over the last month. The t-statistics in parentheses are calculated based on Newey and West (1987) adjusted standard errors.

	(1)	(2)	(3)	(4)	(5)
AbVol	0.342	0.315	0.271	0.269	0.272
	(15.14)	(15.70)	(17.20)	(17.51)	(17.31)
EXVA	-0.382	-0.398	-0.303	-0.302	-0.304
	(-8.83)	(-9.73)	(-9.69)	(-9.76)	(-9.86)
EXVA x AbVol	**-0.072**	**-0.065**	**-0.046**	**-0.045**	**-0.044**
	(-4.23)	**(-3.81)**	**(-2.92)**	**(-2.82)**	**(-2.67)**
MISP	-0.459		-0.278		-0.280
	(-11.34)		(-12.88)		(-13.03)
MISP x AbVol	0.036		0.017		0.020
	(3.24)		(1.64)		(1.85)
Alpha		-0.146		0.052	0.072
		(-2.06)		(1.08)	(1.51)
Alpha x AbVol		0.026		0.006	0.005
		(1.66)		(0.41)	(0.31)
Ret(-1)			-0.305	-0.279	-0.297
			(-7.23)	(-6.89)	(-7.41)
Ret(-12,-1)			0.214	0.324	0.258
			(4.07)	(6.85)	(5.40)
Ret(-36,-12)			-0.005	-0.019	0.025
			(-0.16)	(-0.53)	(0.70)
LOGME			-0.217	-0.191	-0.217
			(-4.84)	(-4.24)	(-4.91)
LOGBM			0.091	0.067	0.086
			(2.24)	(1.66)	(2.16)
Beta			0.071	0.072	0.083
			(1.37)	(1.40)	(1.62)
Ivol			-0.312	-0.336	-0.303
			(-9.40)	(-10.16)	(-9.35)
Turnover			0.053	0.044	0.062
			(1.33)	(1.10)	(1.57)
Illiq			0.113	0.126	0.109
			(3.01)	(3.48)	(2.97)
Intercept	1.118	1.169	1.209	1.236	1.210
	(5.16)	(5.48)	(5.63)	(5.78)	(5.63)